Shape gradients, shape warping and medical application to facial expression analysis

Pierre Maurel

École Doctorale de Sciences Mathématiques de Paris Centre
Équipe Odyssée (ENS / INRIA / ENPC)

Doctorat

Informatique

Shape gradients, shape warping and medical application to facial expression analysis

Pierre MAUREL

Thèse dirigée par Olivier FAUGERAS

Soutenue le 02/12/2008

JURY

Patrick Chauvel	
Christian Choffrut	
Laurent Cohen	
Olivier Faugeras	Directeur
Renaud Keriven	Co-directeur
Nikos Paragios	Rapporteur
Gabriel Peyré	
Monique Thonnat	Rapporteur

Contents

Remerciements

Je tiens tout d'abord à remercier Olivier Faugeras et Renaud Keriven de m'avoir accueilli au sein du projet Odyssée ainsi que pour la qualité de leur encadrement. Un merci particulier à Renaud pour avoir été à l'origine de mon envie d'explorer le domaine de la vision par ordinateur à travers la qualité de son enseignement et sa pédagogie. Je remercie également à ce titre Guillermo Sapiro qui m'a accueilli en stage dans son laboratoire avant mon DEA ce qui a constitué ma première expérience de recherche.

J'adresse tout ma gratitude à Monique Thonnat et Nikos Paragios pour avoir accepté d'être mes rapporteurs, malgré le travail important que cela représente, ainsi que les membres du jury pour avoir accepté d'y participer.

Je voudrais également remercier Guillaume Charpiat et Jean-Philippe Pons pour les échanges et collaborations scientifiques qui apparaissent dans la première partie de cet ouvrage ainsi qu'Aileen McGonigal et Patrick Chauvel pour avoir proposé le thème de la deuxième partie de mon travail et m'avoir patiemment initié à des problématiques qui m'étaient jusque-là inconnues. Je remercie tout particulièrement Aileen pour son aide lors de la rédaction.

Merci à tous les membres du projet Odyssée que j'ai côtoyés et en particulier aux parisiens pour la bonne humeur, la vie quotidienne au labo. Je tiens aussi à remercier Alexandre Gramfort et Christian Benar pour leur aide technique au sujet du traitement du signal SEEG. Et enfin un grand merci à Patrick Labatut pour toute son inestimable aide technique et son fabuleux WasteuhWaïzer.

Introduction

Initially motivated by its applications to robotics, computer vision aimed to equip machines with an artificial visual system imitating as well as possible the human one. In particular it seemed to be essential to obtain three-dimensional informations on objects surrounding the robot, in order to avoid obstacles, recognize them and grab them. Classic robotic sensors was too slow or not precise enough and therefore only methods inspired by biological vision seemed to be appropriate. One typical example have been stereovision methods which retrieve three-dimensional informations from multiple cameras.

Nowadays, principal applications of computer vision are not only robotic: shape identification, remote detection, medical image processing, industrial inspection, satellite images, etc. Among these topics, shape recognition hold an important place and has many applications. For example in the medical domain when one wants to detect and locate tumors in an X-ray image of a brain or reconstruct three-dimensional shapes from MRI images, but also in other domain such as video surveillance for automatic early drowning detection at pool. The studied objects are then generally represented by their contours and the segmentation task consists in finding, in an image, a shape that partition the domain into disjoint sub-regions of interest. Usually this problem is resolved by defining an energy which measures for a given partition of the image, the "homogeneity" of the different parts. This "homogeneity" criterion is defined according to each specific application. It is often based on the intensity of the image inside each region (e.g. through histograms or more complex measures such as texture) and on the shape of the contour defining the partition. In character recognition for example the shape of the contour is the most relevant information to build a measure of quality of the segmentation. Once this energy is defined, the problem of segmenting the image can be boiled down to the problem of minimizing an energy functional with respect to the shape that partitions the image.

The first part of this work focuses on the theoretical point of view of modeling prior knowledge about shapes. We will start in chapter 1, *shapes and shape metrics*, by clarifying what we mean by shapes and defining several shape metrics. Understanding shape and its basic empirical statistics is important both in recognition and analysis, with applications ranging from medicine to security to consumer photography. Constructing metrics on the set of shapes is a first step in this direction since it provides a way to measure the similarity between two shapes. It also enables to define the *variational shape warping* problem, chapter 2. Continuously deforming one shape Γ_1 into another Γ_2 is one of the keys leading to statistical shape analysis. It offers more than a way to compare shapes, but also to compute their mean,

to analyze their variability, and eventually to obtain correspondence between them. This problem is closely related to a question of optimization on shapes, since we want to minimize a given distance between the two shapes Γ_1 and Γ_2 by evolving Γ_1. Once the gradient of a shape energy is defined, this can be done by a gradient descent scheme. A great advantage of the gradient descent method is that the energy functional to minimize can explicitly deal with criteria dedicated to the contour itself. For instance the contour could be asked to be attracted by regions of high intensity gradients in the image intensity, which means that the contour will be attracted by areas of big heterogeneity in the image, at the boundary between two homogeneous regions. But, even more interestingly, criteria concerning the shape of the object can easily be defined. For instance its smoothness is easily measured by the curvature of the boundary. And adding a restriction on the boundary (e.g. it should be smooth) can be seen as a prior on the shape of the object. A new way of adding prior knowledge in the shape optimization process will be presented in chapter 3, *generalized gradient: priors on minimization flows*. In fact, having some control on the form of the deformation vector, the gradient, also means controlling the kind of path that will be followed by the gradient descent scheme. We will first demonstrate that this control on the gradient can be assured by the choice of the inner product structure ruling the space of deformations. Then we will present a generalization of the gradient notion that can take into account priors on the deformation field form even without deriving from an inner product. Imposing constraints on the gradient can be useful in many case. This may be a way to avoid some unwanted local minima but also to insure that the family of shapes resulting from the minimization verifies some properties (e.g. spatial coherence during the evolution by favoring rigid transformation). Since it changes the gradient, it can even be used to transform a unusable deformation field in a better one. In chapter 4, *reconciling landmarks and level sets*, we will present an extension of the very well-known level set method which takes advantage of this property and enables the introduction of correspondences between points on the shapes in the energy functional. The warping is then guided by these landmarks points.

The second part focuses on the application side of our work. Interactions between computer vision and medicine have been increasing in the last years. A vast majority of theses interactions consists in medical image analysis: for a medical imaging modality (e.g. X-ray, PET: positron emission tomography, MRI: magnetic resonance imaging) image processing techniques are applied to help or automate diagnosis. The topic of our work is exploring a new aspect of the possible interactions between these two different scientific domains. We developed a method for facial expression analysis under uncon-

trolled imaging conditions and used it in the concrete case of the study of epilepsy seizures. We will first briefly present in chapter 5, *the context* of this problematic. Epilepsy is among the most common serious brain disorders, can occur at all ages, and is characterized by a variety of presentations and causes. It roughly affects about 1% of the world population. In a large majority of cases, seizures will be controlled by drugs, but for the remaining ones, a surgical treatment can be envisaged if the epileptic seizure is proved to come from a relatively localized area in the brain. This area could then possibly be removed by the surgical operation if this resection will not result in unacceptable problems with memory, vision, language or movement. In addition to the measurement of electrical activity produced by the brain during the seizure, the semiology (the study of signs and symptoms) of the epileptic seizures is therefore important for the medical team as it gives indications on the localization of a possible epileptogenic area. Among all the signs and symptoms studied during the seizure, the patient's facial expression seems to be a promising source of informations and a quantification of these expressions appeared to be a potentially useful tool for the medical team. That is why we developed a method for fitting a three-dimensional face model under uncontrolled imaging conditions (chapter 6, *facial expression analysis: 3D model fitting*). Although this work is not a direct application of the methods presented in the first part, we will see that it can also be written as an optimization problem since we want to find the position of a shape (the three-dimensional face model) that minimize a specific energy. This method alone is already a useful tool for studying the semiology of facial expression that may occur during epileptic seizures, since it produces a quantification of the facial expression: one can for example look into the numerical opening of the eyes or of the mouth and its variation during the seizure. Nevertheless what would be very interesting is to find a way to interrelate electrical activity produced by the brain during the seizure and the facial expressions. In chapter 7, *electro-clinical correlation*, we present a first step in this direction. During pre-surgical exploration, patients are videotaped at the same time as their electroencephalography or stereoelectroencephalography is recorded. The anatomo-electro-clinical correlation is investigated: the "anatomic" information comes from the knowledge of the electrodes localization on or in the brain , the "electrical" data are given by the electroencephalography (EEG) or stereoelectroencephalography (SEEG) recordings and the "clinical" information is the object of semiology and is obtained by interrogating the patient and observing the video. This close observation of clinical features (including the order in which they occured) and their relation to the region involved in the epileptic discharge, allows a temporal-spatial profile of the seizure's origin and propagation patterns to be established and the

corresponding anatomical structures to be defined, thus aiding the decision regarding surgical decision for each patient's particular epilepsy. Our work is an answer to the need of quantification of the electro-clinical correlation.

To conclude, let us note that a part of the material herein has led to several publications: chapters 1 and 2 comes partly from a chapter *Approximations of shape metrics and application to shape warping and empirical shape statistics* in the book *Statistics and Analysis of Shapes*, H. Krim & A. Yezzi editors, 2006, with Guillaume Charpiat, Olivier Faugeras and Renaud Keriven [23] and an article *Distance-based Shape statistics* [25] with the same authors in the proceedings of the *International Conference on Acoustics, Speech, and Signal Processing*, special session *Statistical Inferences on Nonlinear Manifolds with Applications in Signal and Image Processing* held in 2006. Most of chapter 3 comes from the article *Generalized gradients : priors on minimization flows* [26] with Guillaume Charpiat, Jean-Philippe Pons, Renaud Keriven and Olivier Faugeras, published in the *International Journal of Computer Vision*, 2007. The work presented in chapter 4 has been published in the article *Reconciling landmarks and level sets* [93] with Renaud Keriven and Olivier Faugeras in the proceedings of the *International Conference on Pattern Recognition* held in 2006. And finally, the chapter 6 is adapted from an article *3D model fitting for facial expression analysis under uncontrolled imaging conditions* [94] in the proceedings of the *International Conference on Pattern Recognition* that will be held in December 2008.

Almost all our implementations have been done in C++, using libraries developed by the Certis Team, CertisLibs [1], specially one for the Level-Set methods and one for graphics. In chapter 6, we also used a C-code implementing a quasi-Newton optimization method with non-linear constraints developed by Andre Tits' research group [84]. And in chapter 7, the treatment of the SEEG signal is implemented under Matlab, using a version of the "analytic wavelet transform" from the WaveLab toolbox (J. Buckheit and D. Donoho [43]) modified by C. Benar and M. Clerc.

[1] http://certis.enpc.fr/~keriven/CertisLibs/

Part I

Shape Optimization

Chapter 1

Shapes and Shape Metrics

In order to compare shapes or to deform a shape into an other one, a way to quantify the similarity between two shapes must be defined. In this chapter, we present several shapes metrics. These metrics have been introduced and studied by Guillaume Charpiat et al. in [22]. Here we principally review their work.

1.1 Introduction

Understanding shape and its basic empirical statistics is important both in recognition and analysis, with applications ranging from medicine to security to consumer photography. The basic metrics and statistics of static shapes have been the subject of numerous fundamental studies in recent years [15, 23, 45, 99, 127, 135].

There exist various definitions of the term shape in the literature. In 1977 David Kendall [75] for example defines shape as all the geometrical information that remains when location, scale and rotational effects are filtered out from a collection of point coordinates. Most research on statistical shape modeling has been devoted to explicit contour representations: objects are represented by a finite number of salient points or landmarks. We refer to the book by Ian Dryden and Kanti Mardia (1998) [45] for an overview. Working with explicit contour representations, shape metrics are naturally been defined based on this representation. One of the most used shape distance based on discrete representation of the shapes is the Procrustes distance ([69]). Several work have been done by deriving some shape metrics from this distance and it is still inspiring recent work (e.g. [76, 64, 52, 82, 83]). We can also cite works based on the Riemannian metric (see for example [118, 5, 96, 116])

Our work deals explicitly with curves as such, independently of their sampling or even parametrization. In fact, our work bears more resemblance with that of several other authors. Ulf Grenander et al. [65, 67, 66] pioneered the concept of considering shapes as points on an infinite dimensional manifold, representing shape deformations as the action of Lie groups on this manifold and computing statistics on the space of diffeomorphisms. Some more recent advances were done by Alain Trouvé, Laurent Younes and colleagues [147, 148, 138, 99]. For infinite dimensional groups such as diffeomorphisms [47, 137] which smoothly change the objects shapes previous authors have been dependent upon the choice of parameterizations and origins of coordinates. For them, warping a shape onto another requires the construction of families of diffeomorphisms that use these parameterizations.

Our approach, based upon the use of distance functions, does not require the arbitrary choice of parameterizations and origins. This is already nice in two dimensions but becomes even nicer in three dimensions and higher where finding parameterizations and tracking origins of coordinates can be a real problem: this is not required in our case. Another piece of related work is the one from Anthony Yezzi and Stefano Soatto [145] who tackles the problem of jointly extracting and characterizing the motion of a shape and its deformation. In order to do this they find inspiration in the above work on the use of diffeomorphisms and propose the use of a distance between shapes (based on the set-symmetric difference). This distance poses a number of problems which have been addressed in [23]. Therefore we will describe others distances which we believe to be more suitable.

1.2 Characteristic function and distance function to a shape

In our work we use a purely *geometric* definition of a shape. We define a shape Γ to be a smooth manifold of dimension k embedded in $\mathbb{R}^n$, for example a planar curve or a surface in the space $\mathbb{R}^3$. Since we are driven by image applications we also assume that all our shapes are contained in a hold-all regular open bounded subset of $\mathbb{R}^n$ which we denote by Ω. We denote by $\mathcal{S}$ the set of shapes. We refer the reader to [37] and [23] for a more rigorous and complete analysis of the set of shapes.

Since, as mentioned in the introduction, we want to be independent of any particular parametrization of the shape, we use two main ingredients,

the characteristic function of a shape Γ:

$$\begin{array}{rccl} \chi_\Gamma : & \Omega & \rightarrow & \{0,1\} \\ & \mathbf{x} & \mapsto & \left\{ \begin{array}{ll} 1 & \text{if } \mathbf{x} \in \Gamma \\ 0 & \text{if } \mathbf{x} \notin \Gamma \end{array} \right. \end{array} \tag{1.1}$$

and the *distance function* to the shape Γ :

$$\begin{array}{rccl} d_\Gamma : & \Omega & \rightarrow & \mathbb{R}^+ \\ & \mathbf{x} & \mapsto & \inf\limits_{\mathbf{y}\in\Gamma} d(\mathbf{x},\mathbf{y}) \end{array} \tag{1.2}$$

where $d(\cdot,\cdot)$ a distance on $\mathbb{R}^n$. Usually $d(\cdot,\cdot)$ is chosen to be the euclidean distance on $\mathbb{R}^n$. Fig. 1.1 illustrates the definitions of d_Γ.

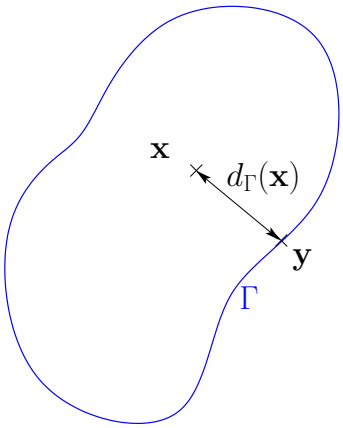


Figure 1.1: Distance between the point $\mathbf{x}$ and the shape Γ.

We also define the *signed distance function* to a shape Γ, denoted by $\tilde{d}_\Gamma$. $\tilde{d}_\Gamma$ is simply equal to d_Γ outside Γ and equal to $-d_\Gamma$ inside Γ:

$$\begin{array}{rccl} \tilde{d}_\Gamma : & \Omega & \rightarrow & \mathbb{R} \\ & \mathbf{x} & \mapsto & \left\{ \begin{array}{ll} -d_\Gamma(\mathbf{x}) & \text{if } \mathbf{x} \text{ is inside } \Gamma \\ d_\Gamma(\mathbf{x}) & \text{otherwise} \end{array} \right. \end{array} \tag{1.3}$$

Fig. 1.2 shows a graphic representation of the signed distance to a 2D-shape (in that case, the shape is the Unit Circle). This function is often used as a shape representation in the framework of the *level set method* (see chapter 4 and [39, 108]). Shape metrics based on signed distance function (or distance function) are then very convenient to compute, as we will see in section 1.3.2. Nevertheless, let us stress the fact that the methods presented

in the following chapters will be *independent* of the choice of the shape representation. We did introduce above the signed distance function to a shape but only in order to use it in the definition of several shape metrics (see the next section). As explained earlier, using this function make the choice of the level set method for representing the shapes the most convenient one and that is why we used this method *for our experiments*. But, again, any shape representation, such as explicit representation (a list of points), parametric representation or implicit representation, will be suitable for our methods.

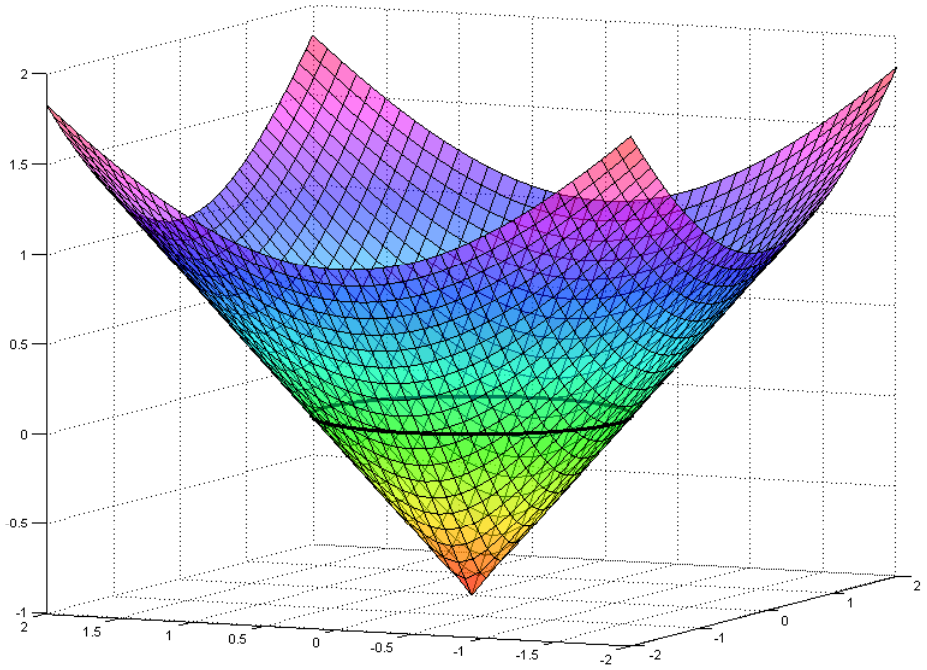


Figure 1.2: Graphic representation of the signed distance to a 2D-shape (the Unit Circle) .

Let us now state some useful simple properties of the three functions we introduced:

- The integral of the characteristic function is equal to the measure

(length, area)of Γ, that we note $|\Gamma|$:

$$|\Gamma| = \int_\Omega \chi_\Gamma(\mathbf{x})\, d\mathbf{x}$$

Note that this integral does not change if we add to or subtract from Γ a measurable set of Lebesgue measure 0 (also called a negligible set).

- the distance functions are continuous, in fact Lipschitz continuous with a Lipschitz constant equal to 1 [36, 38]:

$$\begin{aligned} |d_\Gamma(\mathbf{x}) - d_\Gamma(\mathbf{y})| &= |\inf_{\mathbf{z}\in\Gamma} d(\mathbf{x},\mathbf{z}) - \inf_{\mathbf{z}\in\Gamma} d(\mathbf{y},\mathbf{z})| \\ &\leq |\inf_{\mathbf{z}\in\Gamma} (d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})) - \inf_{\mathbf{z}\in\Gamma} d(\mathbf{y},\mathbf{z})| \\ &\leq |d(\mathbf{x},\mathbf{y}) + \inf_{\mathbf{z}\in\Gamma} d(\mathbf{y},\mathbf{z}) - \inf_{\mathbf{z}\in\Gamma} d(\mathbf{y},\mathbf{z})| \end{aligned}$$

and finally:

$$|d_\Gamma(\mathbf{x}) - d_\Gamma(\mathbf{y})| \leq |\mathbf{x}-\mathbf{y}|,\ \forall(\mathbf{x},\mathbf{y}) \in \Omega^2$$

- Thanks to the Rademacher theorem [51], this implies that $d\Gamma$ is differentiable almost everywhere in Ω, i.e. outside of a negligible set, and that the magnitude of its gradient, where it exists, is less than or equal to 1:

$$|\nabla d_\Gamma(\mathbf{x})| \leq 1, \text{ almost everywhere}$$

1.3 Shape metrics

1.3.1 Similarity measure based on characteristic function

The similarity measure we are about to define is based upon the characteristic functions of the two shapes we want to compare. We denote by $X(\Omega)$ the set of characteristic functions of measurable subsets of Ω. Given two such sets Γ_1 and Γ_2, we define their distance:

$$\begin{aligned} \rho(\Gamma_1,\Gamma_2) &= \|\chi_{\Gamma_1} - \chi_{\Gamma_2}\|_{L^2(\Omega)} \\ &= \left(\int_\Omega \left(\chi_{\Gamma_1}(\mathbf{x}) - \chi_{\Gamma_2}(\mathbf{x})\right)^2 d\mathbf{x}\right)^{\frac{1}{2}} \end{aligned}$$

Since the integral does not change if we modify the values of χ_{Γ_1} or χ_{Γ_1} over negligible sets, we immediately notice that this is not a distance between

the two shapes Γ_1 and Γ_2 but between their equivalence classes $[\Gamma_1]_m$ and $[\Gamma_1]_m$ of measurable sets. Given a measurable subset Γ of Ω, we define its equivalence class $[\Gamma]_m$ as $[\Gamma]_m = \{\Gamma' \mid \Gamma' \text{ is measurable and } \Gamma\Delta\Gamma' \text{ is negligible }\}$, where $\Gamma\Delta\Gamma'$ is the symmetric difference (see Fig. 1.3):

$$\Gamma\Delta\Gamma' = \{\mathbf{x} \in \Omega \mid \mathbf{x} \in \Gamma, \mathbf{x} \notin \Gamma' \text{ or } \mathbf{x} \in \Gamma', \mathbf{x} \notin \Gamma\}$$

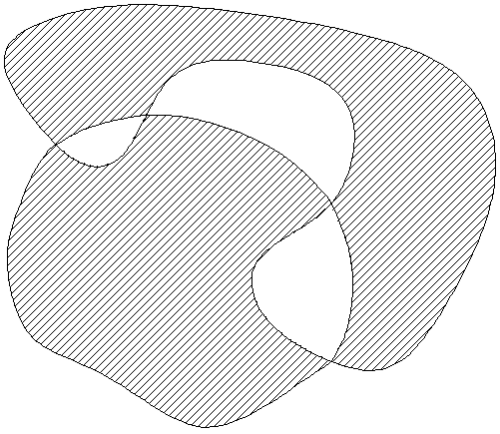


Figure 1.3: The hachured area is the symmetric difference of this two shapes

We have the following property:

$$\begin{aligned}\chi_{\Gamma\Delta\Gamma'}(\mathbf{x}) &= \chi_\Gamma(\mathbf{x})(1-\chi_{\Gamma'}(\mathbf{x})) + \chi_{\Gamma'}(\mathbf{x})(1-\chi_\Gamma(\mathbf{x})) \\ &= \chi_\Gamma(\mathbf{x}) + \chi_{\Gamma'}(\mathbf{x}) - 2\chi_\Gamma(\mathbf{x})\chi_{\Gamma'}(\mathbf{x}) \\ &= \left(\chi_\Gamma\left(\mathbf{x}\right) - \chi_{\Gamma'}\left(\mathbf{x}\right)\right)^2\end{aligned}$$

and finally we can write:

$$\rho(\Gamma_1, \Gamma_2) = \left(\int_\Omega \chi_{\Gamma_1\Delta\Gamma_2}(\mathbf{x}) d\mathbf{x}\right)^{\frac{1}{2}}$$

and, using the first on the property mentioned above:

$$\rho(\Gamma_1, \Gamma_2) = \sqrt{|\Gamma_1\Delta\Gamma_2|}$$

This function has been used as a similarity measure in the work of Soatto and Yezzi [145]. However, in [22], authors stressed that the set of characteristic function does not satisfy important properties such as *compactness* and therefore they introduce different similarity measure.

1.3.2 Similarity measure based on distance functions

(a) d_{L^2} and $d_{W^{1,2}}$

Some other distances are based on signed distance functions. As said previously the signed distance function to a shape can be seen as an implicit representation of the shape. These functions belong to the Sobolev space $W^{1,2}(\Omega)$ of square integrable functions with square integrable derivatives. As a matter of fact the gradient of a distance function is known to be of magnitude equal to 1 except on a subset of measure 0 of Ω. We can then define two metrics based on signed distance functions:

Definition 1. *The L^2 distance, noted $d_{L^2}(\Gamma_1, \Gamma_2)$, between two shapes Γ_1 and Γ_2 is defined as the $\|.\|_{L^2}$ norm between their signed distance function:*

$$\begin{aligned} d_{L^2}(\Gamma_1, \Gamma_2)^2 &= \left\| \tilde{d}_{\Gamma_1} - \tilde{d}_{\Gamma_2} \right\|^2_{L^2(\Omega,\mathbb{R})} \\ &= \left(\int_\Omega \left(\tilde{d}_{\Gamma_1}(\mathbf{x}) - \tilde{d}_{\Gamma_2}(\mathbf{x}) \right)^2 d\mathbf{x} \right)^{\frac{1}{2}} \end{aligned}$$

Definition 2. *The $W^{1,2}$ distance, noted $d_{W^{1,2}}(\Gamma_1, \Gamma_2)$, between two shapes Γ_1 and Γ_2 is defined as the Sobolev norm, $\|.\|_{W^{1,2}}$ between their signed distance function:*

$$\begin{aligned} d_{W^{1,2}}(\Gamma_1, \Gamma_2)^2 &= \left\| \tilde{d}_{\Gamma_1} - \tilde{d}_{\Gamma_2} \right\|^2_{W^{1,2}(\Omega,\mathbb{R})} \\ &= \left\| \tilde{d}_{\Gamma_1} - \tilde{d}_{\Gamma_2} \right\|^2_{L^2(\Omega,\mathbb{R})} + \left\| \nabla \tilde{d}_{\Gamma_1} - \nabla \tilde{d}_{\Gamma_2} \right\|^2_{L^2(\Omega,\mathbb{R}^n)} \\ &= \left(\int_\Omega \left(\tilde{d}_{\Gamma_1}(\mathbf{x}) - \tilde{d}_{\Gamma_2}(\mathbf{x}) \right)^2 d\mathbf{x} \right)^{\frac{1}{2}} + \left(\int_\Omega \left\| \nabla \tilde{d}_{\Gamma_1}(\mathbf{x}) - \nabla \tilde{d}_{\Gamma_2}(\mathbf{x}) \right\|^2 d\mathbf{x} \right)^{\frac{1}{2}} \end{aligned}$$

(b) Hausdorff distance, d_H

One of the broadly used distance between shapes is the *Hausdorff distance*:

Definition 3. *The Hausdorff distance between two shapes Γ_1 and Γ_2 is noted $d_H(\Gamma_1, \Gamma_2)$ and defined by:*

$$d_H(\Gamma_1, \Gamma_2) = \max \left\{ \sup_{\mathbf{x} \in \Gamma_1} d_{\Gamma_2}(\mathbf{x}), \sup_{\mathbf{x} \in \Gamma_2} d_{\Gamma_1}(\mathbf{x}) \right\}$$

The Hausdorff distance is computing by finding the point of the shape Γ_1 which is the farthest of the shape Γ_2 and the point of the shape Γ_2 which is the farthest of the shape Γ_1, and taking the maximum of these two distances. Fig. 1.4 illustrates the definition of the Hausdorff distance. The Hausdorff distance is a metric on the set of shapes. It means that it verifies the following properties:

1. $\forall(\Gamma_1, \Gamma_2, \Gamma_3) \in \mathcal{S}^3,\ d_H(\Gamma_1, \Gamma_2) \leq d_H(\Gamma_1, \Gamma_3) + d_H(\Gamma_3, \Gamma_2)$
2. $\forall(\Gamma_1, \Gamma_2) \in \mathcal{S}^2,\ d_H(\Gamma_1, \Gamma_2) = 0 \Leftrightarrow \Gamma_1 = \Gamma_2$

This distance is one of the most known distance used when there is a need

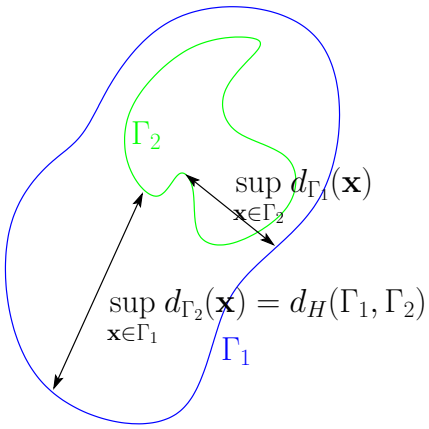


Figure 1.4: Hausdorff Distance between the shapes Γ_1 and Γ_2.

to compare two shapes. Let us also mention the Gromov-Hausdorff distance which measures how far two compact metric spaces are from being isometric. If Γ_1 and Γ_2 are two compact metric spaces, then $d_GH(\Gamma_1, \Gamma_2)$ is defined to be the infimum of all numbers $d_H(f(\Gamma_1), g(\Gamma_2))$ for all metric spaces M and all isometric embeddings $f : \Gamma_1 \mapsto M$ and $g : \Gamma_2 \mapsto M$ (the isometric embedding is understood in the global sense, i.e it must preserve all distances). We will not study in more details this distance but we mention that recently F.Memoli studied the use of Gromov-Hausdorff Distances for Shape Comparison and refer the reader to [100, 101].

Although d_H, d_{L^2} and $d_{W^{1,2}}$ can be shown to be equivalent (see [23]),we will see later that the choice of one or the other is not neutral with respect to warping and to computational complexity.

1.4 Differentiable approximation of these similarity measures

In chapter 2 we will need to take the derivative of the distances with respect to the shapes. Since the Hausdorff distance is not differentiable because it involves infima and suprema, we propose a family of smooth approximations which are continuous with respect to the Hausdorff topology. These family of approximations have been introduced by Charpiat et al. in [22] and extended by us to the 3D-case in [23].

Authors of [22] build a series of smooth approximations of the Hausdorff distance $d_H(\Gamma_1, \Gamma_2)$ of two shapes Γ_1 and Γ_2. The definition of the Hausdorff distance (Def. 3) involves the distance functions to the two shapes. We first show how to approximate those two functions.

First we will introduce a notation. In the following we will have to integrate several functions on shapes.

Definition 4. *For $f : \Gamma \mapsto \mathbb{R}^n$ and a shape Γ we denoted the integral of f on Γ (also called the* curvilinear integral*) by:*

$$\int_\Gamma f(\mathbf{x})\, d\Gamma(\mathbf{x})$$

where $d\Gamma(\mathbf{x})$ stands for the area element of the contour at the point $\mathbf{x} \in \Gamma$ such that the integral over Γ is intrinsic and does not depend on the parametrization. It means that for any parametrization $\mathcal{P}_\Gamma$ of Γ:

$$\begin{array}{rccc} \mathcal{P}_\Gamma : & [0,1] \subset \mathbb{R}^k & \rightarrow & \Gamma \subset \mathbb{R}^n \\ & \sigma & \mapsto & \mathcal{P}_\Gamma(\sigma) \end{array}$$

we have :

$$\int_\Gamma f(\mathbf{x}) d\Gamma(\mathbf{x}) = \int_0^1 f\left(\mathcal{P}_\Gamma(\sigma)\right) |\mathcal{P}'_\Gamma(\sigma)| d\sigma$$

Let us now focus on $d_{\Gamma_1} : \Gamma_2 \mapsto \mathbb{R}^+$. Since d_{Γ_1} is Lipschitz continuous on the bounded set Ω it is integrable on the compact set Γ_2 and we have [121, Chapter 3, problem 4]:

$$\lim_{\beta \to \infty} \left(\frac{1}{|\Gamma_2|} \int_{\Gamma_2} (d_{\Gamma_1}(\mathbf{x}))^\beta \, d\Gamma_2(\mathbf{x}) \right)^{\frac{1}{\beta}} = \sup_{\mathbf{x} \in \Gamma_2} d_{\Gamma_1}(\mathbf{x}) \tag{1.4}$$

where $|\Gamma_2|$ is the length of Γ_2: $|\Gamma_2| = \int_{\Gamma_2} d\Gamma_2(\mathbf{x})$.

Moreover the function $\mathbb{R}^+ \mapsto \mathbb{R}^+$ defined by $\beta \rightarrow \left(\frac{1}{|\Gamma_2|}\int_{\Gamma_2}\left(d_{\Gamma_1}(\mathbf{x})\right)^{\beta} d\Gamma_2(\mathbf{x})\right)^{\frac{1}{\beta}}$ is monotonously increasing [121, Chapter 3, problem 5]. Similar properties hold obviously for d_{Γ_2}.

However the definition of d_{Γ_1} involves itself an infimum, since we have $d_{\Gamma_1}(\mathbf{x}) = \inf_{\mathbf{y}\in\Gamma 1} d(\mathbf{x},\mathbf{y})$. Then we need to build an approximation also for d_{Γ_1}. Let us consider a continuous *strictly decreasing* function $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}^{+*}$. We have, for $\mathbf{y} \in \Omega$:

$$\sup_{\mathbf{x}\in\Gamma_2} \varphi\big(d(\mathbf{x},\mathbf{y})\big) = \varphi\big(\inf_{\mathbf{x}\in\Gamma_2} d(\mathbf{x},\mathbf{y})\big) = \varphi\big(d_{\Gamma_2}(\mathbf{y})\big)$$

and therefore:

$$\lim_{\alpha\to\infty}\left(\frac{1}{|\Gamma_2|}\int_{\Gamma_2}\varphi\big(d(\mathbf{x},\mathbf{y})\big)^{\alpha} d\Gamma_2(\mathbf{x})\right)^{\frac{1}{\alpha}} = \sup_{\mathbf{x}\in\Gamma_2}\varphi\big(d(\mathbf{x},\mathbf{y})\big) = \varphi\big(d_{\Gamma_2}(\mathbf{y})\big)$$

Because φ is continuous and strictly monotonously decreasing, it is one to one andφ^{-1} is strictly monotonously decreasing and continuous. Then:

$$d_{\Gamma_2}(\mathbf{y}) = \lim_{\alpha\to\infty}\varphi^{-1}\left(\left(\frac{1}{|\Gamma_2|}\int_{\Gamma_2}\varphi\big(d(\mathbf{x},\mathbf{y})\big)^{\alpha} d\Gamma_2(\mathbf{x})\right)^{\frac{1}{\alpha}}\right) \tag{1.5}$$

The definition of the Hausdorff distance (Def. 3) involves a last infimum: the discrete maximum between the two sup. We will use the following relation:

$$\lim_{\gamma\to\infty}(a_1^{\gamma} + a_2^{\gamma})^{\frac{1}{\gamma}} = \max(a_1, a_2) \tag{1.6}$$

Using the three previous relation (1.4, 1.5, 1.6) and choosing a strictly decreasing function for φ (such as $\varphi(x) = \frac{1}{x+\epsilon}, x > 0, \epsilon > 0$), we can define a family of approximations of the Hausdorff distance for each triplet $(\alpha,\beta,\gamma) \in \mathbb{R}^{+3}$. For the sake of simplicity, we will first introduce a notation. Let Γ be a shape and $f : \Gamma \mapsto \mathbb{R}^n$ an integrable function on Γ. We denote by $\lfloor f \rfloor_{\Gamma}$ the *average of* f *on* Γ:

$$\lfloor f \rfloor_{\Gamma} = \frac{1}{|\Gamma|}\int_{\Gamma} f(\mathbf{x}) d\Gamma(\mathbf{x})$$

and by $\lfloor f \rfloor_{\Gamma}^{\varphi}$ the φ*-average of* f *on* Γ for f positive integrable function and for any continuous strictly monotonous (hence one to one) function $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}^+$:

$$\lfloor f \rfloor_{\Gamma}^{\varphi} = \varphi^{-1}\left(\lfloor \varphi \circ f \rfloor_{\Gamma}\right) = \varphi^{-1}\left(\frac{1}{|\Gamma|}\int_{\Gamma}\varphi(f(\mathbf{x}))d\Gamma(\mathbf{x})\right)$$

Therefore equation 1.4 can be rewritten as:

$$\sup_{\mathbf{x}\in\Gamma_2} d_{\Gamma_1}(\mathbf{x}) = \lim_{\beta\to\infty} \lfloor d_{\Gamma_1} \rfloor_{\Gamma_2}^{p_\beta}$$

with $p_\beta : r \in \mathbb{R}^+ \mapsto r^\beta \in R^+$. And equation 1.5 as:

$$d_{\Gamma_2}(\mathbf{y}) = \lim_{\alpha\to\infty} \lfloor d(\mathbf{x},\cdot) \rfloor_{\Gamma_2}^{\varphi_\alpha}$$

with $\varphi_\alpha : r \in \mathbb{R}^+ \mapsto \varphi(r)^\alpha \in R^+$.

And finally, by extending this notation to the discrete case:

$$\lfloor a_1, \cdots, a_n \rfloor^\varphi = \varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(a_i)\right)$$

we obtain the following property:

$$d_H(\Gamma_1,\Gamma_2) = \lim_{\alpha,\beta,\gamma\to\infty} \left\lfloor \left\lfloor \lfloor d(\cdot,\cdot)\rfloor_{\Gamma_2}^{\varphi_\alpha}\right\rfloor_{\Gamma_1}^{p_\beta}, \left\lfloor \lfloor d(\cdot,\cdot)\rfloor_{\Gamma_1}^{\varphi_\alpha}\right\rfloor_{\Gamma_2}^{p_\beta} \right\rfloor^{p_\gamma}$$

This naturally yields to an approximation of the Hausdorff distance:

Definition 5. *For a strictly decreasing function φ and a triplet $(\alpha,\beta,\gamma) \in \mathbb{R}^{+3}$, we will note:*

$$\rho_H^{\varphi,\alpha,\beta,\gamma}(\Gamma_1,\Gamma_2) = \left\lfloor \left\lfloor \lfloor d(\cdot,\cdot)\rfloor_{\Gamma_2}^{\varphi_\alpha}\right\rfloor_{\Gamma_1}^{p_\beta}, \left\lfloor \lfloor d(\cdot,\cdot)\rfloor_{\Gamma_1}^{\varphi_\alpha}\right\rfloor_{\Gamma_2}^{p_\beta} \right\rfloor^{p_\gamma}$$

We refer the reader to [23] for a more precise study of this family of function and its properties.

By using equation 1.5, one can also immediately define families of approximations for the L^2 and $W^{1,2}$ distances, replacing in their definition (Def. 1 and Def. 2) the signed distance functions by their approximations:

Definition 6. *For a strictly decreasing function for φ and $\alpha > 0$, we will note:*

$$\rho_{L^2}^{\varphi,\alpha}(\Gamma_1,\Gamma_2) = \left\| \left\lfloor \tilde{d}(\mathbf{x},\cdot)\right\rfloor_{\Gamma_1}^{\varphi_\alpha} - \left\lfloor \tilde{d}(\mathbf{x},\cdot)\right\rfloor_{\Gamma_2}^{\varphi_\alpha} \right\|_{L^2(\Omega,\mathbb{R})}$$

$$\rho_{W^{1,2}}^{\varphi,\alpha}(\Gamma_1,\Gamma_2) = \left\| \left\lfloor \tilde{d}(\mathbf{x},\cdot)\right\rfloor_{\Gamma_1}^{\varphi_\alpha} - \left\lfloor \tilde{d}(\mathbf{x},\cdot)\right\rfloor_{\Gamma_2}^{\varphi_\alpha} \right\|_{W^{1,2}(\Omega,\mathbb{R})}$$

with, for $\mathbf{x} \in \Gamma, \tilde{d}(\mathbf{x},\mathbf{y}) = -d(\mathbf{x},\mathbf{y})$ if $\mathbf{y}$ is inside Γ and $\tilde{d}(\mathbf{x},\mathbf{y}) = d(\mathbf{x},\mathbf{y})$ otherwise.

1.5 What can be done with only shape metrics: Shape Statistics

In the following chapters, we will use these shape metrics essentially for shape warping: a shape energy is derived from a shape metric and the gradient of this energy with respect to the shape is define and computed in order to apply a gradient descent scheme. Let us first briefly present in this section a technique concisely described in [25]. This work and the presented figure has been done by G.Charpiat.

The objective is to build a low-dimensional map of a set of shapes: it is possible to build such a map when given only the distances between shapes, from a purely static geometrical point of view, thanks to the graph Laplacian technique [7].

Let us denote by $(\Gamma_{1 \leqslant i \leqslant n})$ a set of n shapes and consider a distance d on this set, for example one of the distance described above. We fix a positive integer K and search for the K nearest neighbors $N^i_{1 \leqslant l \leqslant K}$ of each shape Γ_i for the chosen distance d. We then define a symmetric weight matrix W by

$$W_{i,j} = \delta_{i,j} \;\; e^{-\frac{d(\Gamma_i,\Gamma_j)^2}{2\sigma^2}}$$

where

$$\delta_{i,j} = \begin{cases} 1 & \text{if} \quad i \in N^j \quad \text{or} \quad j \in N^i \\ 0 & \text{otherwise} \end{cases}$$

and we have chosen for σ the mean distance between neighbors:

$$\sigma = \frac{\sum_{i,j} \; d(\Gamma_i, \Gamma_j) \;\; \delta_{i,j}}{\sum_{i,j} \;\; \delta_{i,j}}.$$

Then, let us consider the symmetric negative semi-definite matrix $L = W - D$ where $D_{i,j} = \sum_i \; W_{i,j} \; \delta_{i,j}$: it is a discrete approximation of the Laplacian operator. Thus, as explained in [7], its eigenvectors F_k of highest (negative) non-zero eigenvalues are the best functions from the shapes Γ_i to $\mathbb{R}$ that could be used as coordinate system of the set of shapes. Consequently, we obtain a natural map in $\mathbb{R}^m$ where each shape Γ_i is represented by a dot with coordinates $(F_{1 \leqslant k \leqslant m}(\Gamma_i))$. Let us try this approach on an artificial dataset. We build a set of rectangles with same center and width but different lengths and orientations, so there are two natural parameters we would expect the algorithm to find. Rectangles are randomly chosen such that the distribution of their corners is the uniform law in the authorized area (orientation between $-\frac{\pi}{6}$ and $+\frac{\pi}{6}$, and length between 2 and 4 times the width). Results vary

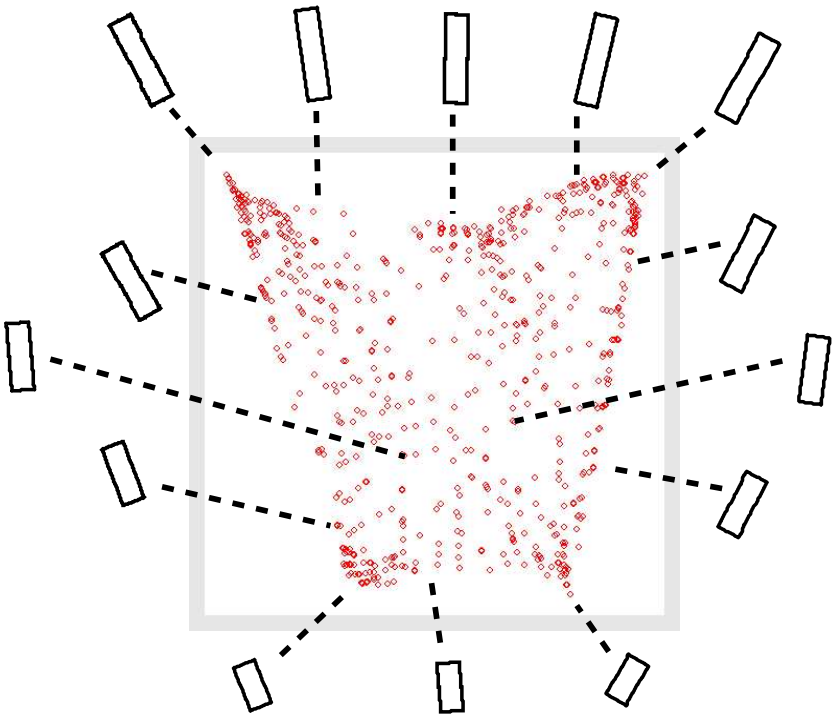


Figure 1.5: Map from the graph Laplacian method for a set of rectangles whose length and orientation have been chosen randomly ($K = 15$).

depending on the distribution density and the value of K: the higher the density, the better the results. Fig. 1.5 has been computed for 700 rectangles.

Let us now study the more complicated case of some different classes in a same connected component. We consider a set of 111 fish from the same database as before. The resulting map for the two first coordinates (see Fig. 1.6) shows some clusters of fish families.

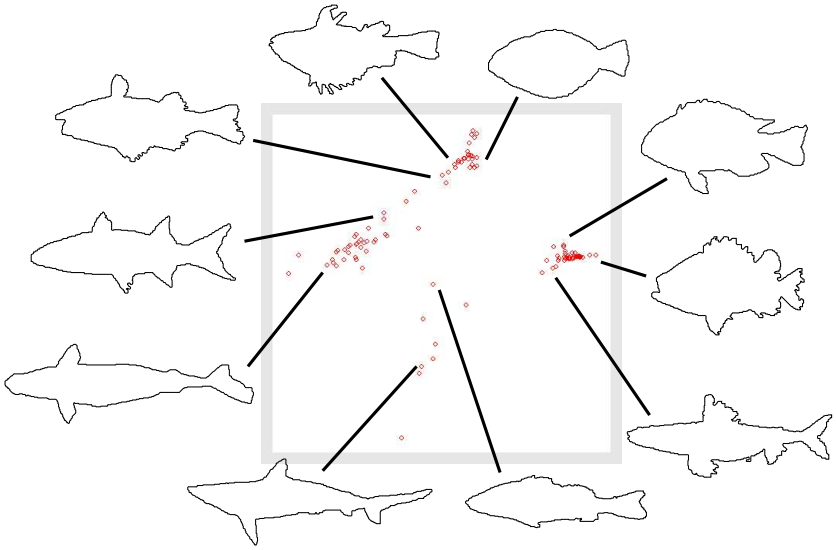


Figure 1.6: Two first coordinates for a set of 111 fish from different classes. The elements from each family are got together into clusters ($K = 25$).

Chapter 2

Variational shape warping

This chapter is dedicated to deforming a shape into another one using the minimization of the distance between them with respect to the shape in a gradient descent framework. The differentiation of an energy with respect to a curve is introduced and applied to the different distances introduced in chapter 1. Results of warping are shown.

Publications related to this chapter: *Approximations of shape metrics and application to shape warping and empirical shape statistics* [23] in the book *Statistics and Analysis of Shapes*, H. Krim & A. Yezzi editors, 2006, with Guillaume Charpiat, Olivier Faugeras and Renaud Keriven and *Distance-based Shape statistics* [25] with the same authors in the proceedings of the *International Conference on Acoustics, Speech, and Signal Processing*, special session *Statistical Inferences on Nonlinear Manifolds with Applications in Signal and Image Processing* held in 2006.

2.1 Introduction

Continuously deforming one shape Γ_1 into another Γ_2 is one of the keys leading to statistical shape analysis [15, 45]. It offers a way to compare shapes, to compute their mean, to analyze their variability, and eventually to obtain correspondence between them. Roughly speaking, the *warping problem* consists in transforming an initial shape into a target one: the result is the family of the intermediate shapes. Slightly different, though closely related, is the *matching problem*, where a correspondence (sometimes one-to-one, but not always) between two given shapes has to be established, regardless to some path from one of the shape into the other one. Indeed, if a warping process keeps track of the motion of every point of the initial shape all along the transformation, it induces a matching.

The diffeomorphic matching problem [99, 138] is often cited in that context. Indeed, as far as applications like brain warping [135, 136] are concerned, shapes (surfaces) and images are closely coupled. In that particular case, considering the images (e.g. MR scans) from which the shapes (e.g. the cortical surface) are extracted is natural and matching the images themselves is justified. Depending on the particular method [147, 58, 60, 59], the shapes are more or less involved in the image matching process, usually as two or three-dimensional curves or points. However, one should not be confused: here, the recovered matching is a diffeomorphism of the spaces in which the shapes are embedded.

Rather different is the original problem of shape warping, motivated by recognition, tracking or segmentation tasks. Since the beginning, it has been formulated in purely *geometric* term of shape distance[37]. In this chapter we present a variational shape warping method based on an energy minimization. We assume that we are given a function $E : S \times S \rightarrow \mathbb{R}^+$, the energy. This energy can be thought of as a measure of dissimilarity between two shapes, such as one of the shape metrics introduced in chapter 1. Warping a shape Γ_1 into another one Γ_2 can be stated as the minimization of the energy $E(., \Gamma_2)$ starting from Γ_1, i.e. finding a family of shapes $\{\Gamma(t),\ t \geq 0\}$ with $\Gamma(0) = \Gamma_1$ and $\Gamma(t)$ following some gradient descent toward Γ_2.

2.2 Gâteaux derivatives

In the following we consider a shape Γ, seen as a manifold of dimension k embedded in R^n. We denote by $E(\Gamma)$ the energy functional to be minimized. In the problem of warping a shape Γ into a target shape Γ_2, we may have $E(\Gamma) = d(\Gamma, \Gamma_2)$ where $d(.,.)$ is for example a distance on the set of shapes. In order to minimize this energy, we first need to define the gradient of this functional. We will use the *Gâteaux differentiation* framework.

Let us start by defining $\mathscr{D}$, the *deformation space*, set of all velocity fields defined on the shape Γ. A *velocity field* is a function defined on the shape Γ with values in $\mathbb{R}^n$:

$$\mathscr{D} = \{v \mid v : \Gamma \rightarrow \mathbb{R}^n\} \tag{2.1}$$

Fig. 2.1 shows a graphic representation of a velocity field v defined on a two-dimensional shape Γ.

Let us then define, for a given $v \in \mathscr{D}$, the shape $\Gamma + \epsilon\, v$. For any parametrization of Γ:

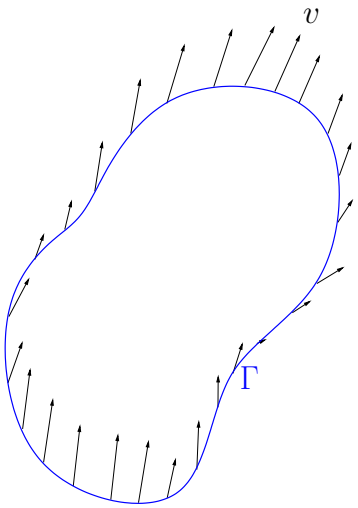


Figure 2.1: Representation of a velocity field v defined on the two-dimensional shape Γ.

$$\begin{array}{rccc} \mathcal{P}_\Gamma : & \mathcal{I} \subset \mathbb{R}^k & \rightarrow & \Gamma \subset \mathbb{R}^n \\ & \sigma & \mapsto & \mathcal{P}_\Gamma(\sigma) \end{array}$$

the shape $\Gamma + \epsilon\, v$ is defined by the parametrization:

$$\begin{array}{rccl} \mathcal{P}_{\Gamma+\epsilon\, v} : & \mathcal{I} & \rightarrow & \mathbb{R}^n \\ & \sigma & \mapsto & \mathcal{P}_{\Gamma+\epsilon\, v}(\sigma) = \mathcal{P}_\Gamma(\sigma) + \epsilon\, v(\mathcal{P}_\Gamma(\sigma)) \end{array}$$

There is no guarantee that $\Gamma + \epsilon\, v$ actually belongs to the set of shapes $\mathcal{S}$ in general but if v is $\mathcal{C}^2$ then we can find an ϵ small enough for $\Gamma + \epsilon\, v$ to be in $\mathcal{S}$. Fig. 2.2 shows, for the same v as in Fig. 2.1 and an $\epsilon < 1$, the shape $\Gamma + \epsilon\, v$.

If we consider the shape Γ as a point in $\mathcal{S}$, the manifold of all admissible shapes, then $\mathscr{D}$, the deformation space, can be seen as the tangent space of Γ. Furthermore, we have the following definition. the Gâteaux derivatives of the energy $E(.)$ at Γ and in the direction of $v \in \mathscr{D}$, denoted by $\mathcal{G}_\Gamma(E(.), v)$ can be defined as:

Definition 7. *The Gâteaux derivatives of the energy $E(.)$ at Γ and in the direction of $v \in \mathscr{D}$, denoted by $\mathcal{G}_\Gamma(E(.), v)$ can be defined as:*

$$\mathcal{G}_\Gamma(E(.), v) = \lim_{\epsilon \to 0} \frac{E(\Gamma + \epsilon\, v) - E(\Gamma)}{\epsilon}\ .$$

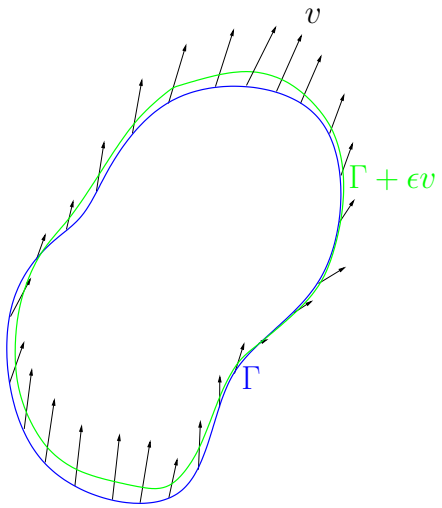


Figure 2.2: The shape $\Gamma + \epsilon\, v$

2.3 Shape gradient

The vector space of velocity fields $\mathscr{D}$ can be considered as a vector subspace of the Hilbert space $L^2(\Gamma, \mathbb{R}^n)$ (of all square integrable functions of Γ with values in $\mathbb{R}^n$) with the usual Hilbert product:

$$\langle v_1 \,|v_2\rangle = \frac{1}{|\Gamma|}\int_\Gamma v_1(\mathbf{x}) \cdot v_2(\mathbf{x})\, d\Gamma(\mathbf{x})$$

where $|\Gamma|$ is the length of Γ: $|\Gamma| = \int_\Gamma d\Gamma(\mathbf{x})$.

Therefore, if for a given shape Γ the application $v \mapsto \mathcal{G}_\Gamma(E(.), v)$ is linear and continuous, it defines a continuous linear form on $\mathscr{D}$ and we can apply *the Riesz's representation theorem* (also called Fréchet-Riesz theorem) [121] to this function. It says that it exists a unique velocity field, which we note $\nabla E(\Gamma) \in \mathscr{D}$, such that for any $v \in \mathscr{D}$, $\mathcal{G}_\Gamma(E(.), v) = \langle \nabla E(\Gamma) \,|v\rangle$. This velocity field $\nabla E(\Gamma)$ is called the *gradient of* $E(\Gamma)$.

Definition 8. *The gradient of the energy $E(\cdot)$ at Γ is the velocity field, noted $\nabla E(\Gamma) \in \mathscr{D}$ such that, for any $v \in \mathscr{D}$:*

$$\mathcal{G}_\Gamma(E(.), v) = \langle \nabla E(\Gamma) \,|v\rangle$$

The linearity and continuity of the function $v \mapsto \mathcal{G}_\Gamma(E(.), v)$ is out of the scope of this chapter. Indeed, it is highly dependent on the properties of each particular energy $E(.)$. Let us just say that each energy functional we are

considering verifies these properties and this fact has been assessed during the explicit computation of the gradient.

In Def. 8, it is important to stress that we implicitly took $\mathscr{D}$ as a subspace of $L^2(\Gamma, \mathbb{R}^n)$ with its usual inner product. Another Hilbert spaces and thus another inner products could be considered and yield to a different gradient. In chapter 3 We will expose the consequences of the change of the inner product and show some example of possible inner product as well as their effects on the gradient.

Equipped some shape gradient, we can define the warping of a shape Γ_1 into another one Γ_2 as finding the family $\Gamma(t)$ solution of the following Partial Differential Equation:

$$\begin{cases} \Gamma(0) = \Gamma_1 \\ \dfrac{d\Gamma}{dt} = -\nabla_{\mathscr{D}} E(\Gamma) \end{cases} \tag{2.2}$$

where $E(\Gamma) = d(\Gamma, \Gamma_2)$ and $d(\cdot, \cdot)$ is a measure of dissimilarity between two shapes.

2.4 Actual computation of shape gradients and tangential velocity field

Let us consider the two-dimensional case of an energy functional which can be written in the form of an integral of the shape:

$$E(\Gamma) = \int_\Gamma f(\mathbf{x}) d\Gamma(x)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ (we can for example think of the approximations of the Hausdorff distance presented in 1.4). This kind of energy functional are only based on *the geometrical form* of the shape.

In order to find a velocity field verifying the Riesz's property (see Def. 8), we have to write the expression of the Gâteaux derivative in the direction of a v as an inner product between v and a velocity field w not depending on v. If the Riesz's representation theorem applies, w can be identified to the gradient of E and it is unique. We recall the definition of the Gâteaux derivative of E at Γ and in the direction of v:

$$\mathcal{G}_\Gamma(E(.), v) = \lim_{\epsilon \to 0} \frac{E(\Gamma + \epsilon\, v) - E(\Gamma)}{\epsilon}\ .$$

Let us note $\Lambda_{\epsilon,v} = \Gamma + \epsilon\, v$ and let $\mathcal{P}_\Gamma : [0,1] \to \Gamma$ be a parametrization of Γ. Then by definition of $\Lambda_{\epsilon,v}$ a parametrization of $\Lambda_{\epsilon,v}$ can be:

$$\begin{array}{rccl} \mathcal{P}_{\Lambda_{\epsilon,v}}: & [0,1] & \rightarrow & \mathbb{R}^n \\ & \sigma & \mapsto & \mathcal{P}_{\Lambda_{\epsilon,v}}(\sigma) = \mathcal{P}_\Gamma(\sigma) + \epsilon\, v(\mathcal{P}_\Gamma(\sigma)) \end{array}$$

and we can write:

$$\begin{aligned} E(\Lambda_{\epsilon,v}) - E(\Gamma) &= \int_{\Lambda_{\epsilon,v}} f(\mathbf{x}) d\Lambda_{\epsilon,v}(\mathbf{x}) - \int_\Gamma f(\mathbf{x}) d\Gamma(\mathbf{x}) \\ &= \int_0^1 f\left(\mathcal{P}_{\Lambda_{\epsilon,v}}(\sigma)\right) \left|\mathcal{P}'_{\Lambda_{\epsilon,v}}(\sigma)\right| d\sigma - \int_0^1 f\left(\mathcal{P}_\Gamma(\sigma)\right) \left|\mathcal{P}'_\Gamma(\sigma)\right| d\sigma \\ &= \int_0^1 \left(f\left(\mathcal{P}_{\Lambda_{\epsilon,v}}(\sigma)\right) \left|\mathcal{P}'_{\Lambda_{\epsilon,v}}(\sigma)\right| - f\left(\mathcal{P}_\Gamma(\sigma)\right) \left|\mathcal{P}'_\Gamma(\sigma)\right| \right) d\sigma \end{aligned} \tag{2.3}$$

Now, if f is differentiable on Ω, we have, with $\nabla f(\mathbf{x})$ the gradient of f at $\mathbf{x}$:

$$f(\mathbf{x} + \epsilon\, v(\mathbf{x})) = f(\mathbf{x}) + \langle \epsilon\, v\,(\mathbf{x}) \,|\nabla f\,(\mathbf{x})\rangle + o\,(\epsilon)$$

and therefore:

$$f\left(\mathcal{P}_{\Lambda_{\epsilon,v}}(\sigma)\right) = f\left(\mathcal{P}_\Gamma(\sigma)\right) + \epsilon \left\langle v(\mathcal{P}_\Gamma(\sigma) \mid \nabla f(\mathcal{P}_\Gamma(\sigma)) \right\rangle + o(\epsilon) \tag{2.4}$$

Besides ($\mathcal{P}'_\Gamma(\sigma)$ is the derivative of $\mathcal{P}_\Gamma$ at σ *on the shape*):

$$\begin{aligned} \mathcal{P}'_{\Lambda_{\epsilon,v}}(\sigma) &= \left[\mathcal{P}_\Gamma(\sigma) + \epsilon\, v(\mathcal{P}_\Gamma(\sigma))\right]' \\ &= \mathcal{P}'_\Gamma(\sigma) + \epsilon\, v'\left(\mathcal{P}_\Gamma(\sigma)\right) \cdot \mathcal{P}'_\Gamma(\sigma) \\ \left|\mathcal{P}'_{\Lambda_{\epsilon,v}}(\sigma)\right| &= \left|\mathcal{P}'_\Gamma(\sigma) + \epsilon\ \langle v'\left(\mathcal{P}_\Gamma(\sigma)\right) | \mathcal{P}'_\Gamma(\sigma)\rangle\ \right| \\ &= \left|\mathcal{P}'_\Gamma(\sigma)\right| + \epsilon \left\langle \frac{\mathcal{P}'_\Gamma(\sigma)}{|\mathcal{P}'_\Gamma(\sigma)|} | v'\left(\mathcal{P}_\Gamma(\sigma)\right) \cdot \mathcal{P}'_\Gamma(\sigma) \right\rangle + o(\epsilon) \\ &= \left|\mathcal{P}'_\Gamma(\sigma)\right| + \epsilon \left\langle \overrightarrow{T}\, | v'\left(\mathcal{P}_\Gamma(\sigma)\right) \cdot \mathcal{P}'_\Gamma(\sigma) \right\rangle + o(\epsilon) \end{aligned} \tag{2.5}$$

where $\overrightarrow{T} = \frac{\mathcal{P}'_\Gamma(\sigma)}{|\mathcal{P}'_\Gamma(\sigma)|}$ is called the *unit tangent vector* to the shape Γ at $\mathcal{P}'_\Gamma(\sigma)$.

Introducing 2.4 and 2.5 in 2.3, we obtain:

$$\begin{aligned} \frac{E(\Lambda_{\epsilon,v}) - E(\Gamma)}{\epsilon} = &\int_0^1 f\left(\mathcal{P}_\Gamma(\sigma)\right) \left\langle \overrightarrow{T} \mid v'\left(\mathcal{P}_\Gamma(\sigma)\right) \cdot \mathcal{P}'_\Gamma(\sigma) \right\rangle d\sigma \ + \\ &\int_0^1 \left\langle \nabla f\left(\mathcal{P}_\Gamma(\sigma)\right) \mid v\left(\mathcal{P}_\Gamma(\sigma)\right) \right\rangle \left|\mathcal{P}'_\Gamma(\sigma)\right| d\sigma \ + \ o(1) \end{aligned} \tag{2.6}$$

The first term of 2.6 can be rewritten as:

$$\int_0^1 \left\langle f\left(\mathcal{P}_\Gamma(\sigma)\right)\ \overrightarrow{T}\ \middle|\ \left[v\left(\mathcal{P}_\Gamma(\sigma)\right)\right]' \right\rangle d\sigma$$

We now apply the method of *integration by parts*, and using the fact that we are considering closed curves we can remove the first term and write:

$$\begin{aligned}
&= -\int_0^1 \left\langle \left[f\left(\mathcal{P}_\Gamma(\sigma)\right)\ \overrightarrow{T}\right]' \ \middle|\ v\left(\mathcal{P}_\Gamma(\sigma)\right) \right\rangle d\sigma \\
&= -\int_0^1 \left\langle \left[\nabla f\left(\mathcal{P}_\Gamma(\sigma)\right)\mathcal{P}'_\Gamma(\sigma)\overrightarrow{T} + f\left(\mathcal{P}_\Gamma(\sigma)\right)\kappa\overrightarrow{N}|\mathcal{P}'_\Gamma(\sigma)|\right] \ \middle|\ v\left(\mathcal{P}_\Gamma(\sigma)\right) \right\rangle d\sigma \\
&= -\int_0^1 \left\langle \left[\left\langle \nabla f\left(\mathcal{P}_\Gamma(\sigma)\right)\middle||\mathcal{P}'_\Gamma(\sigma)|\overrightarrow{T}\right\rangle \overrightarrow{T} + f\left(\mathcal{P}_\Gamma(\sigma)\right)\kappa\overrightarrow{N}|\mathcal{P}'_\Gamma(\sigma)|\right] \ \middle|\ v\left(\mathcal{P}_\Gamma(\sigma)\right) \right\rangle d\sigma \\
&= -\int_0^1 \left\langle \left[\nabla f_{\overrightarrow{T}} + f\left(\mathcal{P}_\Gamma(\sigma)\right)\kappa\overrightarrow{N}\right] \ \middle|\ v\left(\mathcal{P}_\Gamma(\sigma)\right) \right\rangle \left|\mathcal{P}'_\Gamma(\sigma)\right| d\sigma
\end{aligned}$$

where $\kappa(\mathbf{x})$ is the curvature of the shape at $\mathbf{x}$, $\overrightarrow{N}(x)$ the unit normal vector to the shape and $\nabla f_{\overrightarrow{T}}$ the tangential component of ∇f, that is to say:

$$\nabla f_{\overrightarrow{T}} = \left\langle \nabla f\left(\mathcal{P}_\Gamma(\sigma)\right)\middle|\overrightarrow{T}\right\rangle \overrightarrow{T}$$

And finally, going back to 2.6, we have:

$$\frac{E(\Lambda_{\epsilon,v}) - E(\Gamma)}{\epsilon} = \int_0^1 \left\langle \nabla f_{\overrightarrow{N}} - f\left(\mathcal{P}_\Gamma(\sigma)\right)\kappa\overrightarrow{N} \mid v\left(\mathcal{P}_\Gamma(\sigma)\right) \right\rangle \left|\mathcal{P}'_\Gamma(\sigma)\right| d\sigma \ + \ o(1)$$

where $\nabla f_{\overrightarrow{N}}$ is the normal component of ∇f:

$$\begin{aligned}
\nabla f_{\overrightarrow{N}} &= \left\langle \nabla f\left(\mathcal{P}_\Gamma(\sigma)\right)\middle|\overrightarrow{N}\right\rangle \overrightarrow{N} \\
&= \nabla f\left(\mathcal{P}_\Gamma(\sigma)\right) - \left\langle \nabla f\left(\mathcal{P}_\Gamma(\sigma)\right)\middle|\overrightarrow{T}\right\rangle \overrightarrow{T}
\end{aligned}$$

And after all:

$$\boxed{\nabla E(\Gamma)(\mathbf{x}) = \left(\left\langle \nabla f(\mathbf{x})|\overrightarrow{N}(\mathbf{x})\right\rangle - f(\mathbf{x})\kappa(\mathbf{x})\right)\overrightarrow{N}(\mathbf{x})} \tag{2.7}$$

In conclusion, in addition of computing the general form of the gradient of these kind of energy, we prove that for an energy functional based only on the geometrical form of the shape, the shape gradient *lies only on the normal*

vector to the shape (still considering only the L^2 inner product). It means also that the tangential part of a velocity field v *does not have any influence* on the variation of the energy and therefore on the geometrical form on the shape. We will see in chapter 3 that for instance, if we track some landmark points during the evolution and take into account the position of these points in the energy, this property does not hold as expected.

2.5 Numerical experiments

For our experiments we chose to use an implicit representation of the shapes by placing ourself in the *level set framework* [39, 108]. The reader will find a more circumstantial presentation of this method in chapter 4, section 4.3.1. As the energy functional $E(\cdot)$ we chose the square of one of two of the metrics introduced in chapter 1, approximated (see the same chapter) to be differentiable. We therefore computed the gradients of these energies with respect to the shape and apply the gradient descent scheme given by equation 2.2.

2.5.1 Computing the gradient of the approximation to the Hausdorff distance

In [22], authors show that the approximation $\rho_H(\Gamma, \Gamma_0)$ of the Hausdorff distance $d_H(\Gamma, \Gamma_0)$ presented in section 1.4, Def. 5, is differentiable with respect to Γ and compute its gradient $\nabla \rho_H(\Gamma, \Gamma_0)$ in the two-dimensional case. We extended this computation in the three-dimensional case.

To simplify notations we rewrite the expression of this approximation as

$$\rho_H(\Gamma, \Gamma_0) = \left\lfloor \left\lfloor \lfloor d(\cdot,\cdot) \rfloor_{\Gamma_0}^{\varphi} \right\rfloor_{\Gamma}^{\psi}, \left\lfloor \lfloor d(\cdot,\cdot) \rfloor_{\Gamma}^{\varphi} \right\rfloor_{\Gamma_0}^{\psi} \right\rfloor^{\theta} \tag{2.8}$$

and state the result, the reader interested in the proof being referred to [23]

Proposition 9. *For Γ and Γ_0 two curves embedded in $\mathbb{R}^2$ or two surfaces embedded in $\mathbb{R}^3$, the gradient of $\rho_H(\Gamma, \Gamma_0)$ with respect to Γ at any point $\mathbf{y}$ of Γ is given by :*

$$\nabla \rho_H(\Gamma, \Gamma_0)(\mathbf{y}) = \frac{1}{\theta'(\rho_H(\Gamma, \Gamma_0))} \left(\alpha(\mathbf{y})\kappa(\mathbf{y}) + \beta(\mathbf{y})\right), \tag{2.9}$$

where $\kappa(\mathbf{y})$ is the curvature (in the two-dimensional case) or the mean curvature (in the three-dimensional case) of Γ at $\mathbf{y}$ and the functions $\alpha(\mathbf{y})$ and

$\beta(\mathbf{y})$ *are given by*

$$\alpha(\mathbf{y}) = \nu \int_{\Gamma_0} \frac{\psi'}{\varphi'}\left(\lfloor d(\mathbf{x},\cdot)\rfloor_\Gamma^\varphi\right)\ \left[\ \varphi\circ\lfloor d(\mathbf{x},\cdot)\rfloor_\Gamma^\varphi - \varphi\circ d(\mathbf{x},\mathbf{y})\ \right] d\Gamma_0(\mathbf{x}) + |\Gamma_0|\eta\ \left[\psi\left(\left\lfloor\lfloor d(\cdot,\cdot)\rfloor_{\Gamma_0}^\varphi\right\rfloor_\Gamma^\psi\right) - \psi\left(\lfloor d(\cdot,\mathbf{y})\rfloor_{\Gamma_0}^\varphi\right)\ \right], \quad (2.10)$$

$$\beta(\mathbf{y}) = \int_{\Gamma_0} \varphi'\circ d(\mathbf{x},\mathbf{y})\ \left[\nu\frac{\psi'}{\varphi'}\left(\lfloor d(\mathbf{x},\cdot)\rfloor_\Gamma^\varphi\right) + \eta\frac{\psi'}{\varphi'}\left(\lfloor d(\cdot,\mathbf{y})\rfloor_{\Gamma_0}^\varphi\right)\right]\ \frac{\mathbf{y}-\mathbf{x}}{d(\mathbf{x},\mathbf{y})}\cdot\overrightarrow{N}(\mathbf{y})\ d\Gamma_0(\mathbf{x}), \quad (2.11)$$

where $\nu = \dfrac{1}{|\Gamma|\ |\Gamma_0|}\dfrac{\theta'}{\psi'}\left(\left\lfloor\lfloor d(\cdot,\cdot)\rfloor_\Gamma^\varphi\right\rfloor_{\Gamma_0}^\psi\right)$ *and* $\eta = \dfrac{1}{|\Gamma|\ |\Gamma_0|}\dfrac{\theta'}{\psi'}\left(\left\lfloor\lfloor d(\cdot,\cdot)\rfloor_{\Gamma_0}^\varphi\right\rfloor_\Gamma^\psi\right)$.

Note that the function $\beta(y)$ is well-defined even if $\mathbf{y}$ belongs to Γ_0 since the term $\frac{\mathbf{y}-\mathbf{x}}{d(\mathbf{x},\mathbf{y})}$ is of unit norm.

2.5.2 Computation of the gradient of the approximation to the $W^{1,2}$ norm

The gradient $\nabla\rho_{W^{1,2}}(\Gamma,\Gamma_0)$, of our approximation $\rho_{W^{1,2}}(\Gamma,\Gamma_0)$ of the distance $d_{W^{1,2}}(\Gamma,\Gamma_0)$ presented in section 1.4, Def. 6 can also be computed. We simply state the result:

Proposition 10. *The gradient of* $\rho_{W^{1,2}}(\Gamma,\Gamma_0)$ *at any point* $\mathbf{y}$ *of* Γ *is given by*

$$\nabla\rho_{W^{1,2}}(\Gamma,\Gamma_0)(\mathbf{y}) = \int_\Omega \left[B(\mathbf{x},\mathbf{y})\left(C_1(\mathbf{x}) - \frac{\varphi''}{\varphi'}\left(\left\lfloor\tilde{d}(\mathbf{x},\cdot)\right\rfloor_\Gamma^{\varphi_\alpha}\left(\mathbf{C}_2(\mathbf{x})\cdot\nabla\left\lfloor\tilde{d}(\mathbf{x},\cdot)\right\rfloor_\Gamma^{\varphi_\alpha}\right)\right) + \mathbf{C}_2(\mathbf{x})\cdot\nabla B(\mathbf{x},\mathbf{y})\right] d\mathbf{x}, \quad (2.12)$$

where

$$B(\mathbf{x},\mathbf{y}) = \kappa(\mathbf{y})\left(\lfloor\varphi\circ d(\mathbf{x},\cdot)\rfloor_\Gamma - \varphi\circ d(\mathbf{x},\mathbf{y})\right) + \varphi'(d(\mathbf{x},\mathbf{y}))\frac{\mathbf{y}-\mathbf{x}}{d(\mathbf{x},\mathbf{y})}\cdot\overrightarrow{N}(\mathbf{y}),$$

$\kappa(\mathbf{y})$ *is the curvature of* Γ *at* $\mathbf{y}$,

$$C_1(\mathbf{x}) = \frac{1}{|\Gamma|\ \varphi'\left(\left\lfloor\tilde{d}(\mathbf{x},\cdot)\right\rfloor_\Gamma^{\varphi_\alpha}\right)}\left\|\left\lfloor\tilde{d}(\right\rfloor_\Gamma^{\varphi_\alpha} - \left\lfloor\tilde{d}(\right\rfloor_{\Gamma_0}^{\varphi_\alpha}\right\|_{L^2(D)}^{-1}\ \left(\left\lfloor\tilde{d}(\mathbf{x},\cdot)\right\rfloor_\Gamma^{\varphi_\alpha} - \left\lfloor\tilde{d}(\mathbf{x},\cdot)\right\rfloor_{\Gamma_0}^{\varphi_\alpha}\right),$$

and

$$\mathbf{C}_2(x) = \frac{1}{|\Gamma|\ \varphi'\left(\left\lfloor\tilde{d}(\mathbf{x},\cdot)\right\rfloor_\Gamma^{\varphi_\alpha}\right)}\left\|\left\lfloor\tilde{d}(\right\rfloor_\Gamma^{\varphi_\alpha} - \left\lfloor\tilde{d}(\right\rfloor_{\Gamma_0}^{\varphi_\alpha}\right\|_{\mathbf{L}^2(D)}^{-1}\ \nabla\left(\left\lfloor\tilde{d}\right\rfloor_\Gamma^{\varphi_\alpha} - \left\lfloor\tilde{d}\right\rfloor_{\Gamma_0}^{\varphi_\alpha}\right)(\mathbf{x}),$$

2.5.3 Direct minimization of the $W^{1,2}$ norm

An alternative to the method presented in the previous section is to evolve not the curve Γ but its distance function d_Γ. Minimizing $d_{W^{1,2}}(\Gamma, \Gamma_0)$ with respect to d_Γ implies computing the corresponding Euler-Lagrange equation EL. The reader will verify that the result is

$$EL = \frac{d_\Gamma - d_{\Gamma_0}}{\|d_\Gamma - d_{\Gamma_0}\|_{L^2(D)}} - div\left(\frac{\nabla\left(d_\Gamma - d_{\Gamma_0}\right)}{\|\nabla(d_\Gamma - d_{\Gamma_0})\|_{\mathbf{L}^2(D)}}\right) \tag{2.13}$$

To simplify notations we now use d instead of d_Γ. The problem of warping Γ_1 onto Γ_0 is then transformed into the problem of solving the following PDE

$$\begin{aligned} d_t &= -EL \\ d(0,\cdot) &= d_{\Gamma_1}(\cdot). \end{aligned}$$

The problem that this PDE does not preserve the fact that d is a distance function is alleviated by "reprojecting" at each iteration the current function d onto the set of distance functions by running a few iterations of the "standard" restoration PDE [133]

$$\begin{aligned} d_t &= (1 - |\nabla d|)sign(d) \\ d(0,\cdot) &= d_0 \end{aligned}$$

This method has the advantage to be easier and faster to compute but as stressed above, nothing insure us that the family of shapes resulting from this scheme will be the same as the one obtained by a actual gradient descent.

2.5.4 Examples of warping

Fig. 2.3 and Fig. 2.4 show few steps of the result of the warping of one two-dimensional shape into another one. The first one (Fig. 2.3) uses the square of the d_{L^2} shape distance (see Def. 1) as energy functional, and the second one (Fig. 2.4) the square of the $d_{W^{1,2}}$ shape distance (see Def. 2). The effect of the second term in the W^{12} norm, the L^2-norm between the gradient of the distance function, is clear: the gradient descent scheme aim to align the direction of the gradient of the distance function while aligning the value of the function themselves.

Fig. 2.5 shows few steps of the result of the warping of one two-dimensional shape of a rabbit into another one. These silhouettes have been borrowed from a database of several images used as a benchmark for image-indexing [125]. Here the energy functional is chosen to be the square of the approximation of the Hausdorff distance.

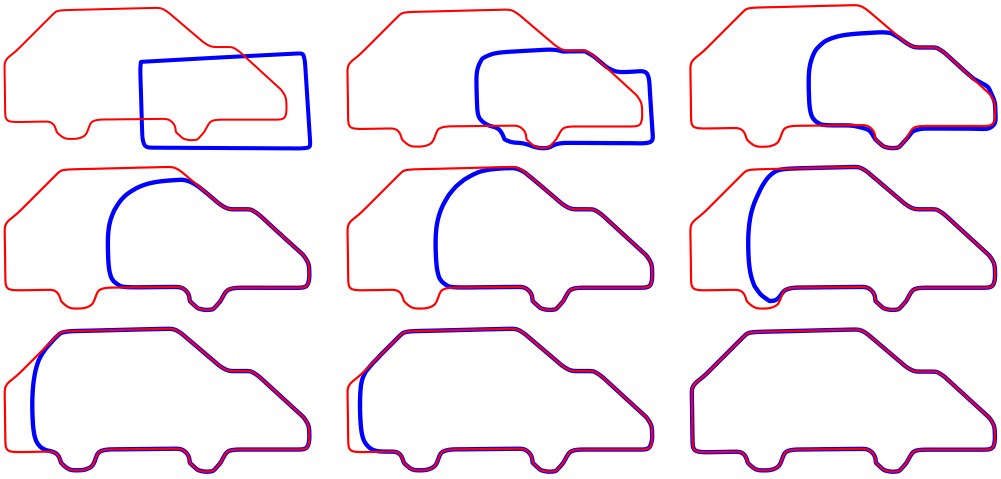


Figure 2.3: L^2 warping of the red shape into the blue one

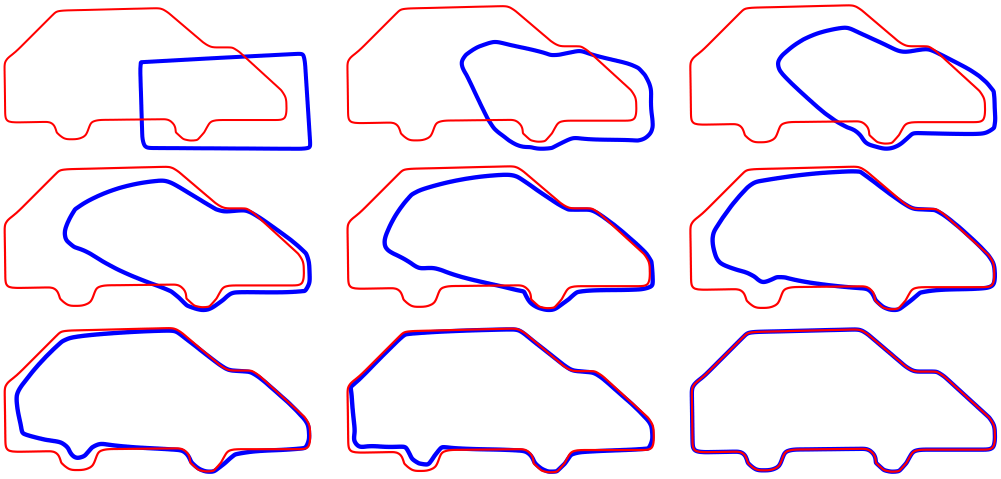


Figure 2.4: $W^{1,2}$ warping of the red shape into the blue one

Fig. 2.6 exemplifies the difference between the L^2 gradient descent scheme and the Hausdorff one. The top row shows few steps of the result of the L^2 warping: the branch of the red shape "collapses on itself". It is naturally because the distance function of the red shape inside the top of its branch is negative and wants to be positive (since it is located outside the blue shape). The bottom row shows few steps of the result of the Hausdorff warping: the points on the top of the red branch are the points farthest to the blue shape and therefore they are the points concerned by the definition of the Hausdorff

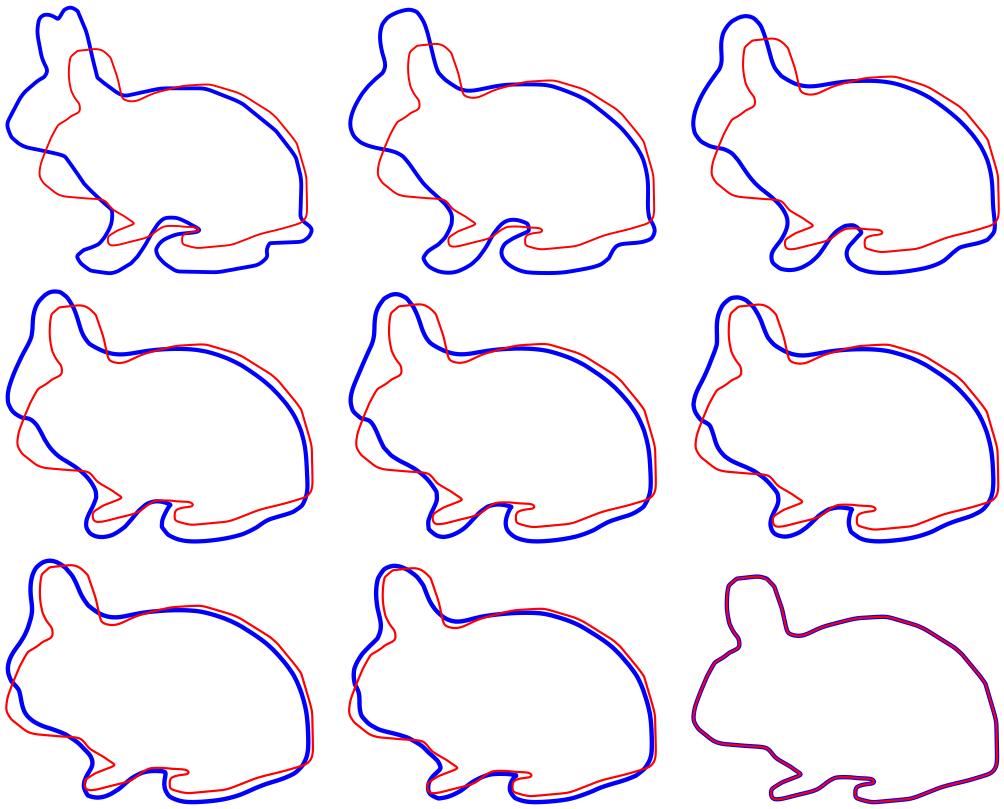


Figure 2.5: Hausdorff warping of the red rabbit into the blue one (shapes of the rabbits come from [125])

distance and will tend to come closer to the blue shape.

Finally Fig. 2.7 shows an example of a warping resulting from a Hausdorff gradient descent scheme in the case of two surfaces embedded in $\mathbb{R}^3$. The initial shape is a teddy bear and it is evolving to Hayao Miyazaki's character Totoro.

2.5.5 Definition and examples of the mean of n shapes

As an example of what can be done with shape metrics and shape warping, we will introduce the notion of the *empirical mean of n shapes* $\Gamma_1, \ldots, \Gamma_n$. Inspired by M. Frechet [57] and H. Karcher [73], we can provide the (classical) following definition:

Definition 11. *Let $d(\cdot, \cdot)$ be a distance on the set of shapes and $\Gamma_1, \ldots, \Gamma_n$, n static shapes. Their empirical mean is defined as any shape $\hat{\Gamma}$ that achieves*

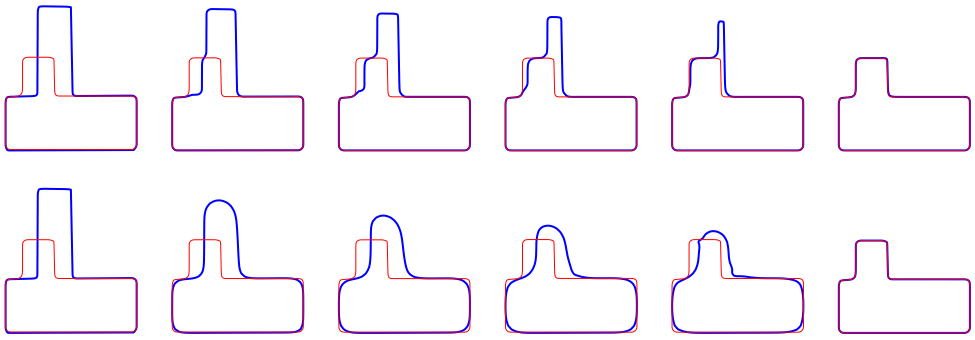


Figure 2.6: Difference between the L^2 warping (*top row*) and the Hausdorff warping (*bottom row*)

a local minimum of the function μ*:*

$$\begin{aligned} \mu : \quad \mathcal{S} &\rightarrow \mathbb{R}^+ \\ \Gamma &\mapsto \frac{1}{n}\sum_{i=1}^{n} d(\Gamma, \Gamma_i)^2 \end{aligned}$$

An empirical mean can then be computed by a gradient descent minimizing $\mu(\Gamma)$. The gradient of this energy is straightforward obtained from the expression of the gradient of the distance $d(\cdot, \cdot)$. Fig. 2.8 shows some examples of empirical mean of shapes. Several works have been done starting from this notion of the mean shape. For example, in [22], authors defined the covariance of a set of shapes. In [95] and in collaboration with Guillermo Sapiro, we extended the notion of mean for *dynamic shapes* using an adaptation of *dynamic time warping* algorithm (often used for speech recognition). In [50] authors used the mean of shapes to define an interpolation between two shapes and then defined the projection of a shape on a manifold in the $\mathcal{S}$ the set of shapes.

2.6 Conclusion

This chapter was dedicated to the derivation of shape-relevant metrics. We introduced the concept of Gateaux derivatives as well as shape gradients toward shape warping using different similarity metrics presented in the previous chapter. This question of deforming a shape onto another one is potentially interesting for every problem of matching (point-to-point) between shapes such as for example shape recognition or the tracking of natural mov-

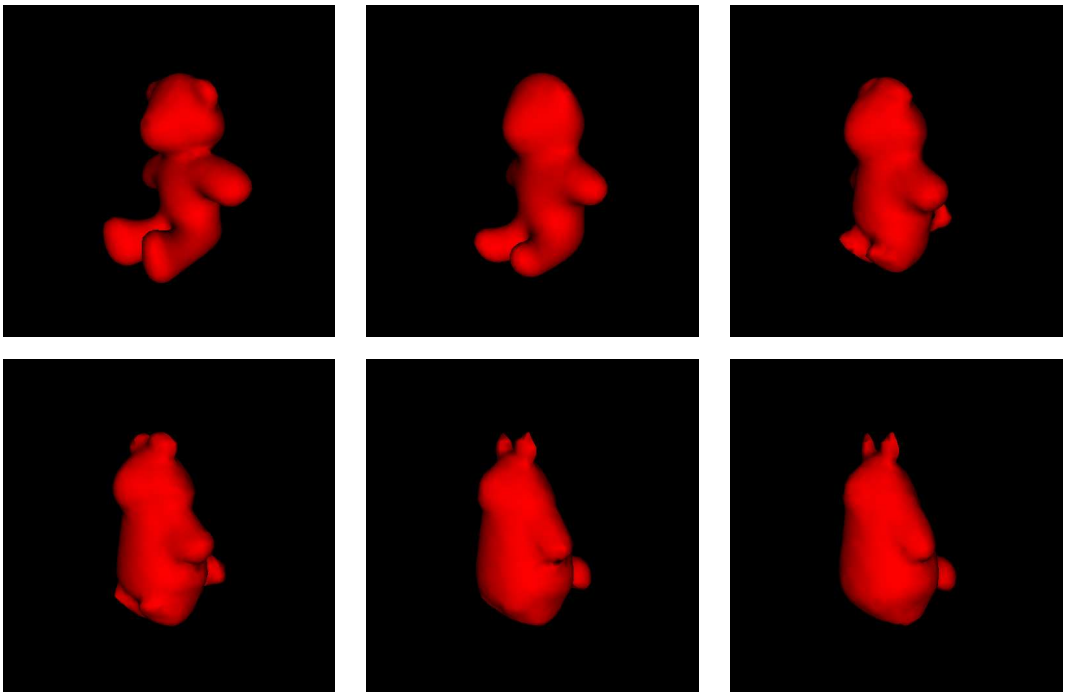


Figure 2.7: Hausdorff warping a closed surface to another one.

ing objects. It enables the comparison between a shape model and an unknown object. Moreover, tracking points during the evolution gives us a matching (not necessarily one-to-one, see chapter 4) between the shapes at no additional cost. In our calculations of shape gradients, we considered only closed curves. Considering an open curve would induce the apparition of Dirac peaks on the extremities of the curve in the expression of the gradient. One solution of this issue could be to use the Sobolev inner product (see chapters 3 and 4) to transform these Dirac peaks in a smooth kernel. Investigating in this direction can probably lead to interesting further developments.

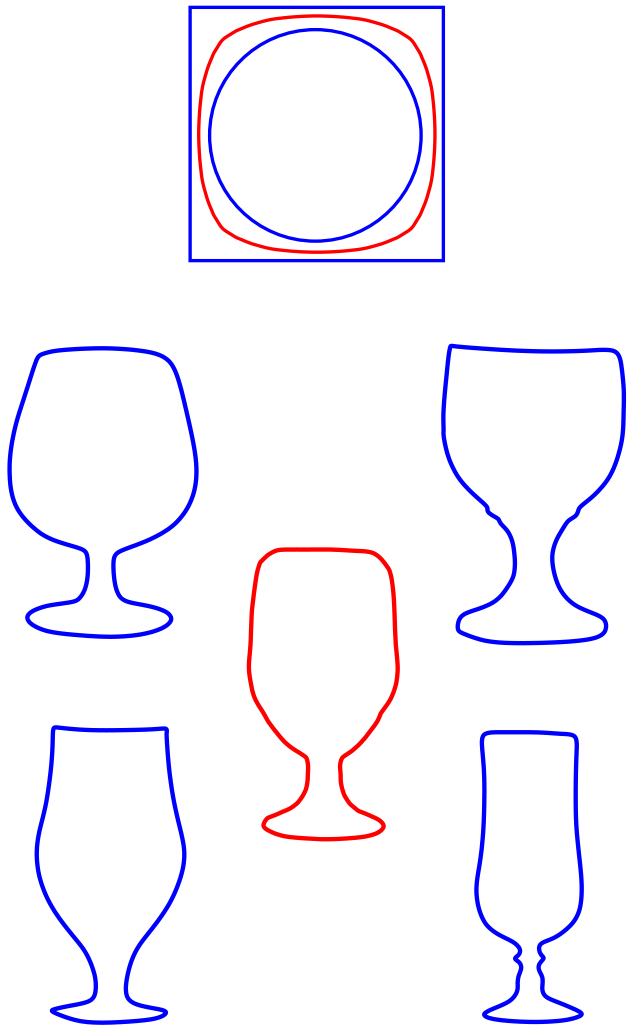


Figure 2.8: Examples of mean: in blue the initial shapes and in red their mean. The distance used is the $d_{W^{1,2}}$

Chapter 3

Generalized gradient: priors on minimization flows

This chapter tackles an important aspect of the variational problem underlying active contours: optimization by gradient flows. As mentioned in chapter 2, the definition of a gradient depends directly on the choice of an inner product structure. This consideration is largely absent from the active contours literature. Most authors, explicitly or implicitly, assume that the space of admissible deformations is ruled by the canonical L^2 inner product. The classical gradient flows reported in the literature are relative to this particular choice. Here, we investigate the relevance of using other inner products, yielding other gradient descents, and moreover, other minimizing flows not deriving from any inner product. We report numerical experiments indicating that the sensitivity of the active contours method to initial conditions, which seriously limits its applicability and efficiency, is alleviated by our application-specific spatially coherent minimizing flows. We show that the choice of the inner product can be seen as a prior on the deformation fields and we present an extension of the definition of the gradient toward more general priors.

This is a joint work with Guillaume Charpiat and Jean-Philippe Pons. The example of the spatially coherent minimizing flows in section 3.4 have been first introduced by G.Charpiat, J.P.Pons et al. in [24]. The generalized gradient method in section 3.6 is a joint work with G.Charpiat.

Publication related to this chapter: *Generalized gradients : priors on minimization flows* [26] with Guillaume Charpiat, Jean-Philippe Pons, Renaud Keriven and Olivier Faugeras, published in the *International Journal of Computer Vision*, 2007.

3.1 Introduction

Many problems in computer vision can advantageously be cast in a variational form, i.e. as a minimization of an energy functional. In this chapter, we focus on variational methods dedicated to the recovery of contours. In this case, the problem consists in finding a shape which corresponds to a global minimum of the energy. Unfortunately, in most cases, the exact minimization of the energy functional is computationally unfeasible due to the huge number of unknowns. Hence, in most cases, a suboptimal strategy must be adopted. A common minimization procedure consists in evolving an initial contour, positioned by the user or resulting from a preprocessing , in the direction of steepest descent of the energy.

This approach, known in the literature as *active contours* or *deformable models*, was pioneered by Kass. et al. in [74] for the purpose of image segmentation. Since then, it has been applied in many domains of computer vision and image analysis, such as image segmentation [20], surface reconstruction [149, 46], stereo reconstruction [54, 71, 61]. Active contours have first been used conjointly with edge-based energies (e.g. [20, 91, 77]) and region-based energies (e.g. [103, 150, 146, 111, 21, 112]). More recently, prior-based energies (e.g. [88, 32, 120, 119]) and energies based on more complex geometrical information (e.g. [78, 104]) have been introduced. This complexification of the energies can be partly seen as a search of a way to avoid unwanted local minima during gradient descent evolution.

In fact, due to the highly non-convex nature of most energy functionals, a gradient descent flow is very likely to be trapped in a local minimum. Also, this local minimum depends on the position of the initial contour. If the latter is far from the expected final configuration, the evolution may be trapped in a completely irrelevant state. This sensitivity to initial conditions seriously limits the applicability and efficiency of the active contours method. A *balloon force* that can either inflate or deflate the model was proposed by L.Cohen in [27] to reduce the requirement to initialize the model near the desired object boundaries. Furthermore there have been efforts in investigating new optimization method that can obtain directly a global minimum of an energy. For example, the minimal path technique [28] was designed to find the global minimal solution of the edge-based energy considered in [20]. More recently the *graph cuts* method has been investigated. It is a powerful energy minimization method which allows to find a *global minimum* or a strong local minimum of an energy. In the last few years, this method has been successfully applied to several problems in computer vision, including stereovision [79] and image segmentation [18]. However, these global methods have a severe limitation: it cannot be applied to an arbitrary energy

function [80], and, when applicable, it can be computationally expensive (for the 3D case for example).

In this chapter we present a framework to specify some constraints on the gradient and therefore on the minimization path induced by the gradient descent. In the first place, we show that it can be done by carefully choosing the inner product involved in the gradient definition. And later we show that we can even define new minimization flows that we call *generalized gradients* which do not derive from any inner product. This may be a way to avoid some unwanted local minima but also to insure that the family of shapes resulting from the minimization verifies some properties (e.g. spatial coherence during the evolution).

3.2 Minimization and inner product

In chapter 2 we defined the Gâteaux derivatives of an energy functional $E(\Gamma)$ by:

$$\mathcal{G}_\Gamma(E(.), v) = \lim_{\epsilon \to 0} \frac{E(\Gamma + \epsilon v) - E(\Gamma)}{\epsilon} \ .$$

In order to apply a gradient descent method, we then would like to pick the gradient as the direction of steepest descent of the energy. However, it is not yet possible at this stage: to be able to assess the steepness of the energy, the deformation space needs additional structure, namely an inner product introducing the geometrical notions of angles and lengths. This consideration is largely absent from the active contours literature: most authors, explicitly or implicitly, assume that the deformation space is ruled by the canonical L^2 inner product on Γ, which is, for two deformation fields u and v in $\mathscr{D}$:

$$\langle u \,|v\rangle_{L^2} = \frac{1}{|\Gamma|} \int_\Gamma u(\mathbf{x}) \cdot v(\mathbf{x}) \, d\Gamma(\mathbf{x}) \ ,$$

where $d\Gamma(\mathbf{x})$ stands for the area element of the contour so that the integral over Γ is intrinsic and does not depend on the parametrization (see chapter 1, Def 4).

Here, for sake of generality, we model the space of admissible deformations as an inner product space $(\mathscr{D}, \langle|\rangle_{\mathscr{D}})$ (not necessarily the L^2 one). If there exists a deformation field $u \in \mathscr{D}$ such that

$$\forall v \in \mathscr{D}, \ \mathcal{G}_\Gamma(E(.), v) = \langle u \,|v\rangle_{\mathscr{D}} \ ,$$

then u is unique, we call it the gradient of E *relative to the inner product* $\langle|\rangle_{\mathscr{D}}$, and we denote by $u = \nabla_{\mathscr{D}} E(\Gamma)$. The existence of u is related to the

smoothness of E, or more exactly to the continuity of $\mathcal{G}_\Gamma(E(.), v)$ with respect to v (Riesz representation theorem, see [121] for more details).

Clearly, each choice of inner product yields its own gradient. This is often neglected and most authors improperly refer to *the* gradient of the energy. Thus, the classical gradient flows reported in the literature (mean curvature flow, geodesic active contours [20, 62, 128], multi-view 3D reconstruction [54, 71, 61]) are relative to the L^2 inner product.

The gradient descent method consists in deforming an initial contour Γ_0 in the opposite direction of the gradient.

$$\begin{cases} \Gamma(0) = \Gamma_0 \\ \dfrac{d\Gamma}{dt} = -\nabla_{\mathscr{D}} E(\Gamma) \end{cases} \tag{3.1}$$

The problem of the existence and the uniqueness of this minimizing flow is out of the scope of this chapter. Indeed, it is highly dependent on the properties of each particular energy functional. If this evolution exists, it decreases the energy:

$$\frac{dE(\Gamma)}{dt} = - \left\| \nabla_{\mathscr{D}} E(\Gamma) \right\|^2_{\mathscr{D}} \leq 0 \, .$$

The standard choice for $\mathscr{D}$ is the Hilbert space of square integrable velocity fields $L^2(\Gamma, \mathbb{R}^n)$ equipped with its canonical inner product. Very few authors in the active contours area have considered using other inner products, whereas this is an established technique in image registration [137]. More recently, in the context of shape representation and analysis, [98, 131] have shown that slightly modifying the L^2 inner product allows to build well-behaved metrics in the space of curves. The particular case of the H^1 inner product has been simultaneously and independently investigated by us [24, 93] and by Sundaramoorthi et al. [130, 132].

Let us notice that the variations on the gradient descent theme, as in [14] (e.g. the conjugate gradient method, Newton's or quasi-Newton's methods), will still be applicable to the new gradients we propose, since these methods are in fact not specific to the particular L^2 gradient.

Minimizing flows not deriving from any inner product, that is to say evolutions that decrease the energy, without any gradient interpretation, have also been overlooked so far. Note that any evolution fulfilling the condition

$$\frac{dE(\Gamma)}{dt} = \left\langle \nabla_{\mathscr{D}} E(\Gamma) \middle| \frac{d\Gamma}{dt} \right\rangle_{\mathscr{D}} \leq 0 \tag{3.2}$$

is a candidate to solve the minimization problem. This idea, proposed in [128], is applied by the same authors to the alignment of curve in images in [109]: a complicated term in the gradient is safely neglected after checking that the evolution still decreases the energy.

The spirit of our work is different. We do not focus either on a specific inner product or on a particular energy functional. We rather explore general procedures to build some new inner products and to compute the associated gradients. We also address the design of non-gradient minimizing flows.

Our motivation is also different. Our primary concern in this work is the sensitivity of the active contours method to initial conditions. There are essentially two ways of dealing with this problem: positioning the initial contour very close to the expected final configuration, or using a multiresolution coarse-to-fine strategy, in other words running the optimization on a series of smoothed and subsampled contours and input data. In this chapter, we pioneer a third way to tackle the problem of unwanted local minima: the careful design of the minimizing flow.

We do not modify the energy, hence the relief of the energy landscape and in particular the "number" of local minima remains unchanged. But by using an evolution that favors certain types of directions, we expect to decrease the probability of falling into unwanted energy basins. Typically, in many applications, spatially coherent motions are to be preferred over erratic evolutions. For example, in the tracking problem, the object of interest is likely to have similar shapes in consecutive frames. So if we initialize the contour with the result of the previous frame, it makes sense to encourage the motions which preserve its overall appearance. This way, it may be easier to dodge unexpected local low-energy configurations. A traditional L^2 gradient descent definitely does not have this desirable property since the L^2 inner product completely disregards the spatial coherence of the velocity field.

3.3 New Inner Products and New Flows

In this section, we suppose that the space $\mathscr{D}$ of all admissible deformations of the shape Γ is initially equipped with the inner product $\langle | \rangle_{\mathscr{D}}$, for example in the standard case we would have $\mathscr{D} = L^2$, and we study how to build new inner products or new minimizing flows from the given one.

3.3.1 Designing new inner products

Definition 12. *For any symmetric positive definite linear operator $\mathcal{L} : \mathscr{D} \to \mathscr{D}$, a new inner product can be defined by*

$$\langle u \,|v\rangle_{\mathcal{L}} = \langle \mathcal{L}u \,|v\rangle_{\mathscr{D}} \ . \tag{3.3}$$

Here, for simplicity, we assume that the domain and the range of $\mathcal{L}$ are equal to $\mathscr{D}$. A similar study is possible if they are strictly smaller than $\mathscr{D}$, under certain conditions, using the Friedrichs extension of $\mathcal{L}$ (see [4] for details). But these technical details are out of the scope of this chapter.

The following observation is central to our work:

Proposition 13. *If $\nabla_{\mathscr{D}} E(\Gamma)$ exists and if $\mathcal{L}$ is also invertible, then $\nabla_{\mathcal{L}} E(\Gamma)$ also exists and we have*

$$\nabla_{\mathcal{L}} E(\Gamma) = \mathcal{L}^{-1} \left(\nabla_{\mathscr{D}} E(\Gamma)\right) \ . \tag{3.4}$$

Proof. Indeed:

$$\begin{aligned} \forall v \in \mathscr{D},\ \mathcal{G}_\Gamma(E(.), v) &= \langle \nabla_{\mathscr{D}} E(\Gamma) \,|v\rangle_{\mathscr{D}} \\ &= \left\langle \mathcal{L}\mathcal{L}^{-1} \nabla_{\mathscr{D}} E(\Gamma) \,|v\right\rangle_{\mathscr{D}} \\ &= \left\langle \mathcal{L}^{-1} \nabla_{\mathscr{D}} E(\Gamma) \,|v\right\rangle_{\mathcal{L}} \ . \end{aligned}$$

□

The above procedure is of great practical interest because it allows to upgrade any existing L^2 gradient flow. However, it is not completely general in the sense than all inner products cannot be expressed in this form.

Nevertheless, if $\mathscr{D}$ is a separable Hilbert space (i.e. complete with respect to the norm $\|\cdot\|_{\mathscr{D}}$), the Riesz representation theorem tells us that any inner product $\langle|\rangle_{\mathcal{L}}$ such that

$$\exists\, C > 0,\ \forall u \in \mathscr{D},\ \|u\|_{\mathcal{L}} \leqslant C \|u\|_{\mathscr{D}}$$

can be written in the form of equation (3.3). This suggests that our procedure accounts for a wide range of inner products.

3.3.2 Designing new minimizing flows

In this subsection, we follow the inverse approach. Instead of working with the inner product, we apply a linear operator $\mathcal{L} : \mathscr{D} \to \mathscr{D}$ to the gradient, and we study the properties of the resulting flow:

$$\frac{d\Gamma}{dt} = -\mathcal{L}\big(\nabla_{\mathscr{D}} E(\Gamma)\big) \ . \tag{3.5}$$

This naturally sets up a hierarchy among different types of operators:

- if $\mathcal{L}$ is positive, the energy is non-increasing along the flow (3.5). Indeed,

$$\frac{dE(\Gamma)}{dt} = -\left\langle \nabla_{\mathscr{D}} E(\Gamma) \,|\mathcal{L}\nabla_{\mathscr{D}} E(\Gamma)\right\rangle_{\mathscr{D}} \leqslant 0 \ .$$

- if $\mathcal{L}$ is positive definite, the energy strictly decreases along the flow (3.5) until a critical point of the original gradient flow (3.1) is reached.
- if $\mathcal{L}$ is symmetric positive definite and invertible, the flow (3.5) coincides with a gradient descent relative to the inner product $\langle|\rangle_{\mathcal{L}^{-1}}$, as defined in equation (3.3).

The third case is contained in Subsection 3.3.1. A useful example of the second case is given in Subsection 3.4.3.

3.3.3 Adding an orthogonal term

The rate of decrease of the energy when following the direction of descent $\frac{d\Gamma}{dt}$ is given by:

$$\frac{dE(\Gamma)}{dt} = \left\langle \nabla_{\mathscr{D}} E(\Gamma) \,\middle|\, \frac{d\Gamma}{dt} \right\rangle_{\mathscr{D}} \leqslant 0$$

In particular, for the usual evolution $\frac{d\Gamma}{dt} = -\nabla_{\mathscr{D}} E(\Gamma)$, we have:

$$\frac{dE(\Gamma)}{dt} = -\|\nabla_{\mathscr{D}} E(\Gamma)\|^2_{\mathscr{D}}$$

If we denote by v any vector field defined on Γ such as $\langle \nabla_{\mathscr{D}} E(\Gamma) \,|v\rangle_{\mathscr{D}} = 0$, then adding such a vector field v to the usual gradient descent term will not change the amount of decreased energy:

$$\frac{dE(\Gamma)}{dt} = \langle \nabla_{\mathscr{D}} E(\Gamma) \,|-\nabla_{\mathscr{D}} E(\Gamma) + v\rangle_{\mathscr{D}} = -\|\nabla_{\mathscr{D}} E(\Gamma)\|^2_{\mathscr{D}}$$

so we can choose the field v as we like to add to the initial gradient. Rather than choosing $v = 0$ as usual, we could for example choose one, noted $\hat{v}$, that minimizes a regularizing criterion $R(-\nabla_{\mathscr{D}} E(\Gamma) + v)$:

$$\hat{v} \;=\; \underset{v \perp \nabla_{\mathscr{D}} E(\Gamma)}{\arg\min}\; R(-\nabla_{\mathscr{D}} E(\Gamma) + v) \tag{3.6}$$

In fact this remark still stands when the choice of the direction of descent is not the gradient itself. If we denote by u the initially proposed deformation field $\frac{d\Gamma}{dt}$, then adding a vector field which is orthogonal to the gradient $\nabla_{\mathscr{D}} E(\Gamma)$ will not change the amount of decreased energy at this step of the gradient descent (but will change the evolution):

$$\frac{dE(\Gamma)}{dt} = \langle \nabla_{\mathscr{D}} E(\Gamma) \,|{-u+v}\rangle_{\mathscr{D}} = \langle \nabla_{\mathscr{D}} E(\Gamma) \,|{-u}\rangle_{\mathscr{D}}$$

Note that the notion of being orthogonal to the gradient is independent from the chosen inner product. Indeed, if F and G are two different inner products, $\nabla_F E$ and $\nabla_G E$ the associated gradients, and $\perp_F$ and $\perp_G$ the associated notions of orthogonality, we have for any $v \in \mathscr{D}$:

$$\langle \nabla_F E(\Gamma) \,|v\rangle_F = \mathcal{G}_\Gamma(E(.), v) = \langle \nabla_G E(\Gamma) \,|v\rangle_G$$

so, consequently:

$$\begin{aligned} \langle \nabla_F E(\Gamma) \,|v\rangle_F = 0 \iff & \qquad \langle \nabla_G E(\Gamma) \,|v\rangle_G = 0 \\ \nabla_F E(\Gamma) \;\perp_F\; v \iff & \qquad \nabla_G E(\Gamma) \;\perp_G\; v\;. \end{aligned}$$

3.4 Some Spatially Coherent Minimizing Flows

This theoretical study has brought us the tools we need to better apprehend minimizing flows and build new ones. We now present some minimizing flows yielding different degrees of spatial coherence. This spatially coherent flows have been introduced by G.Charpiat, J.P.Pons et al. in [24]. We insist on the fact that this spatial coherence has nothing to do with an eventual regularity term in the energy functional. We do not modify the energy, so the regularity constraint on the contour remains unchanged. We modify the trajectory of the minimizing flow, by favoring spatially coherent motions, but this does not condition the regularity of the final contour. In the following, we sometimes use differential geometry. We refer the reader to [42] for the basic notions.

3.4.1 Motion decomposition

A simple and useful procedure, to design new inner products yielding spatially coherent flows, is to decompose the deformation space into a sum of

several mutually orthogonal linear subspaces, and to apply different penalty factors to the different types of motions. Typically, the subspaces are chosen according to an application-specific hierarchy of the motions. For example, rigid/non-rigid, affine/non-affine, etc.

We suppose that such an orthogonal (with respect to $\langle | \rangle_{\mathscr{D}}$) decomposition of the deformation space $\mathscr{D}$ into N closed linear subspaces is available:

$$\mathscr{D} = \mathscr{D}_1 \perp \mathscr{D}_2 \perp \cdots \perp \mathscr{D}_N \ .$$

Then a new inner product is derived from $\langle | \rangle_{\mathscr{D}}$ by applying the procedure of subsection 3.3.1 with

$$\mathcal{L} = \bigoplus_{i=1}^{N} \lambda_i \, \mathrm{Id}_{\mathscr{D}_i} \ ,$$

where $\forall i, \ \lambda_i > 0$. The lower is λ_i, the shorter is the norm of the velocity fields of subspace $\mathscr{D}_i$, and the stronger will be this type of motion in the new gradient flow.

Obviously, $\mathcal{L}$ is symmetric positive definite and invertible. If $\nabla_{\mathscr{D}} E$ exists, so does $\nabla_{\mathcal{L}} E$ and

$$\nabla_{\mathcal{L}} E \ = \ \sum_{i=1}^{N} \frac{1}{\lambda_i} \ \Pi_{\mathscr{D}_i} (\nabla_{\mathscr{D}} E) \ , \tag{3.7}$$

where $\Pi_{\mathscr{D}_i}$ denotes the orthogonal projection on the i^{th} subspace $\mathscr{D}_i$. Of course, if all λ_i are equal to 1, $\nabla_{\mathcal{L}} E$ coincides with $\nabla_{\mathscr{D}} E$.

We apply this general construction to two useful cases. In the first case, we decompose the velocity field into a translation, an instantaneous rotation, a rescaling motion and a non-rigid residual. In the second case, we isolate the instantaneous affine motion.

In the following, we denote by

$$\mathbf{G} = \frac{1}{|\Gamma|} \int_{\Gamma} \mathbf{x} \, d\Gamma(\mathbf{x})$$

the *center of mass* of the shape.

(a) Translation, rotation and scaling

In this paragraph, we focus on the two-dimensional and three-dimensional cases. The expressions below are for the 3D case, but can easily be adapted to 2D.

We denote by T, R and S the subspaces of the translations, the instantaneous rotations around the centroid, and the scaling motions centered on the centroid, respectively, defined on the shape Γ:

$$
\begin{aligned}
T = & \quad \left\{ v : \mathbf{x} \in \Gamma \mapsto \mathbf{t} \mid \mathbf{t} \in \mathbb{R}^3 \right\} , \\
R = & \quad \left\{ v : \mathbf{x} \mapsto (\mathbf{x} - \mathbf{G}) \wedge \omega \mid \omega \in \mathbb{R}^3 \right\} , \\
S = & \quad \left\{ v : \mathbf{x} \mapsto s(\mathbf{x} - \mathbf{G}) \mid s \in \mathbb{R} \right\} .
\end{aligned}
$$

These subspaces are mutually orthogonal for the L^2 inner product. Indeed, the L^2 product of any two fields of any two different subspaces (among T, R and S) is zero. For instance, if $v_1 : \mathbf{x} \mapsto \mathbf{t}_1$ is an element of T and $v_2 : \mathbf{x} \mapsto (\mathbf{x} - \mathbf{G}) \wedge \omega_2$ an element of R, then:

$$
\begin{aligned}
\langle v_1 \,|v_2\rangle_{L^2} &= \frac{1}{|\Gamma|} \int_\Gamma \mathbf{t}_1 \cdot \big((\mathbf{x} - \mathbf{G}) \wedge \omega_2 \big) \, d\Gamma(\mathbf{x}) \\
&= \mathbf{t}_1 \cdot \left(\left[\frac{1}{|\Gamma|} \int_\Gamma \mathbf{x} \, d\Gamma(\mathbf{x}) - \mathbf{G} \right] \wedge \omega_2 \right) \\
&= 0
\end{aligned}
$$

since $\mathbf{G}$ is the center of mass of Γ.

We suppose that these subspaces are included in the space of admissible deformations $\mathscr{D}$, and that the latter is ruled by the L^2 inner product. We denote by N the orthogonal complement of these subspaces: $F = T \perp R \perp S \perp N$. The orthogonal projection of a velocity field u on one of these subspaces can be found by minimizing $\|u - v\|_F$ with respect to v in the considered subspace. As an example, we detail the computation of $(\Pi_R \, u)$.

As for each element v of R there exists an ω such that $v(\mathbf{x}) = (\mathbf{x} - \mathbf{G}) \wedge \omega$ for all $\mathbf{x}$, we minimize the quantity $\|u - (\cdot - \mathbf{G}) \wedge \omega\|_{L^2}$ with respect to ω. We start by computing the derivative of this quantity with respect to ω, ∂_ω:

$$
\begin{aligned}
&\partial_\omega \left(\int_\Gamma \|u(\mathbf{y}) - (\mathbf{y} - \mathbf{G}) \wedge \omega\|^2 d\Gamma(\mathbf{y}) \right) = \\
&\qquad\qquad \int_\Gamma -\big(u(\mathbf{y}) - (\mathbf{y} - \mathbf{G}) \wedge \omega \big) \wedge (\mathbf{y} - \mathbf{G}) d\Gamma(\mathbf{y}) \\
&= -\int_\Gamma u(\mathbf{y}) \wedge (\mathbf{y} - \mathbf{G}) \, d\Gamma(\mathbf{y}) + \\
&\qquad\qquad \left(\int_\Gamma \|\mathbf{y} - \mathbf{G}\|^2 d\Gamma(\mathbf{y}) \right) \omega - \left[\int_\Gamma (\mathbf{y} - \mathbf{G})(\mathbf{y} - \mathbf{G})^T d\Gamma(\mathbf{y}) \right] \omega
\end{aligned}
$$

As this quantity is zero for the ω_u which minimizes $\|u - (\cdot - \mathbf{G}) \wedge \omega\|_{L^2}$, we have:

$$
\omega_u = \mathbf{M}^{-1} \left(\int_\Gamma u(\mathbf{y}) \wedge (\mathbf{y} - \mathbf{G}) \, d\Gamma(\mathbf{y}) \right)
$$

with:

$$\mathbf{M} = \left(\int_\Gamma \|\mathbf{y} - \mathbf{G}\|^2 d\Gamma(\mathbf{y}) \right) \mathrm{Id} - \int_\Gamma (\mathbf{y} - \mathbf{G})(\mathbf{y} - \mathbf{G})^T d\Gamma(\mathbf{y})$$

To guarantee that the linear application $\mathbf{M}$, we prove it is a symmetric positive definite matrix. We have indeed for any $\mathbf{x}$:

$$\mathbf{x}^T \mathbf{M} \mathbf{x} \;=\; \int_\Gamma \|\mathbf{x}\|^2 \|\mathbf{y} - \mathbf{G}\|^2 - \big(\mathbf{x} \cdot (\mathbf{y} - \mathbf{G})\big)^2 \, d\Gamma(\mathbf{y})$$

As for any $\mathbf{x}$ and $\mathbf{z}$ we have $(\mathbf{x} \cdot \mathbf{z}) \leqslant \|\mathbf{x}\| \|\mathbf{z}\|$, with equality only if the two vectors are collinear, and as $\mathbf{x}$ cannot be collinear with all $\mathbf{y} - \mathbf{G}$ for $\mathbf{y}$ in Γ, we obtain $\mathbf{x}^T M \mathbf{x} > 0$ for any $\mathbf{x}$, so M is positive definite and consequently invertible.

Note that if we had not taken for u the L^2 gradient but the gradient for another inner product F, we would have to ensure the subspaces are orthogonal for that inner product F, and compute new projections by minimizing $\|u - v\|_F$.

We apply the method we detailed for the subspace R to the other subspaces T and S, and obtain:

$$\begin{aligned}
(\Pi_T \; u)(\mathbf{x}) &= \overline{u} \;:=\; \tfrac{1}{|\Gamma|} \textstyle\int_\Gamma u(\mathbf{y}) \, d\Gamma(\mathbf{y}), \\
(\Pi_R \; u)(\mathbf{x}) &= (\mathbf{x} - \mathbf{G}) \wedge \omega_u \;, \\
(\Pi_S \; u)(\mathbf{x}) &= \tfrac{\int_\Gamma u(\mathbf{y}) \cdot (\mathbf{y} - \mathbf{G}) \, d\Gamma(\mathbf{y})}{\int_\Gamma \|\mathbf{y} - \mathbf{G}\|^2 \, d\Gamma(\mathbf{y})} (\mathbf{x} - \mathbf{G}) \;, \\
(\Pi_N \; u)(\mathbf{x}) &= u(\mathbf{x}) - (\Pi_T + \Pi_R + \Pi_S)(u)(\mathbf{x}) \;.
\end{aligned} \tag{3.8}$$

In the two-dimensional case, the expressions of the projections are the same, and the expression of ω_u can be simplified in:

$$\omega_u \;=\; \frac{\int_\Gamma (\mathbf{y} - \mathbf{G}) \wedge u(\mathbf{y}) \, d\Gamma(\mathbf{y})}{\int_\Gamma \|\mathbf{y} - \mathbf{G}\|^2 \; d\Gamma(\mathbf{y})} \,.$$

The new gradient is deduced from the L^2 gradient by equation (3.4) with

$$\mathcal{L}^{-1} = \mathrm{Id} + \left(\frac{1}{\lambda_T} - 1 \right) \Pi_T + \left(\frac{1}{\lambda_R} - 1 \right) \Pi_R + \left(\frac{1}{\lambda_S} - 1 \right) \Pi_S \,.$$

The weights λ_T, λ_R and λ_S are adapted to the user's needs in each particular application. For example:

- Boost rigid+scaling motions: $\lambda_T, \lambda_R, \lambda_S \ll 1$,
- Boost rigid motions: $\lambda_T, \lambda_R \ll 1$, $\lambda_S = 1$,
- Boost translations: $\lambda_T \ll 1$, $\lambda_R = \lambda_S = 1$.

(b) Affine motion

We can apply this same idea to the subspace $\mathcal{A}$ of instantaneous affine motions:

$$\mathcal{A} = \left\{ v : \mathbf{x} \in \Gamma \mapsto \mathbf{A}\mathbf{x} + \mathbf{b} \mid \mathbf{A} \in \mathbb{R}^{n \times n},\ \mathbf{b} \in \mathbb{R}^n \right\} .$$

The L^2 orthogonal projection on this subspace writes:

$$\left(\Pi_A\ u\right)(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} ,$$

where

$$\begin{aligned} \mathbf{A} &= \left[\int_\Gamma u(\mathbf{y})(\mathbf{y}-\mathbf{G})^T\ d\Gamma(\mathbf{y})\right]\left[\int_\Gamma (\mathbf{y}-\mathbf{G})(\mathbf{y}-\mathbf{G})^T\ d\Gamma(\mathbf{y})\right]^{-1} , \\ \mathbf{b} &= \overline{u} - \mathbf{A}\ \mathbf{G} . \end{aligned}$$

3.4.2 The Sobolev H^1 gradient flow

We consider the canonical inner product of the Sobolev space $H^1(\Gamma, \mathbb{R}^n)$ of square integrable velocity fields with square integrable derivatives, defined on the shape Γ with values in $\mathbb{R}^n$. For two such fields u and v its expression is:

$$\begin{aligned} \langle u \,|\, v \rangle_{H^1} &= \langle u \,|\, v \rangle_{L^2} + l^2 \langle D_{\mathbf{x}} u \,|\, D_{\mathbf{x}} v \rangle_{L^2} \\ &= \tfrac{1}{|\Gamma|} \textstyle\int_\Gamma u(\mathbf{x}) \cdot v(\mathbf{x}) d\Gamma(\mathbf{x}) + \tfrac{1}{|\Gamma|}\, l^2 \textstyle\int_\Gamma D_{\mathbf{x}} u(\mathbf{x}) \cdot D_{\mathbf{x}} v(\mathbf{x}) d\Gamma(\mathbf{x}) \end{aligned}$$

where $D_{\mathbf{x}}$ denotes the intrinsic derivatives on the contour and l is a characteristic length for the derivation which acts as a weight between the two integrals. The second term of this expression introduces a notion of spatial coherence: not only the length of the velocity field, but also its variations along the contour are penalized. Indeed, $D_{\mathbf{x}} u(\mathbf{x})$ stands for the matrix of the derivative of the vector field u at the point $\mathbf{x}$ on the manifold Γ and consequently expresses how much the field u varies at point $\mathbf{x}$. In the two-dimensional case, $D_{\mathbf{x}} u(\mathbf{x})$ is simply a vector. In the general case, $D_{\mathbf{x}} u(\mathbf{x}) \cdot D_{\mathbf{x}} v(\mathbf{x}) = \sum_{i,j} \left(D_{\mathbf{x}} u(\mathbf{x})\right)_{i,j} \left(D_{\mathbf{x}} v(\mathbf{x})\right)_{i,j}$ is the usual inner product between matrices.

By definition of the gradients of $E(\Gamma)$, and then thanks to an integration by parts on the manifold Γ, we have:

$$\begin{aligned}\forall\, v\,,\ \langle \nabla_{L^2} E(\Gamma) \,|\, v\rangle_{L^2} &= \mathcal{G}_\Gamma(E(.), v) \;=\; \langle \nabla_{H^1} E(\Gamma) \,|\, v\rangle_{H^1} \\ &= \langle \nabla_{H^1} E \,|\, v\rangle_{L^2} + l^2\, \langle D_{\mathbf{x}} \nabla_{H^1} E \mid\, D_{\mathbf{x}} v \,\rangle_{L^2} \\ &= \left\langle \nabla_{H^1} E - l^2 \Delta \nabla_{H^1} E \,\middle|\, v \right\rangle_{L^2}\end{aligned}$$

where Δ denotes the intrinsic Laplacian operator on the contour, often called the Laplace-Beltrami operator.

Thus the H^1 inner product is related to the L^2 inner product as proposed in subsection 3.3.1 through the linear operator $\mathcal{L}(u) = u - l^2 \Delta u$, As a consequence, the H^1 gradient can be obtained from the L^2 gradient by solving an intrinsic heat equation with a data attachment term:

$$l^2\, \Delta u = u - \nabla_{L^2} E\ . \tag{3.9}$$

Interestingly, the solution of equation (3.9) coincides with

$$\arg\min_u \int_\Gamma \|u(\mathbf{x}) \;-\; \nabla_{L^2} E(\Gamma)(\mathbf{x})\|^2 \; d\Gamma(\mathbf{x}) \;+\; l^2 \int_\Gamma \|D_{\mathbf{x}} u(\mathbf{x})\|^2 \; d\Gamma(\mathbf{x}) \tag{3.10}$$

Intuitively, the H^1 gradient is a smoothed version of the L^2 gradient and can be obtained by a process similar to the image restoration process on a manifold Γ, a problem familiar to the image processing community. The factor l^2 acts as a parameter balancing the influences of the data term and the regularizing term. Actually, smoothing a gradient using this particular inner product is a standard "trick", well-known in numerical analysis. As we mentioned previously, this idea has been introduced in computer vision simultaneously by us [24] and by Sundaramoorthi et al. [130]. However, the main point remains that, introducing this smoothing via a modification of the gradient rather than directly from equation (3.10), warrants that the gradient descent will decrease the energy.

In the two-dimensional case, the shape is a curve which can be parametrized by its arc length σ, so that any field u defined on Γ can be seen as an application from $[\,0, |\Gamma|\,]$ to $\mathbb{R}^2$, where $|\Gamma|$ is the length of the curve. The explicit solution of the equation $\Delta u = u - v$ is then known and given by:

$$u(\sigma) \;=\; \frac{1}{2l}\left(e^{\sigma/l}\left(A - \int_0^\sigma e^{-\tau/l}\, v(\tau)\, d\tau\right) \;+\; e^{-\sigma/l}\left(B + \int_0^\sigma e^{\tau/l}\, v(\tau)\, d\tau\right)\right) \tag{3.11}$$

$$\text{with } A = \frac{e^{|\Gamma|/l}}{e^{|\Gamma|/l}-1} \oint_\Gamma e^{-\tau/l}\, v(\tau)\, d\tau$$
$$\text{and } B = \frac{1}{e^{|\Gamma|/l}-1} \oint_\Gamma e^{\tau/l}\, v(\tau)\, d\tau\,.$$

Of course, the choice of the initial point on Γ in order to define its parametrization by the arc length does not interfere with the resulting solution considered as an application from Γ into $\mathbb{R}^2$.

In greater dimensions, we can obtain in practice the H^1 gradient, solution of equation (3.9), from an iterative minimization induced by (3.10). We will see in chapter 4 that since the work introduced in [9], implementing a PDE on a surface is affordable in the implicit framework with the level set method [39, 108].

3.4.3 Intrinsic Gaussian smoothing

We apply the procedure of subsection 3.3.2 to design a useful minimizing flow: it is a smoothed version of the L^2 gradient flow. Hence, to some extent, it resembles the H^1 gradient flow of subsection 3.4.2. However, here, we apply an *ad hoc* procedure to the L^2 gradient without resorting to an inner product.

We define a linear intrinsic smoothing operator which may be seen as the counterpart on the contour of Gaussian smoothing in $\mathbb{R}^{n-1}$, by considering the solution $\tilde{u}$ of the intrinsic heat equation on Γ with initial condition u:

$$\begin{cases} \tilde{u}(.,0) = u \\ \dfrac{\partial \tilde{u}}{\partial \tau} = \Delta\, \tilde{u} \end{cases}, \tag{3.12}$$

where Δ denotes the Laplace-Beltrami operator. We then denote by $\mathcal{L}_\tau\, u$ its solution $\tilde{u}(.,\tau)$ at time $\tau \geq 0$.

On the one hand, $\mathcal{L}_\tau$ is symmetric positive. In particular, a flow (3.5) based on this operator decreases the energy. The larger is τ, the smoother is the flow.

$\mathcal{L}_\tau$ is symmetric:

$$\begin{array}{rcccl} \langle \mathcal{L}_0(u)\,|v\rangle_{L^2} & = & \langle u\,|\mathcal{L}_0(v)\rangle_{L^2} & = & \langle u\,|v\rangle_{L^2}\,, \\ \dfrac{\partial}{\partial\tau}\langle \mathcal{L}_\tau(u)\,|v\rangle_{L^2} & = & \dfrac{\partial}{\partial\tau}\langle u\,|\mathcal{L}_\tau(v)\rangle_{L^2} & = & -\langle D_{\mathbf{x}}u\,|D_{\mathbf{x}}v\rangle_{L^2} \end{array}$$

$\mathcal{L}_\tau$ is positive:

$$\langle \mathcal{L}_\tau(u) \,|u\rangle_{L^2} = \qquad \langle \mathcal{L}_{\tau/2}\mathcal{L}_{\tau/2}(u) \,|u\rangle_{L^2} = \left\|\mathcal{L}_{\tau/2}(u)\right\|^2_{L^2} \geq 0$$

But on the other hand, the inversion of $\mathcal{L}_\tau$ for $\tau > 0$ is an ill-posed anti-diffusive process. So a gradient interpretation is not available.

3.5 Numerical Experiments With The New Inner Products

The approach presented in this chapter can be applied to virtually any active contour evolution. In this section, we have chosen some particular applications to demonstrate the interest of our contribution.

Moreover, the content of this chapter is not specific to a particular implementation of the contour evolution. In our experiments, we have used the level set framework [39, 108, 124, 106, 107], motivated by its numerical stability and its ability to handle topological changes automatically. The implicit framework also offers an elegant formulation of the Laplace-Beltrami operator [9] and of the average of a quantity along the contour [115]. We will see in chapter 4 an elegant example of what can be done using our new gradients in the level set framework.

The additional computational cost of our approach depends on the type of minimizing flow we consider. The extra time is barely noticeable for the rigid plus scaling and affine flows of paragraphs 3.4.1(a) and 3.4.1(b). The latter only require to compute a handful of integrals on the contour. The smooth minimizing flows of subsections 3.4.2 and 3.4.3 are more demanding. In 2D, the implicit diffusion equations (3.9) and (3.12) are equivalent to some convolutions with respect to the curvilinear coordinate on Γ. In 3D (and eventually more), they must be solved with some iterative methods, for each time step.

3.5.1 Shape warping

We illustrate our approach in the problem of shape warping. In this context, the energy functional to be minimized is a measure of dissimilarity between the evolving contour and a target contour. The study of shape metrics is still an active research area [147, 145, 22], and there are many candidates for the dissimilarity measure. In this chapter, we use the differentiable approximation of the well-known Hausdorff distance, as proposed in [22] and presented in chapter 1, to warp the contours of two different hands.

Let us first illustrate an important aspect of our method: Modifying the gradient by changing the inner product involved in its definition and reprojecting the velocity field onto specifics subspaces of $\mathscr{D}$ while favoring particular deformation field (such as spatially coherent one) *does not change the intrinsic informations contained in the field*. Fig.3.1 and Fig.3.2 illustrate this notion. The first one (3.1) shows on top row a "classic" L^2 warping and on bottom row a L^2 warping "rigidified" by our spatially coherent flows (favoring rigid plus scaling motions), while the second one (3.2) shows on top row a "classic" Hausdorff warping and on bottom row a Hausdorff warping with the same "rigidification". In both examples our "rigidification" forced the evolving square to keep being a square during the evolution. However the two evolutions are not the same at all. In the L^2 warping, the initial field contains at first principally a scaling component, although in the Hausdorff warping the initial field contain directly a translating component.

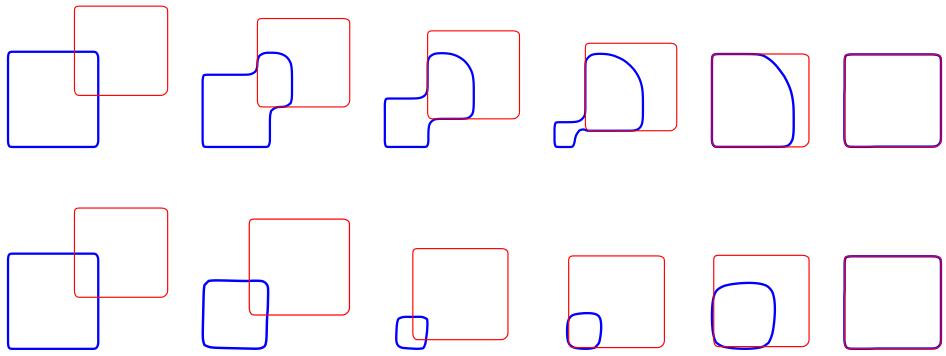


Figure 3.1: L^2 warping (*top row*) and L^2 warping with a modified gradient descent favoring rigid plus scaling motions (*bottom row*)

Fig. 3.3 (by G. Charpiat) compares the evolution of the contour when using the L^2 gradient descent *(top row)* and a modified gradient descent favoring rigid plus scaling motions *(bottom row)* as in paragraph 3.4.1(a).

Both evolutions achieve a perfect warping. However, despite the similarity of the two input shapes, the L^2 gradient flow goes through some states of completely different appearances. The trajectory followed by this flow looks particularly inefficient and unnatural, because the notion of length contained in the L^2 inner product is very far from our intuition. In contrast, the behavior of our gradient flow is natural and visually pleasing.

In Fig .3.4, we show a three-dimensional warping example from a teddy bear to Hayao Miyazaki's character Totoro. We use here the $W^{1,2}$-norm

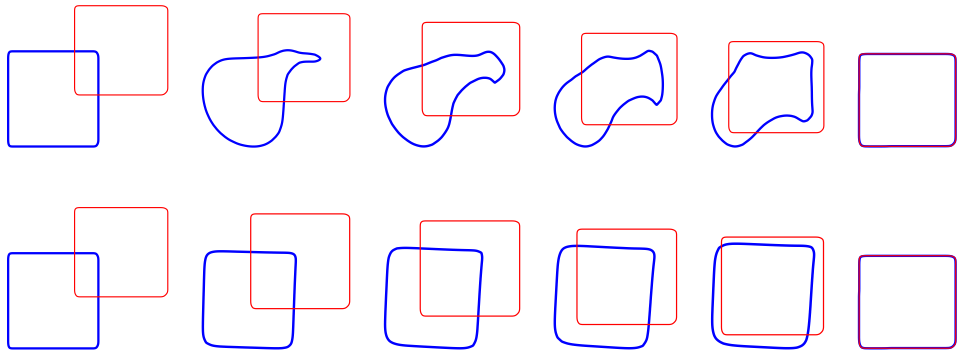


Figure 3.2: Hausdorff warping (*top row*) and Hausdorff warping with a modified gradient descent favoring rigid plus scaling motions (*bottom row*)

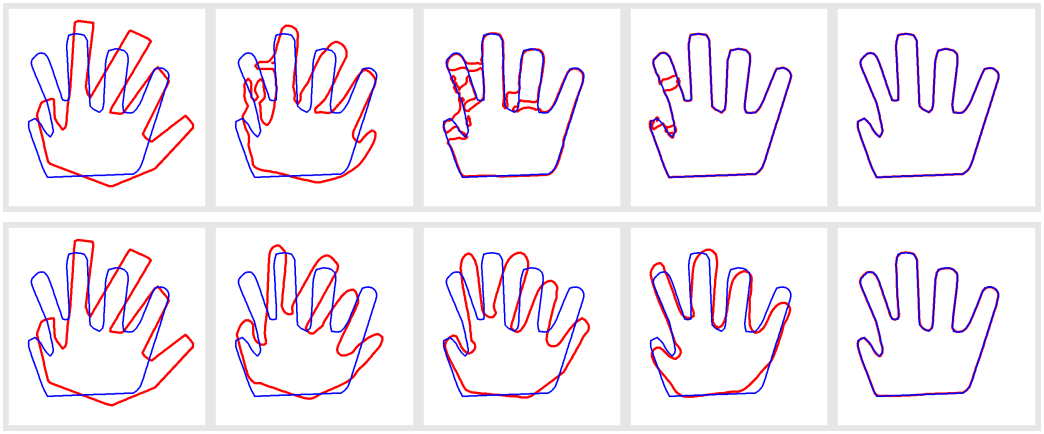


Figure 3.3: Shape warping with the L^2 gradient descent *(top)* and with a modified gradient descent favoring rigid plus scaling motions *(bottom)*: $\lambda_T = \lambda_R = \lambda_S = 0.025$.

of the distance functions as proposed in chapter 1. Despite an initial rigid registration, the L^2 gradient descent is unable to give satisfying results. A modified gradient descent favoring rigid plus scaling motions leads to better results.

This suggests that our approach can infer relevant correspondences between the two contours, as a byproduct of the warping process. This point-to-point matching is obtained by tracking the points along the evolution. It does not make much sense with a L^2 gradient flow, because the latter yields a strictly normal velocity field as showed in chapter 2, section 2.4. But

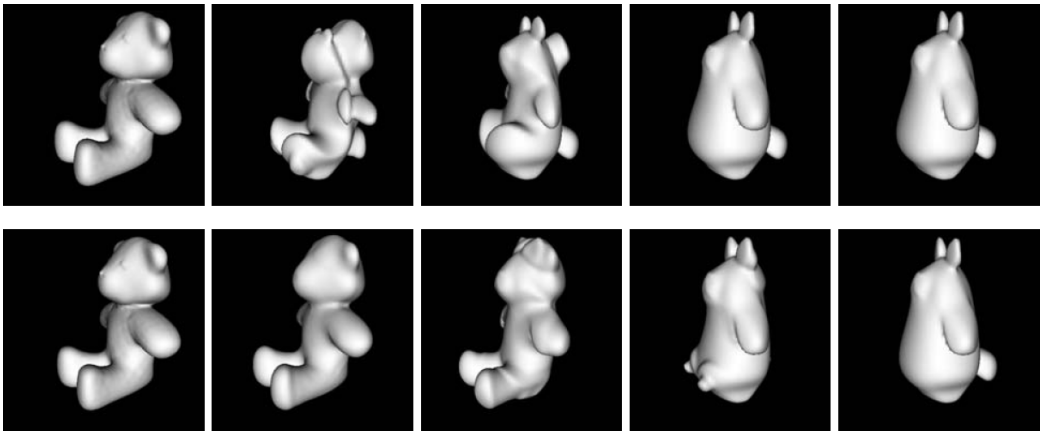


Figure 3.4: 3D shape warping with the L^2 gradient descent *(top)* and with a modified gradient descent favoring rigid plus scaling motions *(bottom)*: $\lambda_T = \lambda_R = \lambda_S = 0.025$.

when using our approach, the velocity field has a meaningful tangential part. Maintaining point correspondences during the evolution is straightforward in an implementation with meshes. It is also feasible in a level set implementation, with an extension proposed in [117]. In chapter 4 we will describe more circumstantially this technique and present a framework using it.

3.5.2 Tracking

We now illustrate the better robustness to local minima of spatially coherent minimizing flows with a naive experiment. We insist on the fact that this example is illustrative: we did not look for the method and the energy that gave the best results of tracking for the particular sequence we worked on; we focus more on the improvements brought by our change of inner product rather than on the results themselves.

We track a moving hand in a monocular video sequence. For each frame, we minimize the contour-based energy of the original geodesic active contours method [20], starting from the result of the segmentation of the previous frame. Note that a region-based approach [113] or a background subtraction method would give better results on our particular test sequence.

Fig. 3.5 (by G.Charpiat) compares the evolution of the contour when using the L^2 gradient descent *(top row)* and a modified gradient descent favoring affine motions *(bottom row)* as in paragraph 3.4.1(b). Due to large displacements between consecutive frames, the L^2 gradient flow fails and the

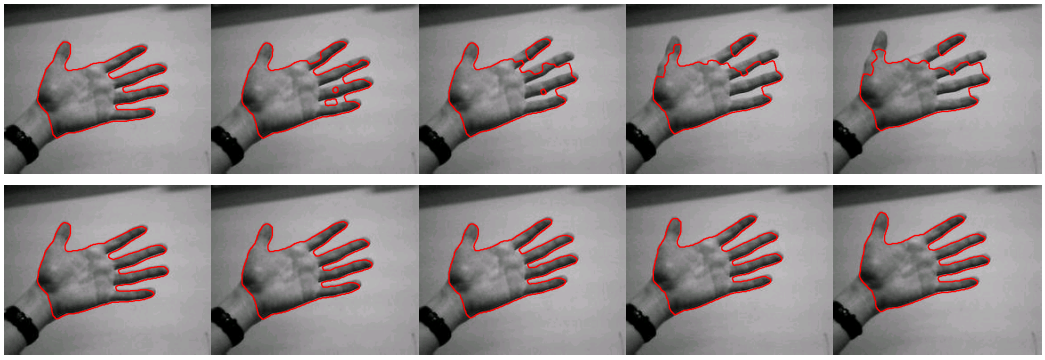


Figure 3.5: Tracking a hand in a video sequence with the L^2 gradient descent *(top)* and with a modified gradient descent favoring affine motions *(bottom)*: $\lambda_A = 0.025$.

contour finally locks between two fingers, whereas our gradient flow manages to dodge this unwanted low-energy configuration.

3.6 A generalized gradient method

In this section, we go further and consider the definition of the gradient of an energy from a new point of view, which leads us to a larger class of minimization algorithms. The thread we follow is the fact that the gradient of the energy can be obtained as the result of another minimization problem.

3.6.1 The gradient seen as the result of a minimization problem

To help developing the reader's intuition let us recall that the usual gradient descent method can be seen, up to first order, as minimizing $E(\Gamma + u)$ with respect to the deformation field u through the linearization of the energy E in the neighborhood of the shape Γ:

$$E(\Gamma + u) \simeq E(\Gamma) + \mathcal{G}_\Gamma(E(.), u)$$

But since $\mathcal{G}_\Gamma(E(.), u)$ is linear with respect to the deformation field u, there is no minimum. This is of course a direct consequence of the first-order approximation. It is therefore more sensible to speak in terms of the direction of the deformation field u. The notion of direction implies the choice of a norm: the set of all directions is the set of all fields with norm equal to 1.

Once a norm F has been chosen (related to an inner product preferably), a natural solution appears as the direction u_F that minimizes the energy $\mathcal{G}_\Gamma(E(.), v)$:

$$u_F = \underset{\{v \in \mathscr{D} \text{ s.t. } \|v\|_F = 1\}}{\arg\min} [\ \mathcal{G}_\Gamma(E, v)\ ] \ = \ -\frac{\nabla_F E(\Gamma)}{\|\nabla_F E(\Gamma)\|_F} \tag{3.13}$$

The main point here is that the opposite of the gradient $-\nabla_F E(\Gamma)$ of the energy E for the inner product related to the norm F is precisely in the direction u_F. This gradient has been introduced previously as the deformation field linked to the continuous linear form $\mathcal{G}_\Gamma(E, \cdot)$ for the inner product F thanks to the Riesz theorem. Note that the influence of the inner product F upon the best direction u_F lies in the fact that it changes the shape of the unit sphere (the set of all directions u with unit norm $\|u\|_F = 1$).

It turns out that the gradient itself (not only its direction) can be obtained as the solution of a minimization problem. This also explicits the link between the norm F and the gradient. This is shown in the following

Theorem 14. *The gradient $\nabla_F E(\Gamma)$ for the inner product F satisfies:*

$$-\nabla_F E(\Gamma) = \underset{v \in \mathscr{D}}{\arg\min} \left[\ \mathcal{G}_\Gamma(E, v) \ + \ \frac{1}{2}\|v\|_F^2 \ \right]$$

Proof. We have indeed, for any v:

$$\begin{aligned} \mathcal{G}_\Gamma(E, v) \ + \ \frac{1}{2}\|v\|_F^2 \ &= \ \frac{1}{2}\left[\|v\|_F^2 + 2\ \langle v\, |\nabla_F E(\Gamma)\rangle_F\ \right] \\ &= \ \frac{1}{2}\left[\|v + \nabla_F E(\Gamma)\|_F^2 - \|\nabla_F E(\Gamma)\|_F^2\right] \end{aligned}$$

So that:

$$\begin{aligned} \underset{v \in \mathscr{D}}{\arg\min} \left[\ \mathcal{G}_\Gamma(E, v) \ + \ \frac{1}{2}\|v\|_F^2 \ \right] \ &= \ \underset{v}{\arg\min} \left[\|v + \nabla_F E(\Gamma)\|_F^2\right] \\ &= \ -\nabla_F E(\Gamma) \end{aligned}$$

□

The expression between brackets in Theorem 14 breaks up into two parts: the first one, $\mathcal{G}_\Gamma(E, v)$, comes from the energy $E(\Gamma)$ and stands for the quantity to minimize, whereas the second one, $R(v) = \frac{1}{2}\|v\|_F^2$, is a regularizing term which imposes to the solution to be smooth and small enough in the sense of the norm F. Different choices of the smoothing term *thanks to different choices of the norm F* imply different final gradients $\nabla_F E(\Gamma)$.

For example, the choice of the H^1 inner product leads to the regularizing term:

$$R(v) = \frac{1}{2}\|v\|_{L^2}^2 + \frac{1}{2}l^2\|Dv\|_{L^2}^2$$

and consequently the gradient $\nabla_{H^1}E(\Gamma)$ is the deformation field which minimizes:

$$\mathcal{G}_\Gamma(E, v) \; + \; \frac{1}{2}\|v\|_{L^2}^2 \; + \; \frac{1}{2}l^2\|Dv\|_{L^2}^2$$

This leads us to an elegant proof of a property of the H^1 gradient stated in section 3.4.2, without considering PDEs:

Proposition 15. *The opposite of the H^1 gradient is the solution of:*

$$\arg\min_{v\in\mathscr{D}} \left[\|u - v\|_{L^2}^2 + l^2\|Dv\|_{L^2}^2\right]$$

where $u = -\nabla_{L^2}E(\Gamma)$ is the opposite of the usual L^2-gradient.

Proof. Indeed, for any v:

$$\|u - v\|_{L^2}^2 = \|u\|_{L^2}^2 - 2\left\langle u \,|\, v\right\rangle_{L^2} + \|v\|_{L^2}^2$$

Therefore,

$$\|u - v\|_{L^2}^2 + l^2\|Dv\|_{L^2}^2 = \|u\|_{L^2}^2 + 2\,\mathcal{G}_\Gamma(E, v) + \|v\|_{H^1}^2$$

since by definition of u, $\mathcal{G}_\Gamma(E, v) = \left\langle -u \,|\, v\right\rangle_{L^2}$. And thanks to theorem 14:

$$\arg\min_v \left[\|u - v\|_{L^2}^2 + l^2\|Dv\|_{L^2}^2\right] = \arg\min_v \left[\mathcal{G}_\Gamma(E, v) + \frac{1}{2}\|v\|_{H^1}^2\right] = -\nabla_{H^1}E(\Gamma)$$

□

3.6.2 Generalization of the regularizing term

We have stressed the influence of the choice of an inner product $\left\langle |\right\rangle_F$ on the associated gradient:

$$-\nabla_F E(\Gamma) = \arg\min_v \left[\; \mathcal{G}_\Gamma(E, v) \; + \; R(v) \;\right]$$

where $R(v) = \frac{1}{2}\|v\|_F^2$, and $\|\cdot\|_F$ is the norm related to the chosen inner product. Since the choice of the inner product is equivalent to the choice of

the regularizing term $R(v)$ and acts qualitatively upon the gradient descent paths, we can see $R(v)$ as a prior on the deformation fields.

Let us now generalize our framework and allow $R(v)$ to be (almost) any positive real function, not necessarily related to an inner product, and compute (when it exists) the associated field which we will denote, with a slight abuse of notation, by $-\nabla_R E(\Gamma)$ (note that if R is related as previously to the inner product F, then $\nabla_F E = \nabla_R E$):

$$-\nabla_R E(\Gamma) = \arg\min_{v \in \mathscr{D}} [\, \mathcal{G}_\Gamma(E, v) \; + \; R(v) \,] \tag{3.14}$$

Under some reasonable assumptions about $R(v)$, the new "gradient" $\nabla_R E(\Gamma)$ exists and has interesting properties.

The question of the existence and unicity of ∇_R in general is not the main point here, it depends on the particular chosen application R. Here, R is supposed to stand for an application approximatively "looking like" the square of a norm; for reasonable choices of R from this point of view, the existence is guaranteed, and so is the uniqueness in most cases.

We now present the fundamental property of the extended gradient $\nabla_R E$ as the

Theorem 16. *If R is differentiable and reaches its global minimum at the zero field, then the flow $-\nabla_R E(\Gamma)$, if it exists, decreases the energy E.*

Proof. We start by proving that $\mathcal{G}_\Gamma(E, -\nabla_R E(\Gamma)) \leqslant 0$.

We have $-\nabla_R E(\Gamma) = \arg\min_v [\, \mathcal{G}_\Gamma(E, v) \; + \; R(v) \,]$. It can be written as:

$$\forall v \in \mathscr{D} : \mathcal{G}_\Gamma(E, -\nabla_R E(\Gamma)) \; + \; R(-\nabla_R E(\Gamma)) \; \leqslant \; \mathcal{G}_\Gamma(E, v) \; + \; R(v)$$

Therefore, in particular, considering the zero field $v = 0$:

$$\mathcal{G}_\Gamma(E, -\nabla_R E(\Gamma)) \; + \; R(-\nabla_R E(\Gamma)) \; \leqslant \; \mathcal{G}_\Gamma(E, 0) \; + \; R(0)$$

$$\mathcal{G}_\Gamma(E, -\nabla_R E(\Gamma)) \leqslant R(0) - R(-\nabla_R E(\Gamma))$$

since $\mathcal{G}_\Gamma(E, 0) = 0$ by definition of the Gâteaux derivatives (Def. 7). As, by hypothesis, $v = 0$ is the global minimum of R, we have $R(-\nabla_R E(\Gamma)) \geqslant R(0)$, so:

$$\mathcal{G}_\Gamma(E, -\nabla_R E(\Gamma)) \leqslant 0$$

Moreover, this last inequality is strict if the usual gradient $\nabla_{L^2} E(\Gamma)$ is not zero. Indeed, in that case, as $R(v)$ reaches its global minimum at the zero field $v = 0$, its derivative is zero for $v = 0$. Consequently, the L^2 gradient of $\mathcal{G}_\Gamma(E, v) + R(v)$ with respect to v at the zero field equals $\nabla_{L^2} E(\Gamma)$, which is not zero by hypothesis. Therefore $\inf_v [\, \mathcal{G}_\Gamma(E, v) \; + \; R(v) \,]$ is not reached at $v = 0$ and all inequalities in the proof are strict. □

Note that the application R is specific to the shape Γ (or, more exactly, to its tangent space) and there is no assumption about a "regularity" of the applications R_Γ with respect to Γ. However, as in the previous part where we had to associate to each shape an inner product and naturally chose the same general expression for all of them, we will restrict ourselves to the case where the application R_Γ has the same general expression $R(\Gamma)$ for all shapes Γ and consequently will commit a slight abuse of notation between R and R_Γ.

3.6.3 Remarks

(a) Addition of an orthogonal term

Note that the method proposed in section 3.3.3, which consists in adding an orthogonal term to the gradient (see equation (3.6)), can be seen as a variation on the extended gradient theme, where the search for the infimum has been restricted to the affine hyperplane $\mathcal{H}$ containing the opposite of the gradient $-\nabla_F E$ and orthogonal to it. Indeed:

$$\begin{aligned}\arg\min_{v\in\mathcal{H}} \left[\mathcal{G}_\Gamma(E,v) + R(v)\right] &= \arg\min_{w;w\perp\nabla_F E} \left[\mathcal{G}_\Gamma(E,-\nabla_F E + w) + R(-\nabla_F E + w)\right] \\ &= \arg\min_{w;w\perp\nabla_F E} R(-\nabla_F E + w)\end{aligned}$$

(b) Directional formulation

We have seen earlier (equation (3.13)) that the direction of the gradient could be defined as the field v of the unit sphere $\mathcal{U}_F = \{v \text{ s.t. } \|v\|_F = 1\}$ which most decreases the energy, and that changing the inner product F was precisely acting on the gradient by changing the unit sphere. One way to generalize the notion of gradient could have been to set any hypersurface $\mathcal{S}$ instead of the unit sphere $\mathcal{U}_F$ and to search for the best field v in $\mathcal{S}$. However, this would lead to some difficulties in practice (how to search for a minimum on an hypersurface of infinite dimension, how to represent this hypersurface?). A slightly better way to do this would be to focus on the hypersurfaces of the form $\mathcal{U}_R = \{v \text{ s.t. } R(v) = 1\}$, which is in the spirit of the level-set method. Note that this approach would be very close in practice to the one we described, the main difference being that we would only obtain a direction, without the magnitude.

(c) Temporal coherence

The application $R(v)$ does not necessarily only deal with spatial coherence and can also be designed to favor temporally coherent motions. For example,

at time step t of an evolution, one could force the new deformation field u_t to resemble the previous one u_{t-1}. If we transport u_{t-1} from the previous shape Γ_{t-1} to the new one Γ_t, we obtain a new field noted $T(u_{t-1})$ defined on the same space as u_t, and we can compare them, e.g., with $\|T(u_{t-1}) - u_t\|$. We are thus led to define $R(v) = \|T(u_{t-1}) - v\|$. This function however does not satisfy in general the condition $R(0) = 0$ which is necessary in theorem 16. Nevertheless this problem can be solved by defining $R(v)$ to be the norm of the projection of v orthogonally to $T(u_{t-1})$.

3.6.4 Computing the extended gradient

If R is simple enough so that the inverse application of $v \mapsto D_v R(v)$ is easily computable, then the computation of the extended gradient $\nabla_R E$ is immediate from the knowledge of the usual L^2-gradient. Indeed, the application $v \mapsto \mathcal{G}_\Gamma(E, v) + R(v)$ has a local minimum at $v = -\nabla_R E(\Gamma)$, so its derivative with respect to v is zero at this point:

$$
\begin{aligned}
D_v \left(\mathcal{G}_\Gamma(E, v) + R(v) \right)\big|_{v=-\nabla_R E} &= 0 \\
D_v \left(\langle \nabla_{L^2} E(\Gamma) \,|\, v \rangle_{L^2} + R(v) \right)\big|_{v=-\nabla_R E} &= 0 \\
\nabla_{L^2} E(\Gamma) + D_v R\big(-\nabla_R E(\Gamma) \big) &= 0 \\
\nabla_R E(\Gamma) = -\left(D_v R\right)^{-1} \big(-\nabla_{L^2} E(\Gamma) \big)&
\end{aligned}
$$

This formula generalizes the one obtained previously in proposition 13 in section 3.3 concerning the relation between the gradient for an inner product P and the usual L^2 gradient. Now, for the extended gradient, the application $(D_v R)^{-1}$ which stands in for $\mathcal{L}$ in this proposition is not supposed to be linear anymore.

In more general cases, if we cannot compute the application $(D_v R)^{-1}$, we can still solve the infimum problem with a Partial Differential Equation (PDE) which is equivalent to a . . . infinitesimal gradient descent! The definition in equation (3.14) can be seen indeed as a minimization problem which leads to the evolution:

$$
\begin{cases}
v(0) = 0 \\
\dfrac{dv}{dt} = -\nabla_{L^2} E(\Gamma) - D_v R(v)
\end{cases}
\tag{3.15}
$$

This evolution leads to a local minimum of $\mathcal{G}_\Gamma(E, v) + R(v)$. Even if this local minimum is not the global one (if R has not been well-chosen) or if the evolution is stopped before the convergence, the final flow v computed will strictly decrease the energy $E(\Gamma)$ (same proof as in theorem 16). This

point may be important in practice. Note also that there exist many other methods [14] than the usual gradient descent to solve that kind of problem, since the quantity to minimize is a sum of a linear term $\mathcal{G}_\Gamma(E, v)$ and another term $R(v)$ which "looks like" a quadratic term since it is supposed to play a role similar to the square of a norm.

3.6.5 Application: the semi-local rigidification

We now present an example for which the previous framework appears to be useful. We consider an energy $E(\Gamma)$ defined on plane curves. These curves are assumed to lie in an image Ω, in fact a bounded subset of $\mathbb{R}^2$. Instead of following a usual gradient descent in order to minimize $E(\Gamma)$ with respect to Γ, we would like to favor the deformation fields which preserve the rigidity of the shape as much as possible, or, more exactly, we would like to favor more rigid deformations, so that some kinds of local minima could be avoided. In section 3.4.1(a) we showed how to change the inner product so as to favor global rigid transformations. In case of articulated movement, this global method may not be sufficient, so we would like to favor fields containing parts close to rigid motions; this leads us to the notion of "semi-local rigidification". We use the expression "semi-local" in order to emphasize the contrast with usual smoothing methods such as Gaussian smoothing or H^1 smoothing, which we will qualify of "local".

Let us consider a shape Γ and any field v defined on it. We would like to find the parts, if any, of the field v which are well approximated by a translation or a rotation acting on the corresponding parts of Γ. In order to model this, we associate to each point $\mathbf{x}$ of Γ a rigid deformation $w_\mathbf{x}$ defined on the whole image Ω. In order to describe $w_\mathbf{x}$ we introduce three functions defined on Γ:

- a translation $T(\mathbf{x})$
- a center of rotation $C(\mathbf{x})$
- the magnitude $A(\mathbf{x})$ of the instantaneous rotation

so that:

$$\forall \mathbf{y} \in \Omega, w_\mathbf{x}(\mathbf{y}) = A(\mathbf{x}) \left(\mathbf{y} - C(\mathbf{x})\right)^\perp + T(\mathbf{x})$$

where $a^\perp$ stands for the vector a rotated by $+\pi/2$. We suppose that this rigid deformation $w_\mathbf{x}$ varies slowly with respect to $\mathbf{x}$, that is to say we suppose the $L^2(\Omega, \mathbb{R}^2)$ norm of its derivative $\|D_\mathbf{x} w_\mathbf{x}(\cdot)\|_{L^2}$ to be small for each point $\mathbf{x}$ of the curve Γ. We consider the $L^2(\Gamma, \mathbb{R})$ norm of this application

defined on Γ and obtain the criterion $\left\| \|D_{\mathbf{x}} w_{\mathbf{x}}(\cdot)\|_{L^2(\Omega,\mathbb{R}^2)} \right\|_{L^2(\Gamma,\mathbb{R})}$ to quantify the smoothness of the field $w_{\mathbf{x}}$ of rigid deformations on Γ.

It is always possible to express any field v on Γ as a member of the class rigid motions:

$$\forall \mathbf{x} \in \Gamma, v(\mathbf{x}) = w_{\mathbf{x}}(\mathbf{x}) = A(\mathbf{x})(\mathbf{x} - C(\mathbf{x}))^{\perp} + T(\mathbf{x}) \tag{3.16}$$

The field v is then completely defined by the knowledge of T, A and C. For a given field v, there exist of course many triplets (T, A, C) satisfying (3.16), the simplest one being $(v, 0, G)$, where G_Ω is the center of mass of the image Ω. In order to lift this ambiguity we define a deformation prior R which depends on T, A and C that should be seen as parameters of v:

$$R(T, A, C) = \|v\|_{L^2}^2 + \left\| \|D_{\mathbf{x}} w_{\mathbf{x}}(\cdot)\|_{L^2(\Omega,\mathbb{R}^2)} \right\|_{L^2}^2$$

which in fact can also be written simpler (by expanding and integrating the expression $\|D_{\mathbf{x}} w_{\mathbf{x}}(\mathbf{y})\|^2$) as:

$$R(T, A, C) = \|v\|_{L^2}^2 + \left\| DT + DA\,(G_\Omega - C)^{\perp} - A\,DC^{\perp} \right\|_{L^2}^2 + \sigma_\Omega^2 \|DA\|_{L^2}$$

where:

$$\sigma_\Omega^2 = \int_\Omega (\mathbf{y} - G_\Omega)^2 d\mathbf{y}$$

is a characteristic squared "length" of the image. The middle term represents the interaction between T, A and C; for example, changing the center of rotation $DC(\mathbf{x})$ has no consequence on the rigid motion $w_{\mathbf{x}}$ if it is compensated by the adequate added translation $DT = A\,DC^{\perp}$. Note that the quantities G_Ω and σ_Ω are the only ones where the influence of the image Ω appears.

In order to compute the generalized gradient ∇_R of an energy $E(\Gamma)$, we first compute the usual L^2 gradient $\nabla_{L^2} E$, initialize $(T, A, C) = (0, 0, G_\Omega)$ so that the corresponding field $v(T, A, C)$ is zero, as required in equation (3.15), and let (T, A, C) evolve to minimize $\mathcal{G}_\Gamma(E(.), v) + R(T, A, C)$. The corresponding PDEs are

$$\left\{ \begin{array}{rcl} \partial_t T(\mathbf{x}) & = & -\big(\nabla_{L^2} E(\Gamma)(\mathbf{x}) + v(\mathbf{x})\big) \;+\; \Delta \overline{w_{\mathbf{x}}} \\ \partial_t A(\mathbf{x}) & = & -\big(\nabla_{L^2} E(\Gamma)(\mathbf{x}) + v(\mathbf{x})\big) \cdot \big(\mathbf{x} - C(\mathbf{x})\big)^{\perp} \\ & & \qquad + \;\big(G_\Omega - C(\mathbf{x})\big)^{\perp} \cdot \Delta \overline{w_{\mathbf{x}}} \;+\; \sigma_\Omega^2 \Delta A(\mathbf{x}) \\ \partial_t C(\mathbf{x}) & = & -A(\mathbf{x})\big(\nabla_{L^2} E(\Gamma)(\mathbf{x}) + v(\mathbf{x})\big)^{\perp} \;+\; A(\mathbf{x}) \Delta \overline{w_{\mathbf{x}}}^{\perp} \end{array} \right.$$

where $\overline{w_{\mathbf{x}}} = w_{\mathbf{x}}(G_\Omega)$ is the mean of the linear application $\mathbf{y} \mapsto w_{\mathbf{x}}(\mathbf{y})$ on Ω.

Let us notice that if we had considered only translations T (and not rotations), we would have $\overline{w_{\mathbf{x}}} = T(\mathbf{x}) = v(\mathbf{x})$ and the algorithms would act as an H^1 smoothing.

3.6.6 Numerical Example

We now apply this method to a specific choice of the energy $E(\cdot)$ to minimize. We would like to warp a given initial shape Γ_1 onto a given target shape Γ_2, that is to say, we would like to minimize the shape distance between Γ_1 and Γ_2 with respect to Γ_1. We choose for $E(\cdot)$ the smooth approximation of the Hausdorff distance described in chapter 1, which we will denote here by $d_H(\Gamma_1, \Gamma_2)$.

This energy E achieves generally good warping between any two shapes which are relatively close one to the other, but, in case of large deformations, it can suffer from an important lack of spatial coherence if a part A of the moving shape Γ_1 has to cross a part B of the target one on its way to another parallel part C of the target shape (see Fig; 3.6 for an example), because the part A tries to minimize its distance to both parts B and C at the same time. The major problem of this smooth approximation is hat its usual gradient sometimes lacks spatial coherence.

A global coherence can nonetheless be recovered by an adequate change of inner product which favors rigid transformations, as presented before. However, this is not sufficient for dealing with local deformations. The methods of Gaussian or H^1 smoothing studied in sections 3.4.3 and 3.4.2 could be helpful, since their action is local. But even if their influence is appreciable, these smoothing techniques do not favor semi-locally rigid deformations like the movements of an articulated object. We have noticed that, in practice, the quality of the matching between two shapes Γ_1 and Γ_2 generally depends on the quality of the path that has been followed during the evolution from one shape to the other, or, more precisely, on how natural a human observer would judge this path. This statement is very intuitive and qualitative but we believe that this quality precisely relies on notions such as articulated motion. There is clearly room here for further work. In any case this is the reason why we think that methods like the ones proposed in this chapter, which allow to set priors on the deformation fields can have interesting practical applications.

We use the framework presented above and compare the evolutions resulting from three different approaches on a difficult example of shape warping in the case of the Hausdorff distance: the usual L^2 gradient method, the H^1 smoothing method of section 3.4.2 (for the best value of the smoothness parameter l in equation (3.10)) and the semi-local rigidification method (Fig. 3.6). The last one achieves the best path and the best correspondences (Fig. 3.7).

The gradient descent framework in the case of an extended gradient ∇_R could have needed some important additional time if we had to wait until

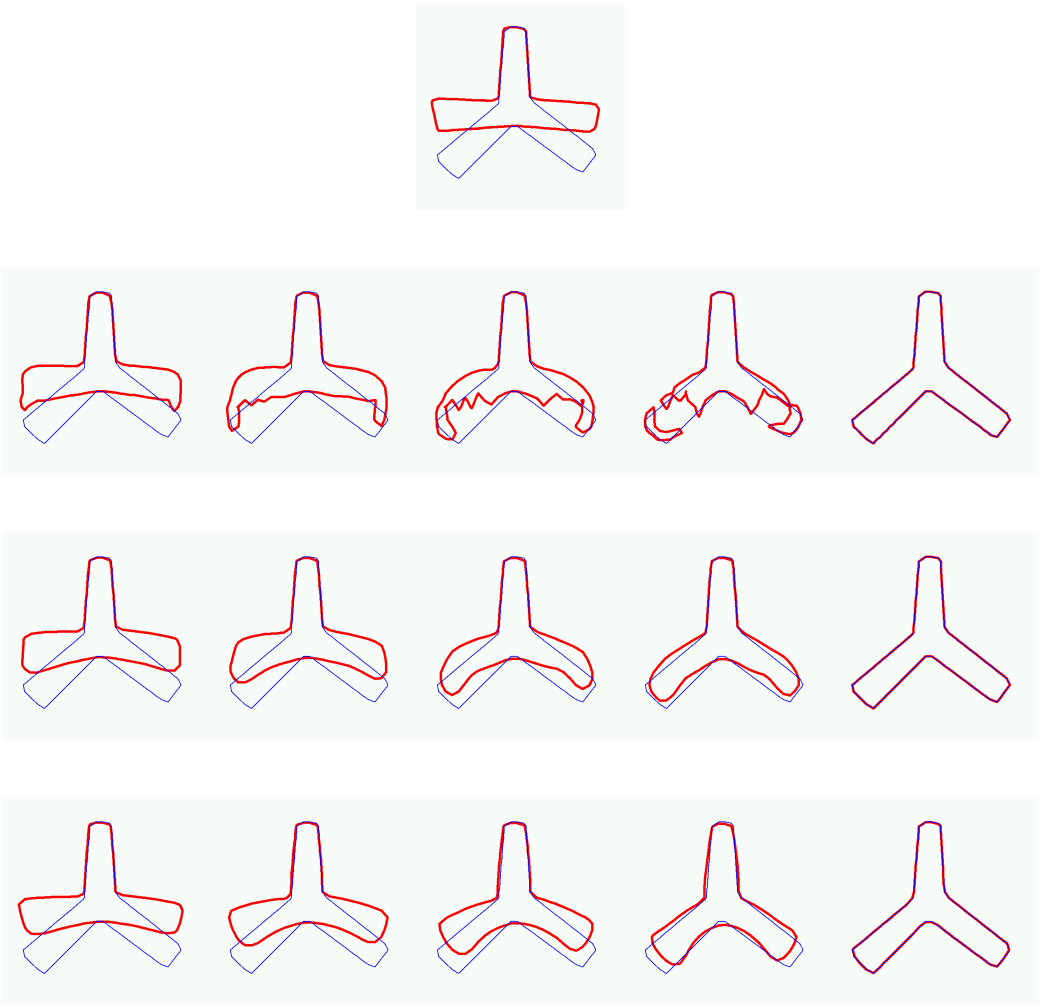


Figure 3.6: Warping the red shape onto the blue one *(top)* with the L^2 gradient descent *(first row)*, with a H^1 gradient descent *(second row)* and with a modified gradient descent favoring semi-local rigid motion *(third row)* for the same energy (Hausdorff distance). All evolutions converge to the same shape, but with different paths.

the evolution of ∇_R converges at each time step of the global evolution of Γ_1. Fortunately, when necessary, thanks to the remark in section 3.6.4, we can choose to stop the evolution of ∇_R before convergence in order to keep the additional cost into reasonable limits. The result presented here was computed so that the total evolution time was multiplied by two, but the

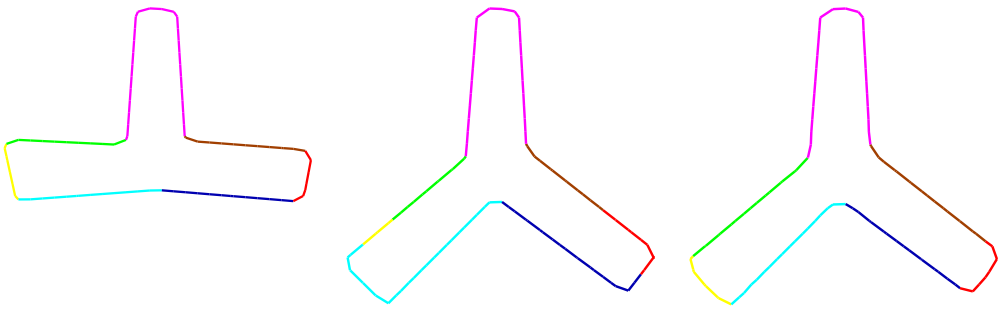


Figure 3.7: Comparison of the correspondences between the initial curve *(left)* and two final curves resulting from the H^1 gradient evolution *(middle)* and from the semi-local rigidification *(right)*. The different parts of the curves are shown with different colors, so that their respective evolutions can be followed. The correspondences for the semi-local rigidification case are more geometrically meaningful.

effect of the semi-local rigidification is already noticeable for an added cost time of 10%.

For the particular example presented in Fig. 3.6, one could object we should have considered other distances, such as the L^2 norm between the signed distance functions of the shapes, which always leads to very smooth evolutions. However, those smooth evolutions are not very sensible, in that they often miss entirely the similarity between the two shapes to match (see Fig. 3.8). As a consequence their gradient does not contain a lot of geometric information and cannot be very much improved by changes of inner products. This is why, despite the sometimes irregular behavior of the gradient of the Hausdorff distance, we prefer to use it in combination with new inner products, because this has both advantages of guaranteeing smoothness and making geometric sense.

In Fig. 3.9 we show an example with real contours from hand segmentation of pictures. As in the previous example, we show the evolution path obtained by minimization of the approximation of the Hausdorff distance between the two curves, with the semi-local rigidification approach. The evolution mainly consists in four local rotations (arms and legs), which fits well our intuition. We have colored, as previously, different parts of the initial curve in order to follow them through the evolution and notice how relevant the correspondences are. A usual gradient descent for this energy would have faced the same irregularity problems as in the first evolution presented in Fig. 3.6, and the choice of other usual energies, like the L^2 norm between

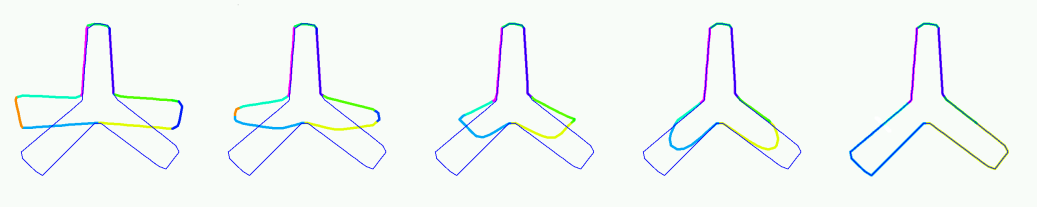


Figure 3.8: Comparison with a L^2 gradient descent on the L^2 norm of the signed distance functions associated to the curves. The gradient is naturally smooth but lacks geometric information.

the signed distance functions to the curves, would lack geometric sense, as in Fig. 3.8.

3.7 Conclusion

The impact of the inner product structure of the deformation space on the behavior of the active contours method had been overlooked so far in the computer vision community. We have explored several families of inner products, as well as some minimizing flows not deriving from any inner product by extending the notion of gradient. Given an energy, we now have several ways to minimize it, each of the proposed flows being a minimizing flow but leading to different kinds of evolutions. The inner products and the extended gradients should consequently be seen as priors on the deformation fields, that is to say priors on the evolution paths. They can be used to introduce different degrees of spatial coherence (local, semi-local or global) in the evolution of the contour. We have shown, with some numerical experiments, that these evolutions better fit our intuitive notion of deformation cost and that they can mimic the behavior of the objects of interest. As a result, they are at the same time more meaningful and more robust to local minima attraction.

In our experiments, we only considered shapes with no auto-occlusion. And yet, most of natural moving objects (e.g. the profile silhouette of a walking person) present auto-occlusion, the topology of the 2D shape is therefore always changing and as a consequence the matching between the corresponding 2D shapes is quite difficult. Future work could investigate the application of our methods on the 3D moving shapes and develop techniques to use the corresponding warping and matching for the projective 2D shapes.

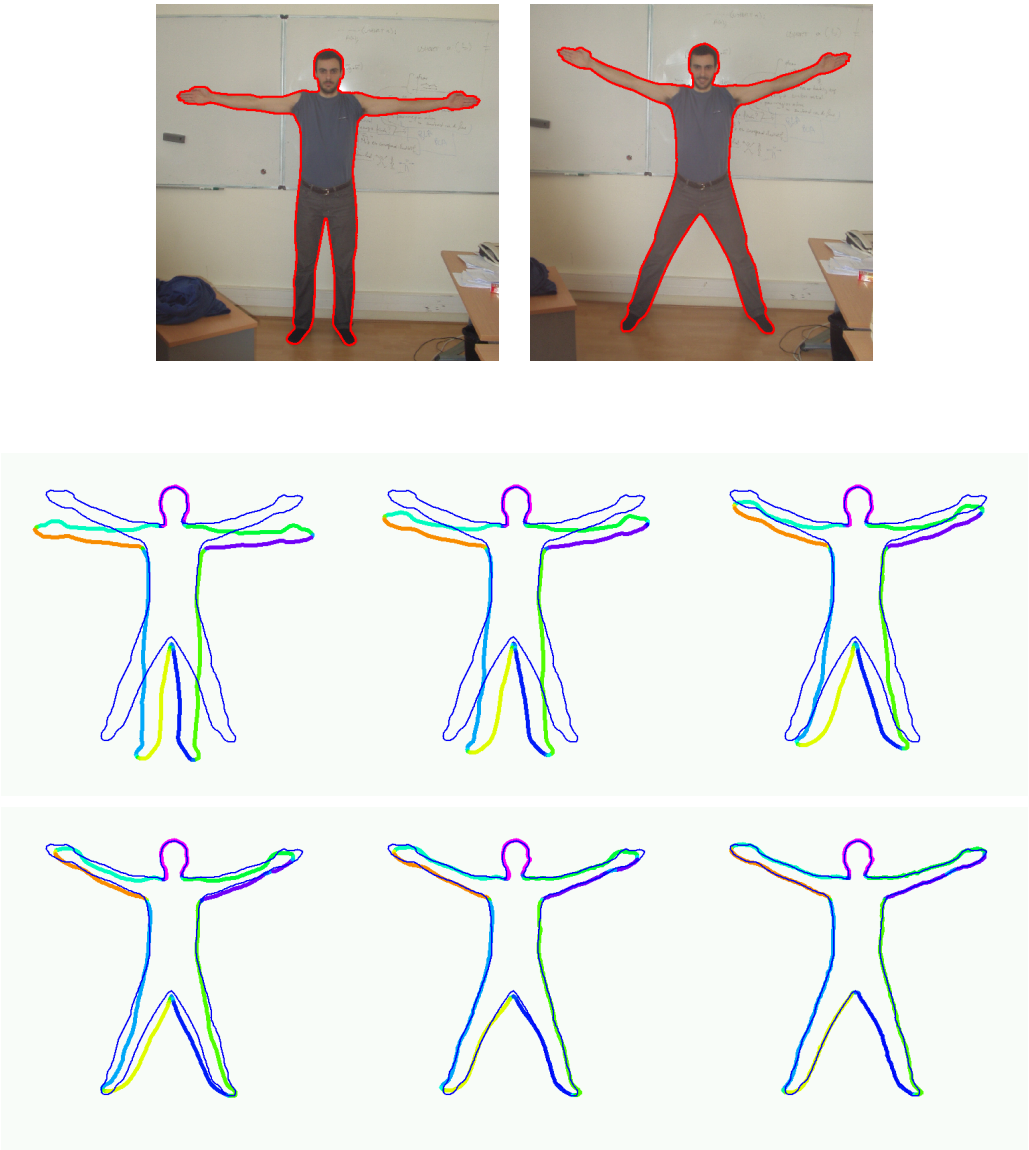


Figure 3.9: Warping real contours by minimization of the approximation of the Hausdorff distance with the semi-local rigidification approach. The colors show the correspondences between the moving curve and the initial one.

Chapter 4

Reconciling landmarks and level sets

Shape warping is a key problem in statistical shape analysis. This chapter proposes a framework for geometric shape warping based on both shape distances and landmarks. Taking advantage of the previously presented spatially coherent flows (see chapter 3), our method is mathematically well-posed and uses only intrinsic shape information, namely some similarity measure between shapes and the correspondence of landmarks provided on the shape surface. No extrinsic quantity is considered, neither a diffeomorphism of the embedding space nor point correspondences in this space. Thanks to a recent extension of the level set method allowing point tracking and tangential velocities, our method is compatible with implicit representations. Moreover, a matching between shape surfaces is provided at no additional cost. Although some recent work deals with implicit representations and landmarks, it is, to our knowledge, the first time that landmarks and shape distances are reconciled in a pure geometric level set framework. The feasibility of the method is demonstrated with two- and three-dimensional examples. Combining shape distance and landmarks, our approach reveals to need only a small number of landmarks to obtain improvements on both warping and matching.

Publication related to this chapter: *Reconciling landmarks and level sets* [93] with Renaud Keriven and Olivier Faugeras in the proceedings of the *International Conference on Pattern Recognition* held in 2006.

4.1 Introduction

As seen in chapter 2, the problem of shape warping is one of the keys leading to statistical shape analysis [15, 45]. It offers a way to compare shapes,

to compute their mean, to analyze their variability, and eventually to obtain correspondence between them. This problem, motivated by recognition, tracking or segmentation tasks, has been since the beginning formulated in purely *geometric* term of shape distance[37]. *Landmarks* have most of the time been part of the proposed methods, sometimes as guides, given by the user (e.g. anatomical landmarks) or geometrically determined (e.g. high curvature points), sometimes to define the shapes themselves (pseudo-landmarks between the landmarks). Among successful methods, let us cite the Procrustes analysis [69, 76], the thin-plate splines [16, 17] (still inspiring more recent work like [142]), the active shape models [30] (later extended into the active appearance models [29], in which the underlying images are considered).

Introduced as a way to cope with interface evolution simulation, the level set method [39, 108] is based on an implicit representation of surfaces. The natural choice for the implicit representation is often the signed distance function to the closed surface. In the original version of this method, the signed distance function was only a convenient choice to get an implicit representation of the initial surface. Then, some methods were designed [63, 115] to preserve this distance during the whole evolution: the evolving implicit function remains the distance to its zero level set. Consequently, the emergence of shape statistics in the implicit framework is not surprising. The pioneering piece of work considered the distance function as the only object of analysis: warping, matching, or statistical analysis were directly performed on the distance functions [87, 114, 145]. Nevertheless, mainly because the combination of two distance functions is not a distance function anymore, these methods could not be considered as definitive answers.

More recently, in a large collection of papers [90, 89, 86, 85], the authors proposed to mix implicit representations and landmarks. Yet, their work belongs to the diffeomorphic matching family: a space diffeomorphism is recovered thanks to the implicit representation of the landmarks (closed curves, but also open curves or points). It does not consider any surface evolution, thus, to our opinion, cannot be really seen as a true level set method application.

As we saw in chapter 1 the Hausdorff distance [37, 22] can also be used as shape similarity measure. The warping induced by the minimization of such shape distances can be improved in some manner using spatially coherent minimizing flows presented in chapter 3. The Hausdorff distance has always been one of the most considered shape similarity measure [37]. Hopelessly, trying to make a shape evolve toward another one minimizing their Hausdorff distance yields an irregular motion. That is why, the authors of [22] introduced a family of smoothing approximations of the Hausdorff distance that

makes this warping possible (see chapter 1). Although mathematically well justified, the resulting warping does not seem to be completely satisfying, in the sense that it does not reveal to be the one a human observer would have chosen. As shown in chapter 3, the *generalized gradient* method can be a way to correct this unwanted behavior.

Though, in the case of complex shapes, providing corresponding landmarks on both the initial and the target shapes reveals to be inevitable [55, 68]. In [1], the authors proposed a first try to modify a curve evolution in the implicit framework in a way that encourages the motion of the landmarks toward their respective targets. Turning back to simple ideas, the natural way to guide an evolution with landmarks is to try to minimize the distance between the landmarks on the evolving shape and the corresponding ones on the target shape. Again, this yields an irregular motion. In this chapter, we present a novel usage of the *generalized gradients* introduced in chapter 3 that turns this motion into a regular and well posed one.

Remarkably, recent advances in the level set method make our shape evolution compatible with it. To do this, two techniques are required. First, a way to simulate a Partial Differential Equation (PDE) embedded on a surface. This has been affordable for some times[9]. Second, a way to deal with *surface evolutions involving non normal velocities* (and, but this is related, to *track the surface points along time*). This last need is not usual in the level set framework, but a recent work[117] proposed a solution to that problem. As a result, we present what is, to our knowledge, the first usage of landmarks for true geometrical shape warping in the level set framework. Moreover, tracking points during the evolution gives us a matching (again, not necessarily one-to-one) between the shapes at no additional cost.

As a benchmark, we test the effect of adding our landmark-guided force to the distance-based spatially coherent evolution presented in previous chapters. Our method proves to be robust and gives the expected improvements: the obtained warping is more satisfying. Although not required, we track points in the original method too, and show that adding landmarks greatly improves the induced matching, even in cases where the original warping seems to be good enough. Again, our landmark-based evolution could be added to any other shape warping scheme. Yet, we have found the tested scheme, namely shape distance plus landmarks (plus eventually spatially coherent flows), an interesting combination that reveals to need only a small number of landmarks to obtain improvements on both warping and matching.

4.2 Landmarks-guided warping

In chapter 2 we presented the variational shape warping framework. In this chapter we will keep using the same notations. We place ourselves in the case where we want to warp an initial shape Γ_1 into a target one Γ_2 by minimizing a given shape energy $E(\Gamma_1, \Gamma_2)$ and by specifying some points on the two shapes that "correspond" each other. It means that we want those points to be matched at the end of the evolution.

The addition of landmark points can be necessary for example if the considered shapes are complex and we want to *guide the evolution* by this additional information. It can also be useful if the user want to specify some correspondences which are not only geometrical: one can think for instance of the case of two squares on which correspondence between corners are given by an external information (e.g. if the square come from a segmentation on two images), see Fig. 4.2.

Provided by the user (anatomical landmarks), or automatically extracted (geometric landmarks), we assume that we are given p pairs of corresponding points on the initial and on the target shapes :

$$\text{landmark points: } \left\{ (\mathbf{x}_{1,i}\ ,\ \mathbf{x}_{2,i}) \in \Gamma_1 \times \Gamma_2, 1 \leq i \leq p \right\} \tag{4.1}$$

We would like to use the information given by theses correspondences to guide the evolution.

4.2.1 A naive definition

As an example, let us suppose we want to provide landmark guidance to the evolution presented in chapter 2 given by equation (2.2):

$$\begin{cases} \Gamma(0) = \Gamma_1 \\ \dfrac{d\Gamma}{dt} = -\nabla_{\mathscr{D}} E(\Gamma, \Gamma_2) \end{cases}$$

where $\nabla_{\mathscr{D}} E(\Gamma)$ is the gradient of $E()$ relative to some inner product on $\mathscr{D}$. We do this by adding a landmark term $E_{\mathscr{L}}$ to the energy, obtaining a new energy, noted E_{tot}:

$$E_{tot}(\Gamma, \Gamma_2) = E(\Gamma, \Gamma_2) + E_{\mathscr{L}}(\Gamma, \Gamma_2) \tag{4.2}$$

The new equation of the evolution is then:

$$\begin{cases} \Gamma(0) = \Gamma_1 \\ \dfrac{d\Gamma}{dt} = -\nabla_{\mathscr{D}} E_{tot}(\Gamma, \Gamma_2) \end{cases} \tag{4.3}$$

A first idea would be to track the evolution of the landmarks with points $\mathbf{x}_i(t)$. We would like $\mathbf{x}_i(t)$ to be the position of the point on $\Gamma(t)$ which corresponds to $\mathbf{x}_{1,i}$ at time t :

$$\begin{aligned} \mathbf{x}_i(0) &= \mathbf{x}_{1,i} \\ \frac{\partial \mathbf{x}_i(t)}{\partial t} &= -\nabla_{\mathscr{D}} E_{tot}(\Gamma, \Gamma_2)(\mathbf{x}_i(t)) \end{aligned} \tag{4.4}$$

and to simply choose for $E_{\mathscr{L}}$ the L^2 square norm between the vectors $(x_i(t))$ and $(\mathbf{x}_{2,i})$:

$$E_{\mathscr{L}}(\Gamma(t), \Gamma_2) = \sum_{i=1}^{p} d(\mathbf{x}_i(t), \mathbf{x}_{2,i})^2 \tag{4.5}$$

Hopelessly, this is not well posed: the existence of point $\mathbf{x}_i(t)$ is not even guaranteed! The framework of viscosity solutions adapted to PDEs like equation (4.3) allows some points to disappear (see [31, 124]).

4.2.2 A correct definition

In fact, forward correspondences may not exist if the interface evolution forms shocks; the interface may even collapse and merely disappear. On the contrary, backward correspondences are guaranteed: each point of the evolving interface comes from one point at time 0 (see [117]).

We note $\psi_t : \Gamma(t) \to \Gamma_1$ the family of functions giving for each point $\mathbf{x}$ of $\Gamma(t)$ the point $\psi_t(\mathbf{x})$ on Γ_1 from which $\mathbf{x}$ comes. Let $\gamma_i(t)$ be the subset of $\Gamma(t)$ coming from $\mathbf{x}_{1,i}$ (possibly empty):

$$\gamma_i(t) = \psi_t^{-1}(\{\mathbf{x}_{1,i}\})$$

Equipped with this correspondence, we are now able to define a correct landmark-based energy as the sum, for each landmark of Γ_2, of the squared distance between this point and the corresponding set $\gamma_i(t)$:

$$E_{\mathscr{L}} = \sum_i d(\mathbf{x}_{2,i}, \gamma_i(t))^2 \tag{4.6}$$

with the convention that the distance to an empty set is zero. Note that some landmarks might disappear (shock) or become a continuous infinity of points (rarefaction). Actually, we conjecture that, for reasonable choices of the landmarks, rarefaction does not happen with smooth shapes. Yet, depending on the initial energy E, there might be some shocks, even with smooth shapes. Such considerations are beyond the scope of our work and

might still be open questions. So far, we can only mention that we have observed in our experiments that with the addition of our landmark-based energy $E_{\mathscr{L}}$ landmarks points are not likely to disappear.

In the sequel, we will suppose that either an initial landmark $\mathbf{x}_{1,i}$ remains one point $\mathbf{x}_i(t)$ ($\gamma_i(t) = \{\mathbf{x}_i(t)\}$), or it disappears ($\gamma_i(t) = \emptyset$). Under these hypothesis, the difference between our energy (equation (4.6)) and the "naive" one (equation (4.5) boils down to the possible disappearance of some landmarks:

$$E_{\mathscr{L}} = \sum_{\{i,\gamma_i(t)\neq\emptyset\}} d(\mathbf{x}_i(t), \mathbf{x}_{2,i})^2 \tag{4.7}$$

keeping in mind that point $\mathbf{x}_i(t)$ are not given by evolution (4.4), but come from the backward correspondences ψ_t.

4.2.3 Adapted gradient

Formally, the energy given by equation (4.7) yields Dirac peaks in the expression of the usual L^2-gradient of the energy:

$$\nabla_{L^2} E_{tot}(\mathbf{x}) \;=\; \nabla_{L^2} E(\mathbf{x}) \;+\; \sum_{\{i,\gamma_i(t)\neq\emptyset\}} \delta_{\mathbf{x}_i(t)}(\mathbf{x})(\mathbf{x}_i(t) - \mathbf{x}_{2i}) \tag{4.8}$$

where $\delta_{\mathbf{x}}$ denotes the Dirac function centered at point $\mathbf{x}$. This is indeed not a good candidate for a gradient descent.

The solution here is to change the inner product and to use the H^1 inner product as introduced in chapter 3. Let us detail what is actually a standard "trick" in numerical analysis. We change the inner product which appears in the definition of the gradient. Let $H^1(\Gamma, \mathbb{R}^n)$ be the Sobolev space of square integrable velocity fields with square integrable derivatives. We consider the canonical inner product of $H^1(\Gamma, \mathbb{R}^n)$:

$$\langle f | g \rangle_{H^1} = \int_\Gamma f(\mathbf{x}) \cdot g(\mathbf{x})\, d\Gamma(\mathbf{x}) + l^2 \int_\Gamma D_{\mathbf{x}} f(\mathbf{x}) \cdot D_{\mathbf{x}} g(\mathbf{x})\, d\Gamma(\mathbf{x})$$

where $D_{\mathbf{x}}$ denotes the intrinsic derivatives on the contour and l is a characteristic length for the derivation which acts as a weight between the two integrals. Recalling the definition of the gradient given chapter 2, we get:

$$\begin{aligned} \forall\, \mathbf{v}\,, \quad \langle \nabla_{L^2} E_{tot} \mid \mathbf{v} \rangle_{L^2} \;&= \mathcal{G}_\Gamma\big(E(\Gamma, \Gamma_2), \mathbf{v}\big) = \langle \nabla_{H^1} E_{tot} \mid \mathbf{v} \rangle_{H^1} \\ &= \langle \nabla_{H^1} E_{tot} \mid \mathbf{v} \rangle_{L^2} + \langle D_{\mathbf{x}} \nabla_{H^1} E_{tot} \mid D_{\mathbf{x}}\, \mathbf{v} \rangle_{L^2} \end{aligned}$$

We get that the H^1 gradient can be obtained from the L^2 gradient by solving an intrinsic heat equation with a data attachment term: $\nabla_{H^1} E_{tot}$ is solution of

$$\Delta_\Gamma\, u \;=\; u \;-\; \nabla_{L^2} E_{tot} \tag{4.9}$$

where Δ_Γ denotes the intrinsic Laplacian operator on the surface, often called the Laplace-Beltrami operator. The solution of this equation coincides with:

$$\arg\min_u \int_\Gamma |\, u(\mathbf{x}) \;-\; \nabla_{L^2} E_{tot}(\mathbf{x}) \,|^2 \, d\Gamma(\mathbf{x}) \;+\; \int_\Gamma |\;\; D_{\mathbf{x}} u(\mathbf{x}) \,|^2 \, d\Gamma(\mathbf{x}) \quad (4.10)$$

and the H^1 gradient is finally a smoothed version of the L^2 gradient, given by an image restoration process on a manifold Γ, a problem familiar to the image processing community. However, let us make two remarks:

- Introducing this smoothing via a modification of the gradient rather than directly from equation (4.10) warrants that the gradient descent will decrease the energy (as shown in chapter 3).
- We have knowingly omit to mention the space of admissible velocities. All the computations have been written formally. A rigorous demonstration would have involved the theory of distributions. Sketchily, in an appropriate distribution space, we should have considered the PDE obtained from (4.9) when replacing the formal symbol $\nabla_{L^2} E_{tot}$ by the second term of equation (4.8). The solution of this PDE can be shown to be a member of H^1 and exactly the desired smooth gradient $\nabla_{H^1} E_{tot}$.

In summary, starting from the irregular gradient $\nabla_{L^2} E_{tot}$ given by equation (4.8), we obtain a smooth gradient $\nabla_{H^1} E_{tot}$, given by the PDE (4.9) and mathematically justified by an adapted choice of inner product that guarantees a decrease of the energy.

4.2.4 Matching

Let us suppose that the warping process of Γ_1 into Γ_2 has converged. More precisely, we suppose there exists some time T such that $\Gamma(T)$ is very close to Γ_2 (e.g. $d_H(\Gamma(T), \Gamma_2) < \epsilon$ or $E_{tot}(\Gamma(T), \Gamma_2) < \epsilon'$), and a way to assimilate points of Γ_2 to points of $\Gamma(T)$ (e.g. taking the closest point[1]). Then, the backward correspondence ψ_T supplies a natural matching from Γ_2 to Γ_1. This matching is not one to one if some points of Γ_1 have disappeared during the evolution (shocks).

4.3 Level set implementation

Implementing our scheme in the level set framework [39, 108] requires two adaptations of the original method: implementing a PDE on an implicit

[1] This could be a problem if the evolution gets stuck into some local minimum. Yet, we have never experienced this case.

surface and being able to track points during the evolution.

4.3.1 The original level set method

Let us present briefly the broadly known level set method (we refer the reader to [108, 105] for more details). The key idea of the level set method is to use an implicit representation of the shape $\Gamma(t)$ i.e. to represent Γ as the zero level of some function $\phi(\mathbf{x}, t)$ defined for $\mathbf{x} \in \Omega$:

$$\Gamma(t) = \left\{ \mathbf{x} \in \Omega \mid \phi(\mathbf{x}, t) = 0 \right\}$$

Usually ϕ is negative inside $\Gamma(t)$ and positive outside. It can be proved that, if the level set function ϕ evolves according to:

$$\frac{\partial \phi(\mathbf{x}, t)}{\partial t} \; + \; v(x) \cdot \nabla \phi \; = \; 0 \tag{4.11}$$

then, its zero level, which is $\Gamma(t)$ by hypothesis, evolves according to the required equation $\frac{d\Gamma(t)}{dt} = v$. Here, $v(x)$ is defined on Ω and is equal to the desired velocity on $\Gamma(t)$ and is theorically arbitrary elsewhere (see below).

The advantages of the level set method are well known: stability, accuracy, convergence to the correct solution, easy extension to higher dimensions, correct handling of topological changes such as breaking and merging.

An important issue is that the velocity field $v \in \mathscr{D}$ is only defined on the shape Γ in the partial differential equation $\frac{d\Gamma(t)}{dt} = v$. In many cases, v has a natural extension everywhere in domain Ω, so that equation $\frac{\partial \phi(\mathbf{x},t)}{\partial t} \; + \; v(x) \cdot \nabla \phi \; = \; 0$ is defined. For instance, when $v(\mathbf{x})$ is the curvature of Γ at point $\mathbf{x} \in \Gamma$, one could choose, at each point $\mathbf{x} \in \Omega$, $v(\mathbf{x})$ equal to the curvature of the level set of ϕ going through $\mathbf{x}$. In some other cases, like ours, v can only be computed on Γ and some extension procedure has to be used to expand v everywhere in Ω. This is now classical [115, 2, 63].

Useful intrinsic geometric properties of the shape may be easily determined from the level function ϕ. For example, at any point of the shape, the normal vector $\overrightarrow{N}$ is given by:

$$\overrightarrow{N} = \frac{\nabla \phi}{|\nabla \phi|}$$

and the curvature κ is easily obtained from the divergence of the gradient of the unit normal vector to front, i.e.,

$$\kappa = \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = - \frac{\phi_{xx} \phi_y^2 - 2 \phi_x \phi_y \phi_{xy} + \phi_x^2 \phi_{yy}}{(\phi_x^2 + \phi_y^2)^{3/2}}$$

In our case, the velocity field we consider often involves the distance function to the shape Γ which we would like to be known without extracting the zero level of ϕ. Usually the initial value of ϕ is taken equal to the signed distance function to the initial shape $\Gamma(0)$. As a matter of fact this function verify all the necessary properties for a level set function. Therefore, $\phi(\cdot, 0)$ can be used to compute our velocity field at time $t = 0$. However, $\phi(\cdot, t)$ has no reason to remain the signed distance to $\Gamma(t)$. In fact the solution to equation 4.11 is not a distance function ([63]). In [63], the authors propose a method in which the implicit representation always remains a distance function by construction during the evolution. We chose to use this method and we can then use at any time t, $\phi(\cdot, t)$ as the signed distance function to $\Gamma(t)$.

Let us finally notice that in the original level set method, the velocity field is often considered as collinear with the normal field of the shape Γ. In our case, because of the addition of the landmark-based energy term $E_{\mathscr{L}}$, we will obtain a non-normal velocity field and it is important to keep the general expression of this velocity field.

4.3.2 H^1 gradient

We have shown in section 4.2.3 that the H^1 gradient is solution of equation (4.9). Since the work introduced in [9], implementing a PDE on a surface is affordable in the implicit framework. The only hard point in our case could be the Dirac peaks in the data term. We indeed use a smooth approximation of them.

It should also be mentioned that, in the two dimensional case, the explicit solution of the equation $\Delta_\Gamma\, u = u - v$ is known and given by:

$$u(x) = \frac{e^{\sigma(x)}}{2}\left(A - \int_0^{\sigma(x)} e^{-\sigma}\, v(\sigma)\, d\sigma\right) + \frac{e^{-\sigma(x)}}{2}\left(B + \int_0^{\sigma(x)} e^{\sigma}\, v(\sigma)\, d\sigma\right)$$

$$\text{with } A = \frac{e^{|\Gamma|}}{e^{|\Gamma|}-1} \oint_\Gamma e^{-\sigma}\, v(\sigma)\, d\sigma$$

$$\text{and } B = \frac{1}{e^{|\Gamma|}-1} \oint_\Gamma e^{\sigma}\, v(\sigma)\, d\sigma$$

where σ is the arc length and $|\Gamma|$ the length of the shape. Using this explicit solution might be attractive to avoid the iterative minimization giving u. Yet, it requires the extraction of the zero level set Γ of the implicit function, a process generally considered awkward in the level set community.

4.3.3 Point Correspondences

Because it codes interfaces with implicit representations, the original level set method can not follow the evolution of each point of the initial interface.

Only the geometric location of the whole interface is recovered. For the same reason, considered velocities are usually normal to the interface (or projected onto the normal for simplification reasons because it does not modify the location of the interfaces). In our case, we need to follow the landmark points through the backward correspondences ψ_t and to cope with the non normal velocity $-\nabla_{H^1} E_{tot}$. In [117], the authors propose an approach to maintain an explicit backward correspondence from the evolving surface to the initial one. Let the shape $\Gamma(t)$ be represented by a level set function $\phi(\cdot, t) : \mathbb{R}^n \to \mathbb{R}$, and $\mathbf{v}$ be the (non necessarily normal) velocity fields. ϕ evolves according to:

$$\begin{aligned} \phi(\mathbf{x}, 0) &= \tilde{d}_{\Gamma_1}(\mathbf{x}) \\ \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi &= 0 \end{aligned}$$

Let us consider a function $\Psi : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n$ such as:

$$\begin{aligned} \Psi(\mathbf{x}, 0) &= \mathbf{x} \\ \frac{\partial \Psi}{\partial t} + D\Psi \ \mathbf{v} &= 0 \end{aligned} \tag{4.12}$$

where $D\Psi$ stands for the Jacobian matrix of Ψ. It is shown in [117] that $\Psi(\mathbf{x}, t)$ holds the position that this point was occupying at time $t = 0$. Our needed backward correspondence is then straightforward: $\psi_t(\mathbf{x}) = \Psi(\mathbf{x}, t)$.

4.4 Experiments

As a benchmark, we warp some artificial two dimensional curves with the original energy $E = d_{W^{1,2}}$ and test how our landmark-guided force modifies the warping and the final matching. Figure 4.1 shows the warping of a rectangle into another one. The different parts of the curves are shown with different colors, so that their respective evolution can be followed. Although the initial warping without any landmark seems natural, it fails discovering the matching between the edges of the rectangles, a matching indeed recovered when providing landmarks.

Fig. 4.2 shows the warping of a square onto another one, the energy is here the Hausdorff distance and we used the spatially coherent flows in addition of landmarks on the corners.

Figure 4.3 shows the warping between two hand shapes. The energy $E = d_{W^{1,2}}$ yields an unnatural warping. Adding spatially coherent flows makes the warping a bit better but still fails in some parts, mainly because

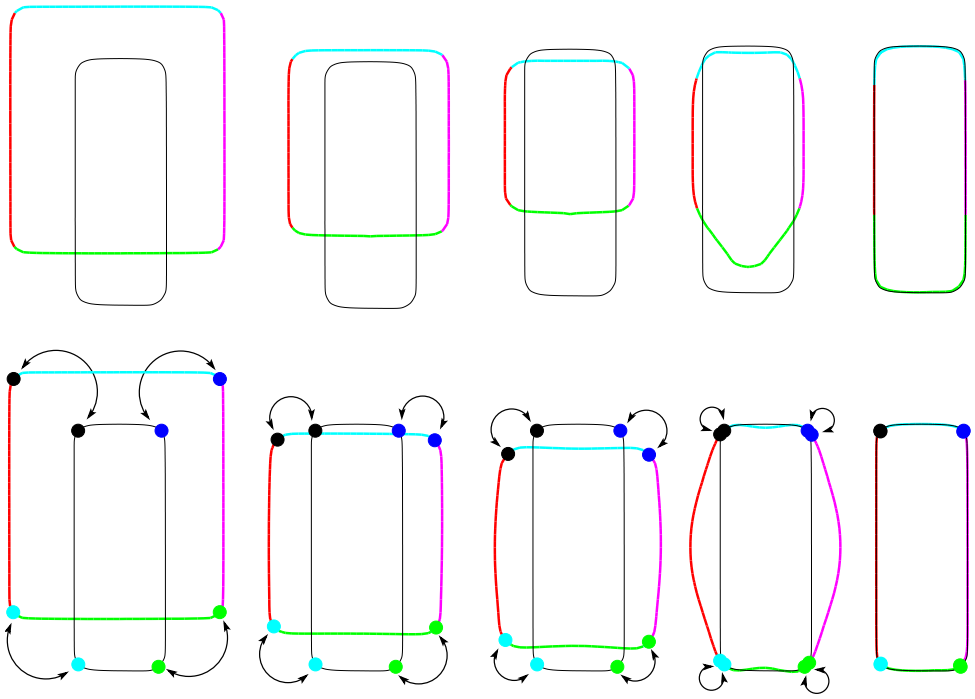


Figure 4.1: Warping of a rectangle shape into another one. Top row: evolution with $E = d_{W^{1,2}}$. Bottom row: evolution with the same energy, augmented with four provided landmarks, marked by color spots. The colors on the evolving curve shows the evolution of different parts of it. See text for comments.

the difference between the two shapes can not be summed up to a global motion. With three landmarks only, both a satisfying warping and a good matching are recovered.

Figure 4.4 shows the warping of a teddy bear into a cartoon character. Without any landmarks, the top row evolution fails matching the ears and arms of the characters. The bottom row shows the evolution with four landmarks. Red spots allow to check a good matching between landmarks.

4.5 Conclusion

We proposed a framework for shape warping based on both shape distances and landmarks. Our method is purely geometric and no extrinsic quantity like a space diffeomorphism has to be considered. Thanks to recent advances in the level set techniques, a level set implementation is possible, reconciling

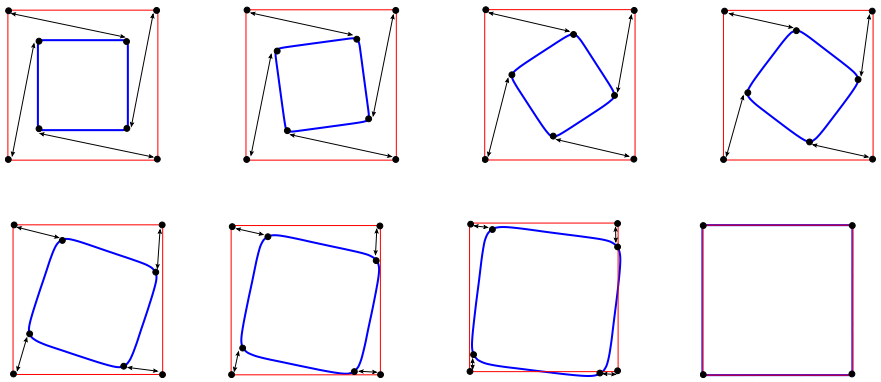


Figure 4.2: Warping of the blue square into the red one. Evolution with the Hausdorff distance, plus spatially coherent flows, plus landmarks points on the corners (marked by black spots).

landmarks and the level set methods. Moreover, a matching between shapes is provided at no additional cost. Two- and three-dimensional examples, combining shape distance and landmarks, demonstrate the improvement brought by our approach on both warping and matching, even with a small number of landmarks. Further work includes investigating for a one-to-one matching between shapes, and a way to cope with other landmarks, such as curves on surfaces in $\mathbb{R}^3$.

One could say that this method implies a loss of autonomy in the warping operation since several landmark points have to be specified. However, in addition of allowing to process more complex shapes, let us stress again that only a few number of landmark points is enough to largely improve the results (in our examples, the number of specified landmarks is typically less than 5). Moreover we can imagine a way to stay completely automated by using a detector of characteristic points (e.g. points of high curvature) and by taking the result of this detection as potential landmarks points.

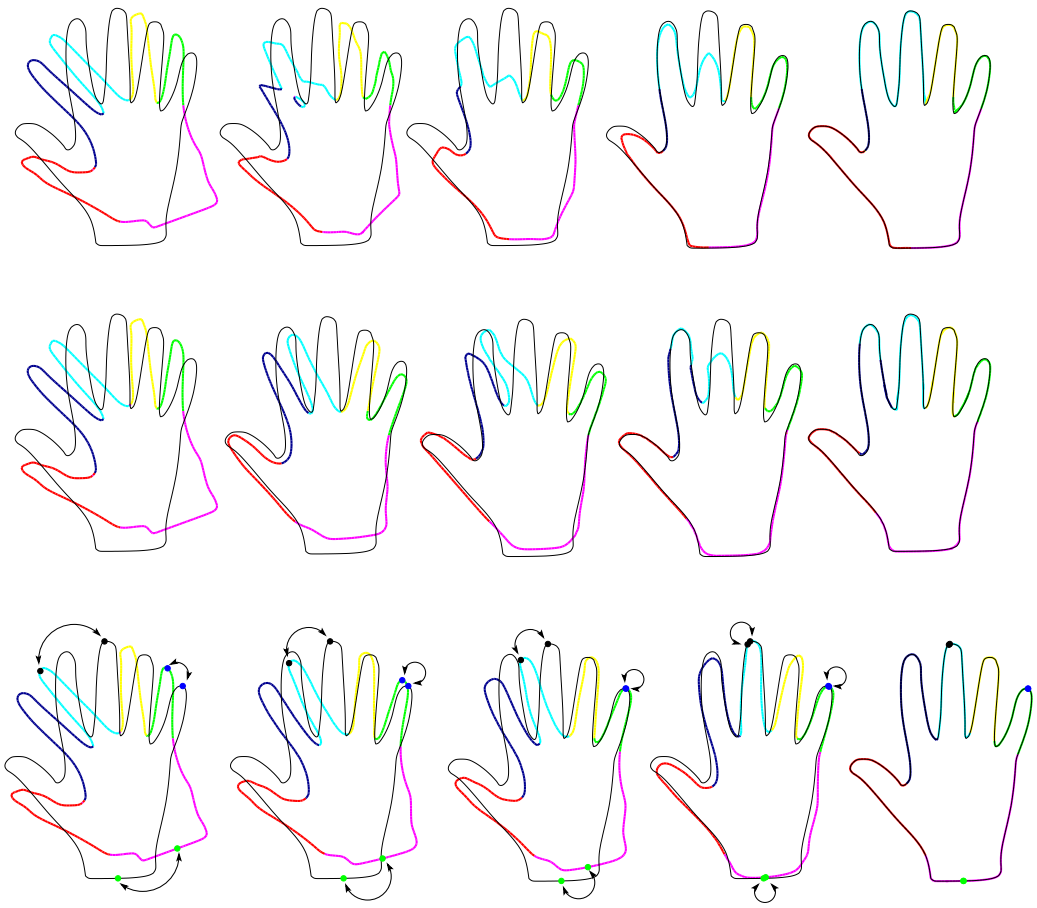


Figure 4.3: Warping of a hand shape into another one. Top row: evolution with $E = d_{W^{1,2}}$. Middle row: evolution with the same energy plus spatially coherent flows. Bottom row: evolution with the same energy plus coherent flows plus three provided landmarks. See text for comments.

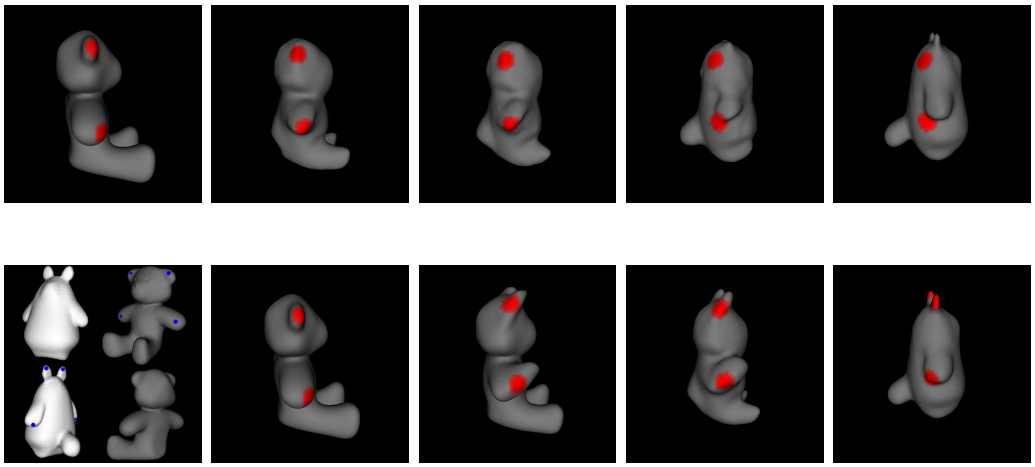


Figure 4.4: Warping of a teddy bear into a cartoon character. Top row: evolution with $E = d_{W^{1,2}}$. Bottom row, first image: four landmarks provided on the two shapes, indicated by blue spots. Bottom row, remaining images: evolution with $E = d_{W^{1,2}}$ plus the provided landmarks. In red, some parts of the shapes are tracked. See text for comments.

Part II

Facial expression and epilepsy

Chapter 5

The context

Epilepsy is among the most common serious brain disorders, can occur at all ages, and is characterized by a variety of presentations and causes. In this chapter we will present the context of our work on facial expression analysis during epileptic seizures. After a general presentation of epilepsy and epileptic seizures, we will characterize our specific problem.

5.1 General presentation of epilepsy and epileptic seizures

5.1.1 Definitions

Epilepsy is the name of a brain disorder characterized predominantly by recurrent and unpredictable interruptions of normal brain function, called epileptic seizures. Epilepsy is not a singular disease entity but a variety of disorders reflecting underlying brain dysfunction that may result from many different causes. In [56] the authors propose a general definition of those two terms:

Definition 17. *An **epileptic seizure** is a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain.*

The epileptic neuronal activity is a specific dysfunction, characterized by abnormal synchronization, excessive excitation and/or inadequate inhibition, and can affect small or large neuronal populations. The clinical manifestations are sudden, transient and usually brief. They include motor, psychic, autonomic and sensory phenomena, with or without alteration in consciousness or awareness. As we will see later, the symptoms depend on the part

of the brain involved in the epileptic neuronal discharge, and the intensity of the discharge.

Definition 18. ***Epilepsy*** *is a disorder of the brain characterized by an enduring predisposition to generate epileptic seizures and by the neurobiologic, cognitive, psychological, and social consequences of this condition. The definition of epilepsy requires the occurrence of at least one epileptic seizure.*

Etymologically the word *seizure* comes from the Greek meaning *to take hold*. Nowadays this word is used to design any sudden and severe event. Many of these paroxysmal events can look like an epileptic seizure but do not involve abnormal, rhythmic discharges of cortical neurons. They can be caused by either physiological or psychological conditions. To emphasize this difference we will preferentially refer to an *epileptic* seizure. Epileptic Seizures can be divided into two types according to whether the source of the seizure within the brain is localized, *partial seizures*, or distributed throughout the cortex, *generalized seizures*.

5.1.2 Epidemiology

Epidemiology is a branch of medical science that deals with the incidence, distribution, and control of disease in a population. In epidemiology context, the *incidence* is a measure of the risk of developing some new condition within a specified period of time and the *prevalence* of a disease is defined as the total number of cases of the disease in the population at a given time divided by the number of individuals in the population.

The estimated incidence of epilepsy is approximately one case per 2000 persons in the Western population per year and the prevalence of active epilepsy (with recent seizures) is around $5 - 10$ per 1000. Yet unexplained, the incidence of epilepsy is highest in the first year of life and for those over 50 years of age (see Fig. 5.1). The cumulative incidence, that is the chance of having epilepsy during a lifetime of 80 years, is about 3%. The fact that prevalence is much lower than cumulative incidence demonstrates that in many cases epilepsy remits. In fact within 5 years of the onset of seizures, $50 - 60\%$ of patients will have entered long remission. However, in about 20% of cases, epilepsy, once developed, never remits.

5.1.3 The causes of epilepsy

Epilepsy is often multifactorial. The range of causes can vary in different age groups, patient groups and geographical locations. Broadly speaking, *genetic, congenital, and developmental* conditions are the most common causes

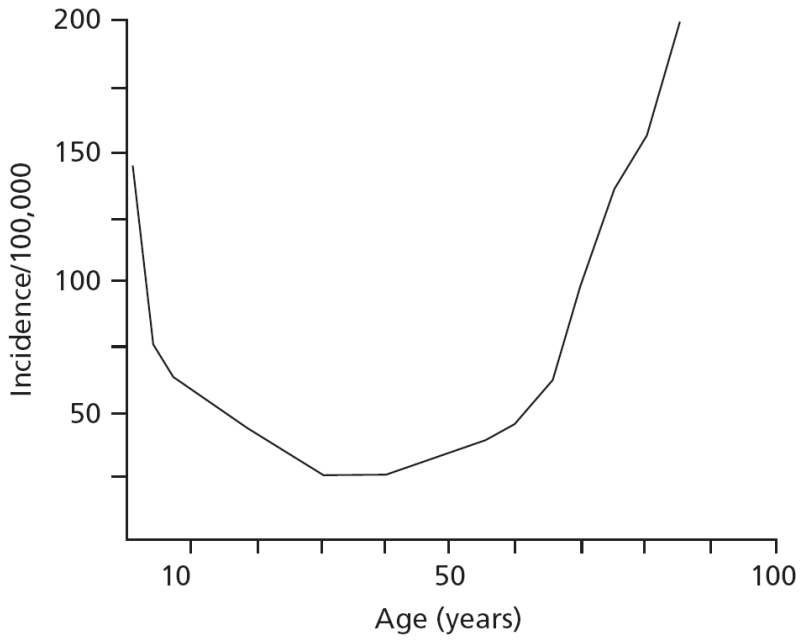


Figure 5.1: Age-specific incidence rates based on combined results from studies in USA, Iceland and Sweden (figure taken from [126])

of early childhood onset epilepsy, whereas in adult life epilepsy is more likely to be due to *external nongenetic* causes [126]: tumor and vascular disease is increasingly common. Head trauma and central nervous system infections may occur at any age and in certain parts of the world, endemic infections (such as tuberculosis, cysticercosis, HIV) are common causes. There are still many people for whom the cause of their epilepsy cannot, as yet, be identified. In such cases, the theory most commonly accepted is that this epilepsy is the result of an unbalance of certain chemicals in the brain (especially chemical messengers known as neurotransmitters) causing them to have a low convulsive threshold.

5.1.4 Why and how to treat epilepsy ?

(a) The aims of drug treatment

- Seizure control: The antiepileptic drugs are remarkably effective in suppressing seizures. Population surveys have shown that between 80 and 90% of patients will gain $1-2$ year remissions from seizures when

therapy is started, and long-term remission is gained in about 70% of all patients treated with antiepileptic drugs.

- Avoidance of social consequences of epilepsy and secondary handicap: epilepsy has a number of potential social consequences which can be more important than the immediate effects of individual seizures. "Being epileptic" can be far worse than simply having seizures.

- Reduction of mortality and morbidity: Epilepsy is a potentially life-threatening condition. Deaths associated with epilepsy can be classified into three categories: (i) those caused directly by the seizures (such as accidental death or SUDEP: sudden unexpected death in epilepsy), (ii) those related indirectly, or only partly, to epilepsy (for instance suicide), and (iii) those due to other factors, for example the underlying causes of the epilepsy (such as a tumor or infection). Successful antiepileptic drug (AED) therapy should prevent deaths in the first category, may prevent some deaths in the second category, but will have no preventive effect in the third category.

(b) Pharmacologic treatment

The first 5 years of treating epilepsy in patients with new-onset epilepsy is crucial. The modern antiepileptic drugs (AEDs) provide satisfactory control of seizures for most patients with epilepsy. According to [49], about 65% of patients with new-onset epilepsy respond, seizure recurrence occurs in 5%, and 35% have uncontrolled epilepsy (see Fig. 5.2).

(c) Surgical treatment

For patients with epilepsy not controlled by adequate attempts with multiple medications, surgical treatment can be an option. Generally patients selected for a pre-surgical observation present a partial epilepsy because this kind of epileptic seizures comes from a relatively localized area in the brain. This area could possibly be removed by a surgical operation. The pre-surgery exploration consists principally in trying to accurately locate the area at the origin of the seizures (see the next section) and assessing whether removal of concerned brain tissue will result in unacceptable problems with memory, vision, language or movement.

The evolution of the surgical treatment of epilepsy in the past century has been dependent on technical developments. The first resective surgery was performed in 1880, following improvements in anesthetics and surgical

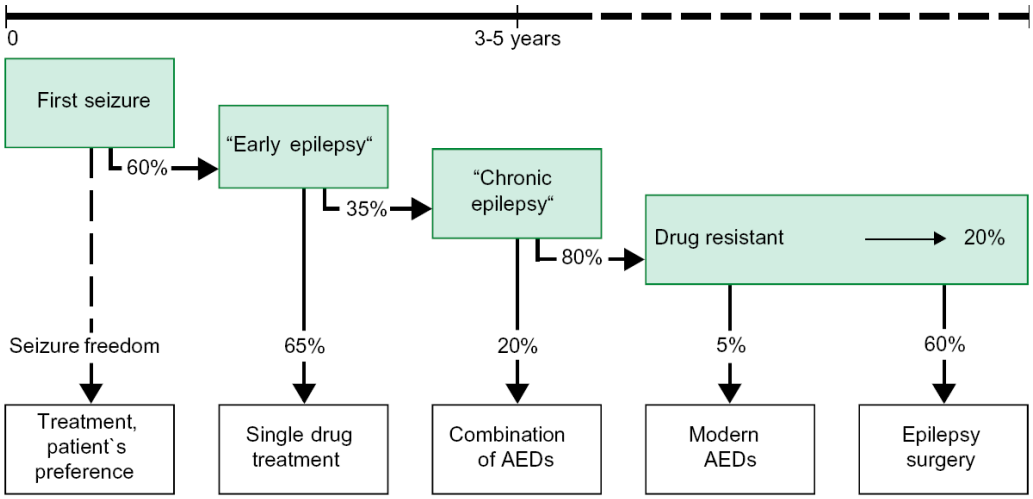


Figure 5.2: Overview on management principles and estimates of seizure freedom after treatment in the first $3-5$ years of new-onset epilepsy (figure taken from [49])

instrumentation. In the late 1930s, the introduction of Electroencephalography (EEG) provided the first objective method for localizing epileptic tissues. EEG and clinical localization were combined and this approach, which is still the basis of epilepsy surgery today, improved both diagnostic accuracy and surgical outcome. Nowadays approximately 80% of patients who had a surgical operation will not present any seizure in the next three years. But this good statistics should be relativized by the fact that the most important determinant of a successful surgical outcome is the good pre-selection of patients.

5.2 Our specific problem

In La Timone hospital, Marseille, France, Dr. Patrick Chauvel and his team receive patients in pre-surgical exploration. These patients present a pharmacoresistant epilepsy: medical therapy failed to sufficiently control seizures, and therefore patients no longer have the ability to maintain their quality of life. Hence the possibility of an surgical treatment is considered and patients stay under observation from several days to few weeks.

5.2.1 Clinical tools

(a) Magnetic resonance imaging (MRI)

This is a medical imaging technique introduced in the early 1980's. It produces a two- or three-dimensional view of a body part, the brain among others. Each patient with new-onset epilepsy usually have an MRI to detect possible structural lesions caused by for example cortical malformation, traumatic brain injury, brain tumor, and cerebrovascular disease, which are the most common causes of *symptomatic* (with a known cause, by opposite of *idiopathic*) epilepsy.

(b) Electroencephalography (EEG)

EEG is the measurement of electrical activity produced by the brain as recorded from electrodes placed on the scalp. It is a completely painless and non-invasive method. Human EEG were first done by Hans Berger in the 1920's, but its use for medical purpose and in particular epilepsy really began from the 1950's. EEG presents a number of advantage, compared for example to functional MRI: the subject does not have to hold as still during the recording and the temporal resolution is much better. However since the electrical activity is recorded from the surface of the scalp, the spatial resolution is poor and the problem of reconstructing intracranial sources for a given EEG signal, referred as the *inverse problem* is very difficult.

(c) Stereoelectroencephalography (SEEG)

This is the recording of electroencephalographic signal via *depth electrodes*, i.e. electrodes surgically implanted into the brain tissue. 12 to 14 electrodes, with 10 to 15 recording points each, are usually implanted in several locations selected using a preliminary EEG study. It has a much higher spatial resolution than surface EEG and thus contributes to define with accuracy the boundaries of the "epileptogenic zone", i.e. the area of brain generating the seizures.

(d) Video-EEG or Video-SEEG

In video-EEG (video-SEEG), patients are videotaped at the same time as their EEG (SEEG) is recorded. The recording is carried out for a long period of time, often several days. The doctor usually views the video and EEG (SEEG) images side by side. In this way the doctor can see precisely how the behavior of the patients during seizures is related to the electrical activity in the brain. This is closely related to the semiology of the seizure.

5.2.2 Semiology of the epileptic seizures

The semiology is the branch of linguistics concerned with signs and symptoms. In our case it designs the study of every sign and symptom, and their temporal evolution, during the epileptic seizure. During focal epileptic seizures patients may manifest abrupt modifications of behavior [126] such as: motor manifestations (e.g. jerking, spasm or posturing), special sensory manifestations (e.g. tingling, numbness, pain or feeling of heat), autonomic manifestations (e.g. changes in skin color, blood pressure, heart rate, pupil size and piloerection), psychic manifestations (e.g. dysphasic symptoms, disturbance of memory, affective symptoms including fear, depression, anger and irritability, ...).

The observation, study and analysis of all the signs and symptoms occurring during the epileptic seizure is an important tool for the clinical team. Along with the EEG or SEEG recordings, it constitutes a crucial *set of clues* which can confirm or refute a localization of a possible epileptogenic area. For this reason, epileptologists pay very careful attention to these signs during patient observation: the seizures are videotaped and can thus be rewatched, the patients are encouraged to described in detail their sensory experiences and their feelings. In addition examination is carried out where possible during and immediately after the seizure, allowing assessment of various aspects including level of responsiveness, language function, memory as well as motor and sensory function.

In spite of the fact that the EEG or SEEG recordings are naturally the most important source of informations among this *set of clues*, other signs are fundamental too. This set of clues can be very useful to get a rough idea about in which part of the brain is located the epileptogenic zone and therefore it can for example enable the clinical team to reduce the number of SEEG-electrodes implanted in the brain. If only for that reason, it proves its importance: the implantation of SEEG-electrodes is already a surgical operation with the associated potential risks.

5.2.3 Facial expressions during epileptic seizures

The appearance of facial and ocular motor signs during seizures, including modification of facial expression, would appear to offer important opportunity for detailed semiological study and electro-clinico-anatomical correlations. Indeed facial emotional expression has long been considered a subject of scientific interest, with work in the 19th century including Charles Darwin [34] and Duchenne amongst others being the point of departure for research in this domain. Perhaps surprisingly, relatively little detailed study has so far

been performed on the semiology of facial expression that may occur during epileptic seizures. One of the main expressions studied so far has been that of fear particularly in the context of frontal lobe seizures ([10], see Fig. 5.3).

However as well as the study of complex facial expressions, analysis of individual components of facial motor modifications during seizures could be of great interest. Although certain facial signs such as in "versive" seizure (meaning a partial seizure associated with head and eye deviation to one side) or asymmetric facial contraction may occur relatively commonly in seizures, little previous work has focused on the electroclinical correlations of such semiological features. Previous authors have considered aspects such as the value of eye version in lateralising frontal seizures [6, 72] and observed the pattern of eye closure in partial epileptic seizures as compared to in non-epileptic seizures of psychogenic origin ([40]) but more detailed analysis of patterns of facial movements with regards to EEG data is so far lacking. In addition the subtle and sometimes transitory nature of certain facial signs can make analysis difficult even with the benefit of simultaneous video-EEG recording, especially with regards to comparison between different patients presenting similar facial modifications.

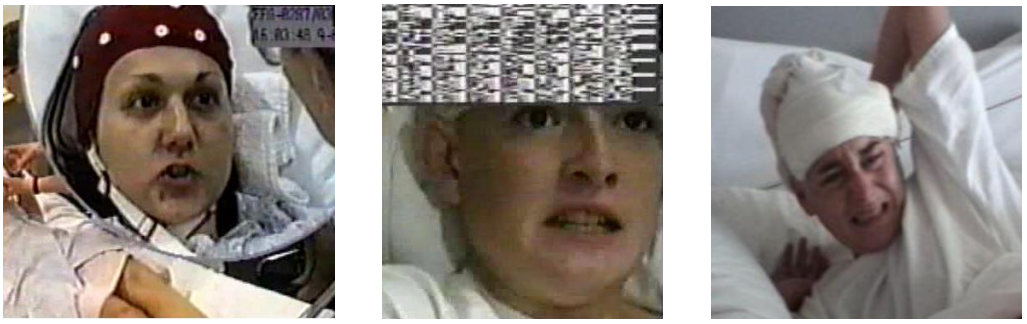


Figure 5.3: Example of fear expression during epileptic seizures

This aspect leads Dr Chauvel and his team to search for a way to quantify this symptom: the facial expression. This is the purpose of this part of our work: to develop a tool for automatically analyzing the facial expressions during the epileptic seizure (see chapter 6) and quantifying them in order to enable for example the numerical correlation between them and other signals, such as SEEG recordings (see chapter 7).

Chapter 6

Facial expression analysis: 3D model fitting

This chapter addresses the recovering of 3D pose and animation of the human face in a monocular single image under uncontrolled imaging conditions. Our goal is to fit a 3D animated model in a face image with possibly large variations of head pose and facial expressions. Our data were acquired from filmed epileptic seizures of patients undergoing investigation in the videotelemetry unit, La Timone hospital, Marseille, France [1].

Publication related to this chapter: *3D model fitting for facial expression analysis under uncontrolled imaging conditions* [94] in the proceedings of the *International Conference on Pattern Recognition* that will be held in December 2008.

6.1 Introduction

Facial expression analysis has been an active research topic for behavioral scientists and psychologists since the work of C.Darwin in 1872 [34, 48]. In 1978, Suwa et al. [134] presented a preliminary investigation on automatic facial expression analysis by tracking the motion of several identified spots on an image sequence. Since then considerable progress has been made in building computer systems that attempt to automatically analyze and recognize facial motions [53, 110].

Two principal classes of approaches have been developed for facial expression analysis: Image-based [12, 97] and Model-based approaches [41, 44]. Image-based methods extract features from images without relying on elab-

[1] Written informed consent was obtained from all patients for the use of their video recordings, including for publication.

orate knowledge about the object of interest, here, a face. Their principal qualities are their quickness and their simplicity. However, if the data images are very diverse (e.g. variation of illumination, of view, of head pose) image-based approaches can become erratic and unsatisfactory. On the other hand model-based methods use models which maintain the essential characteristics of the face (position of the eyes relative to the nose for example), but which can deform to fit a range of possible facial shapes and expressions.

In this work, we are interested in analyzing the facial expression of several patients during epileptic seizures. In fact, detailed study of such facial expressions produced during epileptic seizures could help in understanding the cerebral organization of the seizures (see chapter 5). Because of the unsupervised nature of the data acquisition (a fixed camera in a hospital room), we chose to use a model-based approach. A large class of methods developed in the last decade was based on the Active Appearance Model [29] and more recently on 3D Morphable Model [13]. These methods construct a model from a learning set of several images of different persons showing different expressions. A certain number of points of interest has to be placed in the images of the learning set and a Principal Component Analysis is applied: on the position of the points in the case of the Active Shape Model [30] and on the position of the points and on the textures of the images in the case of the Active Appearance Model. In our case, the expressions of the epileptic patients during their crises could be individual and complex and consequently a model built from a learning set of common expressions (typically such as anger, sadness or happiness) would not have been sufficient. This is why we chose to use the Candide [3] model in our work.

We first introduce the 3D face model we used in this work. Then we present our method to fit this model on a facial image. The next section deals with the analysis of a facial expression. The final section shows some of the results on the real data.

6.2 The 3D face Model

6.2.1 Candide face model

(a) The three dimensional mesh

In this work we use a modified version of the Candide 3D face model [3]. Candide is a parameterized face mask specifically developed for model-based coding of human faces. The original Candide Model contains 113 vertices and 184 triangles. Fig. 6.1 shows the mesh of the model. Nevertheless our

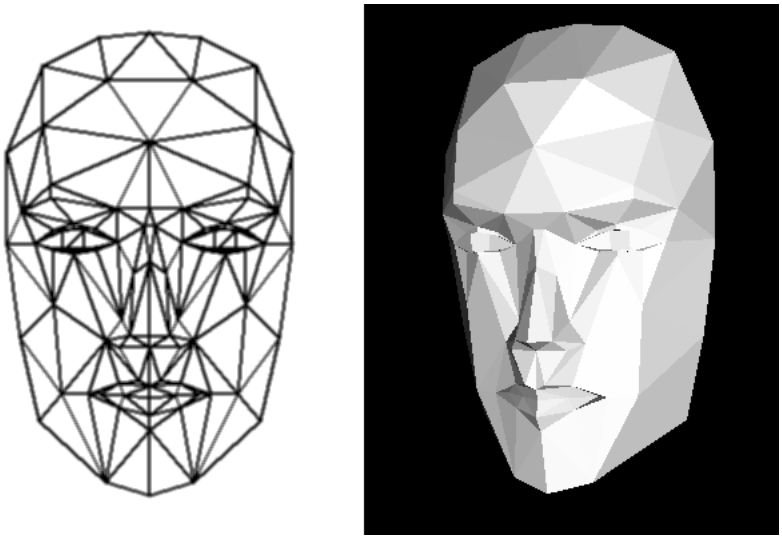


Figure 6.1: The original Candide face model

method does not depend on the choice of the model and we state these two notations:

$(\overline{P}_i)_{1\leq i\leq n_p} \in (\mathbb{R}^3)^{n_p}$ is the set of n_p points of the initial mesh $\overline{\mathcal{M}}$.

$\big(T_i = (a_i, b_i, c_i)\big)_{1\leq i\leq n_t} \in (\mathbb{N}^3)^{n_t}$ the set of n_t triangles of the mesh $\mathcal{M}$.
It means that the i-th triangle of the mesh has $P_{a_i}, P_{b_i}, P_{c_i}$ as vertices.

(b) Controlling the model

To control the model, 14 shape units, noted $(\mathcal{SU}_i)$, and 71 animation units, noted $(\mathcal{AU}_i)$, are also provided. Fig. 6.2 shows some examples of these deformation vectors, and we have the following notations:

$$\Big(\mathcal{SU}_i = \big(\mathbf{v}^s_{i,1}, \ldots, \mathbf{v}^s_{i,n_p}\big)\Big)_{1\leq i\leq n_{su}} \text{ are the } n_{su} \text{ shape units,}$$

$$\Big(\mathcal{AU}_i = \big(\mathbf{v}^a_{i,1}, \ldots, \mathbf{v}^a_{i,n_p}\big)\Big)_{1\leq i\leq n_{au}} \text{ are the } n_{au} \text{ animations units,}$$

where $\mathbf{v}^s_{i,j}$ and $\mathbf{v}^a_{i,j} \in \mathbb{R}^3$ are the deformation vectors for the point P_j :

$$\forall i \in \{1, \ldots, n_{su}\},$$
$$\mathcal{SU}_i(\lambda) : P_j \mapsto \overline{P}_j + \lambda \mathbf{v}^s_{i,j} \qquad \forall j \in \{1, \ldots, n_p\}$$

and similarly for $\mathcal{AU}_i$:

$$\forall i \in \{1, \ldots, n_{au}\},$$
$$\mathcal{AU}_i(\lambda) : P_j \mapsto \overline{P}_j + \lambda \mathbf{v}^a_{i,j} \qquad \forall j \in \{1, \ldots, n_p\}$$

Therefore we can write:

$$\mathbf{g} = \overline{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha \tag{6.1}$$

where $\overline{\mathbf{g}}$ is a $3n_p$ dimensional vector containing the $(\overline{P}_{ix}, \overline{P}_{iy}, \overline{P}_{iz})$ coordinates of the vertices. And the column of $\mathbf{S}$ and $\mathbf{A}$ are the Shape and Animation units respectively and thus the vectors σ and α contain the shape and animation parameters.

We also want to perform global motion, we need a few more parameters for rotation, scaling, and translation. Thus, we replace 6.1 with:

$$\mathbf{g} = s\mathbf{R}(\overline{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha) + \mathbf{t} \tag{6.2}$$

where $\mathbf{R}$ is a rotation matrix, s is the scale, and $\mathbf{t}$ a translation vector. The geometry of our model is thus parameterized by the parameter vector: $\mathbf{p} = [\mathbf{v}, \sigma, \alpha]$, where $\mathbf{v}$ is the vector of global motion parameters:

$$\mathbf{v} = [\theta_x, \theta_y, \theta_z, s, t_x, t_y, t_z].$$

(c) Our modifications

For our work we chose to slightly modify the original model, principally in order to remove the top part of the head (see Fig. 6.3) and to add some specific animation units. These modifications were motivated by the specificity of our images. The patients in fact wore EEG or SEEG electrodes on the head and the forehead was often hidden. Furthermore, the expressions encountered during the crises were often asymmetric and some additional animation parameters were necessary to cover all the scope of the possible expressions. For example the last animation unit presented in Fig 6.2 (on the bottom right) was not in the initial model.

6.2.2 The Reference Texture

We also added a *reference texture* to the model. This texture has been computed as the average between a few number of faces on which the model has been manually placed. We chose to use a monochromatic texture. Fig. 6.4 illustrates this process and shows this *reference texture* on the model. We note I_{ref} this image and $(P_i^{ref})_{1 \leq i \leq n_p}$ the points of the 2D mesh on this image (the triangles are the same as for the 3D one).

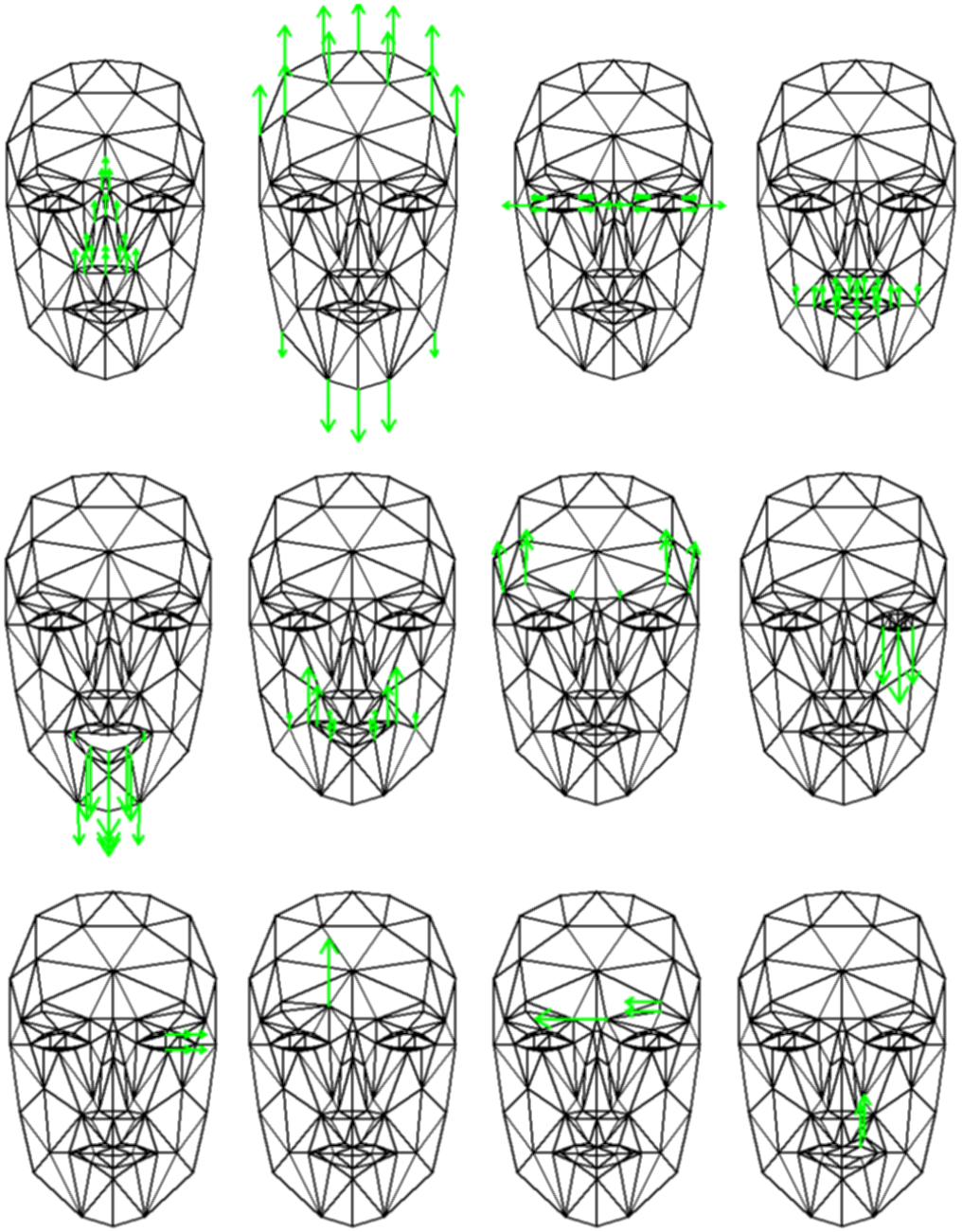


Figure 6.2: first row: four examples of shape units, second and third row: eight examples of animation units

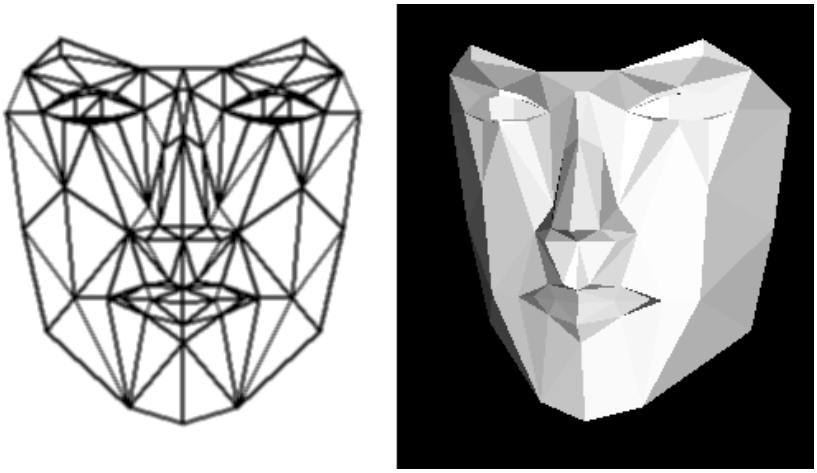


Figure 6.3: Our modified version of the mesh of the Candide face model

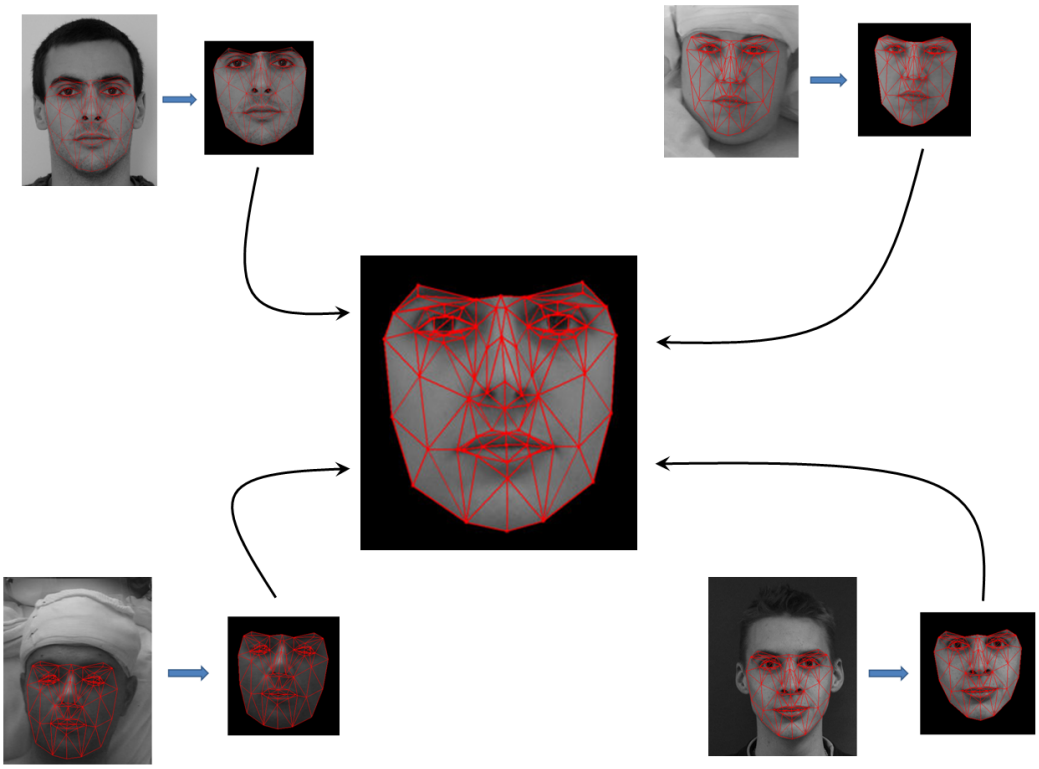


Figure 6.4: Reference texture, I_{ref}

6.2.3 The pros and the cons of our model

There is a lot of related work in the field of facial expression analysis. We have briefly presented the active appearance models and the morphable faces in the introduction (see section 6.1). These methods construct a model from a learning set of several images of different persons showing different expressions. A certain number of points of interest has to be placed in the images of the learning set and a Principal Component Analysis is applied. In our case, the expressions of the epileptic patients during their crises could be individual and complex and consequently a model built from a learning set of common expressions (typically such as anger, sadness or happiness) would not have been sufficient in order to represent the set of possible expressions. On the contrary the positions of our model are more free and can represent a larger set of expressions. On the other hand, by not using one of these statistical models, we lose some model capacity regarding to the texture: in our case we just have a simple average model for the texture. This drawback will sometimes give cause for some problems which we will address in the result section. In a word let us say that our priority was to have a model able to model large variations between faces even if it means that we sometimes will have to manually place some points of the model on the neutral image.

6.3 Model Fitting

6.3.1 Energy

For a new image of a face, we want to find the best position of the 3D model and the best values of the shape and animation parameters in the sense that the projection of the mesh $\mathcal{M}$ in the image matches the face (see Fig. 6.5). We need to define an energy which measures the quality of a given position of the 3D model (and its projection).

Let us first introduce some notations, we denote by Π the projection function:

$$\begin{array}{rccc} \Pi: & \mathbb{R}^3 & \rightarrow & \mathbb{R}^2 \\ & P & \mapsto & \Pi(P) \end{array}$$

with $\Pi(P) = \Pi\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \begin{pmatrix} \frac{X}{Z} \\ \frac{Y}{Z} \end{pmatrix}$ where: $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{M} \cdot \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix}$ and $\mathbf{M}$ is the 3×4 projection matrix. We chose a constant matrix and make the assumption that the difference between the chosen matrix and the actual matrix of the camera used to take the image will be compensated by the fact that we are

just interested in the 2D projection of the mesh and that we allow rotation, translation and scaling of the 3D mesh.

And lastly, an application $\varphi_{(P_i)}$ from the reference image and the actual one (see Fig. 6.6):

$$\begin{aligned} \varphi_{\mathbf{p}} : \quad & I_{text} \rightarrow I \\ & P \mapsto \varphi_{\mathbf{p}}(P) \end{aligned}$$

with, for $P \in I_{ref}$ and T_j such as P lies inside the triangle T_j^{ref}:

$$T_j^{ref} = (P_{a_j}^{ref}, P_{b_j}^{ref}, P_{c_j}^{ref}) = (A, B, C),$$

$$\varphi_{\mathbf{p}}(P) = \alpha(P)\Pi(P_{a_j}) + \beta(P)\Pi(P_{b_j}) + \gamma(P)\Pi(P_{c_j})$$

and

$$\begin{aligned} \alpha(P) &= \frac{P \wedge C - P \wedge B - B \wedge C}{BA \wedge AC} \\ \beta(P) &= \frac{P \wedge A - P \wedge C - C \wedge A}{BA \wedge AC} \\ \gamma(P) &= \frac{P \wedge B - P \wedge A - A \wedge B}{BA \wedge AC} \end{aligned}$$

Notice that $A \wedge B$ is the vector product between A and B and is taken as a simple scalar in our case, where A and B are 2D-vector. Although $\varphi_{\mathbf{p}}$ depends

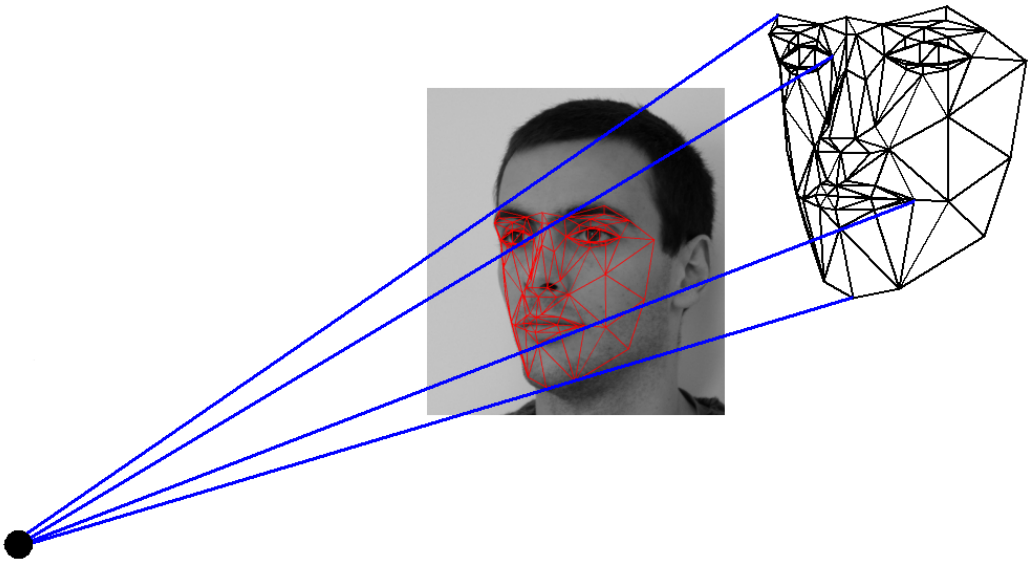


Figure 6.5: Projection of the 3D model in a 2D image, I

on the parameters $\mathbf{p}$ (and thus on the position of the (P_i)), we will sometimes denote it only by φ with a slight abuse of notation. Furthermore φ is only defined on $\bigcup T_j^{ref}$ and we supposed here that for each point $P \in I_{ref}$ there is no more than *one* triangle in which P lies ($T_j^{ref} \cap T_k^{ref} = \emptyset$ for any j, k). That is true if there were no *occlusions* in the meshes which were used to compute the reference texture. We will talk about occlusions in section 6.3.3.

Now, in order to measure the quality of a given position of the 3D model, we will use the cross-correlation between the *reference texture*, I_{ref}, and the image in which we want to fit the model, I, using the projection Π and the function φ. The support of the cross correlation is, for each vertex P_i, the set of triangles $\mathcal{V}_i$ incidents to this vertex (see Fig. 6.6):

$$\mathcal{V}_i = \left\{ T_j = (a_j, b_j, c_j) \;\middle|\; a_j = i \text{ or } b_j = i \text{ or } c_j = i \right\}$$

$$\rho_{\varphi,i} = \sum_{T_j \in \mathcal{V}_i} \int_{T_j^{ref}} \left(I_{ref}(P) - \overline{I_{ref}}_{,i} \right) \left(I(\varphi(P)) - \overline{I_{\varphi}}_{,i} \right) dP$$

with T_j^{ref} is the triangle of I_{ref} composed by $P_{a_j}^{ref}, P_{b_j}^{ref}$ and $P_{c_j}^{ref}$ and:

$$\overline{I_{ref}}_{,i} = \frac{1}{\sum |T_j^{ref}|} \sum_{T_j \in \mathcal{V}_i} \int_{T_j^{ref}} I_{ref}(P) dP$$

$$\overline{I_{\varphi}}_{,i} = \frac{1}{\sum |T_j^{ref}|} \sum_{T_j \in \mathcal{V}_i} \int_{T_j^{ref}} I(\varphi(P)) dP$$

Let us then define:

$$\langle I_{ref}, I \rangle_{\varphi} = \sum_{i=0}^{n_p} \rho_{\varphi,i}$$

And finally the energy:

$$\rho_{I,\varphi}(\mathcal{M}) = \frac{\langle I_{ref}, I \rangle_{\varphi}}{\sqrt{\langle I_{ref}, I_{ref} \rangle} \sqrt{\langle I, I \rangle_{\varphi}}} \tag{6.3}$$

This energy has the advantage of being invariant to any affine transformation of the histogram of the images.

6.3.2 Energy minimization

(a) Computation of the gradient of the energy

We now will compute the derivatives of the energy $\rho_{I,\varphi}(\mathcal{M})$ with respect to the 3D global position of the model, and to the shape and animation

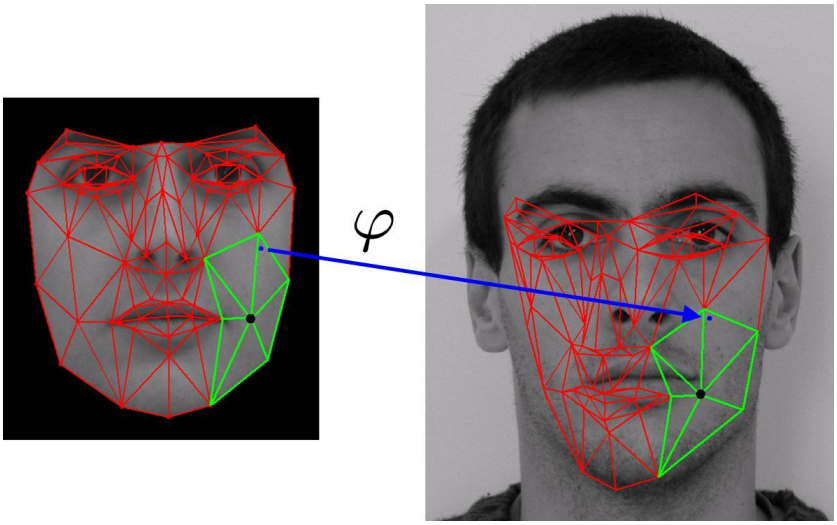


Figure 6.6: Support of the cross correlation and φ, the application between the reference texture and the current projection of the model in the image

parameters. Let start by computing the derivative of the projection function Π at point P.Which means that we want to find an application (or matrix) $\partial\Pi(P)$, such as, for any P and Q in $\mathbb{R}^3$, we have:

$$\Pi(P+Q) = \Pi(P) + \partial\Pi(P) \cdot Q + o(Q)$$

$$\mathbf{M} \cdot \begin{pmatrix} P+Q \\ 1 \end{pmatrix} = \mathbf{M} \cdot \begin{pmatrix} P \\ 1 \end{pmatrix} + \mathbf{M} \cdot \begin{pmatrix} Q \\ 0 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix}$$

$$\Pi(P+Q) = \begin{pmatrix} \frac{X+X'}{Z+Z'} \\ \frac{Y+Y'}{Z+Z'} \end{pmatrix} = \begin{pmatrix} X/Z - XZ'/Z^2 + X'/Z + o(Q) \\ Y/Z - YZ'/Z^2 + Y'/Z + o(Q) \end{pmatrix}$$

$$= \Pi(P) + \begin{pmatrix} 1/Z & 0 & -X/Z^2 \\ 0 & 1/Z & -Y/Z^2 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} + o(Q)$$

Therefore:

$$\Pi(P+Q) = \Pi(P) + \begin{pmatrix} 1/Z & 0 & -X/Z^2 \\ 0 & 1/Z & -Y/Z^2 \end{pmatrix} \mathbf{M_{3\times 3}} \cdot Q + o(Q)$$

with $\mathbf{M} = \begin{pmatrix} \mathbf{M_{3\times 3}} & \begin{matrix} V_x \\ V_y \\ V_z \end{matrix} \end{pmatrix}$ and $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{M} \cdot P$. And we finally have:

$$\partial\Pi(P) = \begin{pmatrix} 1/Z & 0 & -X/Z^2 \\ 0 & 1/Z & -Y/Z^2 \end{pmatrix} \mathbf{M_{3\times 3}} \tag{6.4}$$

($\Pi(P)$ is a 2×3 matrix).

Now we have to compute the derivative of φ with respect to the position of the three-dimensional point P_i, $\partial_i\varphi$. We have, for P such as for all $T_j \in \mathcal{V}_i$, $P \notin T_j^{ref}$, $\varphi_{P_i+\delta P_i}(P) = \varphi_{P_i}(P)$. And if there exists $T_j \in \mathcal{V}_i$ such as $P \in T_j^{ref}$, it means that, by definition of $\mathcal{V}_i$, $T_j = (a_j, b_j, c_j)$, with $a_j = i$ or $b_j = i$ or $c_j = i$. Let suppose $a_j = i$. We write:

$$\begin{aligned}
\varphi_{P_i+\delta P_i}(P) &= \alpha(P)\,\Pi\big(P_i + \delta P_i\big) + \beta(P)\,\Pi\big(P_{b_j}\big) + \gamma(P)\,\Pi\big(P_{c_j}\big) \\
&= \alpha(P)\,\big(\Pi(P_i) + \partial\Pi(P_i)\cdot\delta P_i\big) + \beta(P)\,\Pi(P_{b_j}) + \gamma(P)\,\Pi(P_{c_j}) \\
&= \varphi_{P_i}(P) + \alpha(P)\,\partial\Pi(P_i)\cdot\delta P_i
\end{aligned}$$

As a matter of fact, $\alpha(P)$, $\beta(P)$ and $\Gamma(P)$ does not depend on the position of the P_i since they are defined from P_i^{ref} which are fixed. And therefore:

$$\partial_i\varphi(P) = \begin{cases} \chi(P)\,\partial\Pi(P_i) & \text{if } \exists\ T_j \in \mathcal{V}_i \text{ such as } P \in T_j^{ref} \\ 0 & \text{otherwise.} \end{cases} \tag{6.5}$$

where:

$$\chi(P) = \frac{P \wedge P_{c_j} - P \wedge P_{b_j} - P_{b_j} \wedge P_{c_j}}{P_{b_j}P_{a_j} \wedge P_{a_j}P_{c_j}}$$

($\partial_i\varphi(P)$ is a 2×3 matrix).

We finally compute the derivative of the correlation energy with respect to the position of the three-dimensional point P_i. Let us write the expression of the derivative of the different term composing this energy:

- $\partial_i I(\varphi(P)) = \nabla I(\varphi(P))^{\mathbf{T}} \cdot \partial_i\varphi(P)$. $\partial_i I$ is a 3d-vector and is equal to zero if P does not lie inside a triangle which has P_i for one of its vertices.
- for $k \in \{1, \ldots, n_p\}$:

$$\partial_i \overline{I_{\varphi,k}} = \frac{1}{\sum_{j|T_j\in\mathcal{V}_k} |T_j^{ref}|} \sum_{j|T_j\in\mathcal{V}_k, T_j\in\mathcal{V}_i} \int_{T_j^{ref}} \partial_i I(\varphi(P))dP$$

  The second sum implies that:

  - if $k = i$, every triangles in $\mathcal{V}_i$ (for which P_i is a vertex) are considered,

- if P_k is a neighbor of P_i, only the triangles having P_i and P_k as vertices are considered,
- otherwise $\partial_i \overline{I_{\varphi,k}} = 0$.

- $$\partial_i \langle I, I \rangle_\varphi = \partial_i \Big[\sum_{k=0}^{n_p} \sum_{T_j \in \mathcal{V}_k} \int_{T_j^{ref}} \Big(I(\varphi(P)) - \overline{I_{\varphi,k}} \Big)^2 dP \Big]$$

$$= 2 \sum_{k=0}^{n_p} \sum_{j | T_j \in \mathcal{V}_k, T_j \in \mathcal{V}_i} \int_{T_j^{ref}} \Big(I(\varphi(P)) - \overline{I_{\varphi,k}} \Big) \Big(\partial_i I(\varphi(P)) - \partial_i \overline{I_{\varphi,k}} \Big) dP$$

$$= 2 \sum_{k=0}^{n_p} \sum_{j | T_j \in \mathcal{V}_k, T_j \in \mathcal{V}_i} \left(\int_{T_j^{ref}} I\big(\varphi(P)\big) \partial_i I(\varphi(P)) \; dP - \overline{I}_k \int_{T_j^{ref}} \partial_i I(\varphi(P)) \; dP \right)$$

in fact we have: $\displaystyle\int_{T_j^{ref}} \Big(I(\varphi(P)) - \overline{I_{\varphi,k}} \Big) \; dP \; = \; 0)$

- $$\partial_i \langle I_{ref}, I \rangle_\varphi = \partial_i \Big[\sum_{k=0}^{n_p} \sum_{T_j \in \mathcal{V}_k} \int_{T_j^{ref}} \Big(I_{ref}(P) - \overline{I_{ref}}_{,k} \Big) \Big(I(\varphi(P)) - \overline{I_{\varphi,k}} \Big) dP \Big]$$

$$= \sum_{k=0}^{n_p} \sum_{j | T_j \in \mathcal{V}_k, T_j \in \mathcal{V}_i} \int_{T_j^{ref}} \Big(I_{ref}(P) - \overline{I_{ref}}_{,k} \Big) \Big(\partial_i I(\varphi(P)) - \partial_i \overline{I_{\varphi,k}} \Big) dP$$

$$= \sum_{k=0}^{n_p} \sum_{j | T_j \in \mathcal{V}_k, T_j \in \mathcal{V}_i} \int_{T_j^{ref}} \Big(I_{ref}(P) - \overline{I_{ref}}_{,k} \Big) \partial_i I(\varphi(P)) dP$$

And finally, using all these terms, we can write:

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_i} = \rho_{I,\varphi}(\mathcal{M}) \Big(\frac{\partial \langle I_{text}, I \rangle_\varphi}{\langle I_{text}, I \rangle_\varphi} - \frac{1}{2} \frac{\partial \langle I, I \rangle_\varphi}{\langle I, I \rangle_\varphi} \Big)$$

We now just have to write the derivatives of the energy with respect to each of the evolving parameters in $\mathbf{p}$. Let us start by the derivative with respect to the translation vector $\mathbf{t} = (t_x, t_y, t_z)$:

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial \mathbf{t}} = \sum_{j=1}^{n_p} \frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j}$$

For $i \in \{1, \ldots, n_{su}\}$, the derivative with respect to the i-th shape parameter is:

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial \sigma_i} = \sum_{j=1}^{n_p} \Big(\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j} \cdot s\mathbf{R}\mathbf{v}_{i,j}^{s} \Big)$$

For $i \in \{1, \ldots, n_{au}\}$, the derivative with respect to the i-th animation parameter is:

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial \alpha_i} = \sum_{j=1}^{n_p} \Big(\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j} \cdot s\mathbf{R}\mathbf{v}_{i,j}^{a} \Big)$$

The derivative with respect to s, the scale parameter, is:

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial s} = \sum_{j=1}^{n_p} \Big(\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j} \cdot \mathbf{R}(\overline{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha) \Big)$$

And finally the derivatives with respect to $\theta_x, \theta_y, \theta_z$ the rotation parameters are:

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial \theta_x} = \sum_{j=1}^{n_p} \Big(\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j} \cdot s\frac{\partial \mathbf{R}}{\partial \theta_x}(\overline{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha) \Big)$$

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial \theta_y} = \sum_{j=1}^{n_p} \Big(\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j} \cdot s\frac{\partial \mathbf{R}}{\partial \theta_y}(\overline{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha) \Big)$$

$$\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial \theta_z} = \sum_{j=1}^{n_p} \Big(\frac{\partial \rho_{I,\varphi}(\mathcal{M})}{\partial P_j} \cdot s\frac{\partial \mathbf{R}}{\partial \theta_z}(\overline{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha) \Big)$$

(b) The optimization procedure

We then perform a multi-scale gradient descent and we use a quasi-Newton method with non-linear constraints [84]. The multi-scale method means producing reduced versions of the image and applying the minimization process to the smallest one and then using the result as the initial state for the minimization on the next image. It makes the process less dependent on the first initialization of the minimization and quicken the convergence. Furthermore we imposed several constraints in the minimization algorithm:

- arbitrary boundaries on the shape and animation parameters to keep the mesh in the set of "reasonable" faces and "reasonable" facial expressions (e.g. we do not want the mouth to be "inside" the nose),
- some specific conditions on the mesh: the eyelids and the lips must not pass through each other,

- and the fact that the projection of the mesh should not exceed the borders of the image.

6.3.3 Occlusions

As mentioned earlier, in our formulas we supposed that when projecting the mesh in an image, there were no overlaps, meaning no occlusions. To a first approximation, the mesh representing the face can be considered as a 2D-plane (i.e. with no occlusions). But occlusions can occur in two cases: if the face is rotated (in profile or almost) and if animation units make a part of the mesh to occlude an other part (principally true for the eyelid when eye closure). Therefore we cope with this problem thanks to two modifications: first we make sure that the reference texture is defined on the whole mesh (by averaging images with no occlusions) and secondly we compute for each position of the mesh $\mathcal{M}$ an *occlusion function* $\mathcal{O}()$ which for a point P lying on $\mathcal{M}$ says if this point is "visible" in the image I or "occluded", and in every integral composing the energy or the gradient expression we do not take into account "occluded" points . Unfortunately this function is not differentiable and that is why we did not include this function in the energy expression. Although our mathematical formulas are not exact anymore, we did not encounter any related problems in our numerical experiments.

6.4 Facial Expression

Once the model is fitted in an initial image of a subject, we define a new *reference texture* using the image and the projection of the fitted model. If there were occlusions in the position of the mesh we chose to use the previous reference texture (obtained by average) for the occluded parts. Given a new image of the same subject with a new facial expression, we use this new *reference texture* in order to fit the model in this image. The minimization is now done only with respect to global position and animation parameters (the shape parameters are supposed to be constant for two images of the same person). The facial expression can then be represented as the variation of the animation parameters between the neutral image of the subject and the expressive one. Our first idea was to used this variation of the animation parameters to compare different facial expressions by different persons. As a matter of fact taking the *variation* of the parameters and not directly the *value* makes the result independent on the shape of the person face. For example if the person has naturally (meaning even in "neutral state") his eyebrows in higher (further from the pupil) than an other person and these

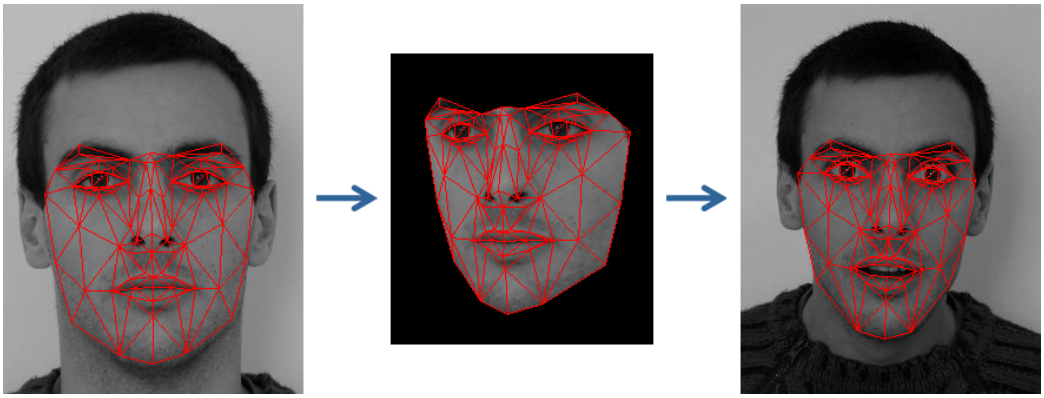


Figure 6.7: the two steps of the analysis of an expression. Left: the model is fitted on a neutral image. Middle: the new reference texture. Right: the model is fitted on the expressive image.

two persons raise their eyebrows (see Fig. 6.8), we want to be able to compare the eyebrows raise without taking into account their initial position. That is the purpose of considering the variation. We can then for example apply these variations to an average face and compare these two facial expressions, see Fig. 6.9.

6.5 Results

During seizures, large movements can occur and occlusion of the image may be produced by intervention of medical personnel. Therefore we manually selected several images for each patient: one of a neutral expression and some during seizures. Fig. 6.10 and 6.11 show some of the results of the algorithm: the selected images and the final fit of the 3D model. We will talk in the next chapter about what can be done with video on which the model can fitted during the all seizures.

In some unusual cases (such as thick beard or dark glasses), we had to help the minimization procedure for the first image (the neutral one). It is made easier by the fact that the points of reference texture are two dimensional: we can do a first minimization, take the result as the new reference texture and then modify the position of a few number of misplaced points and finally apply a new minimization procedure. As explained earlier we then just have to take the result of the fitting process as the new reference texture and to

[3] http://cvc.yale.edu/projects/yalefaces/yalefaces.html

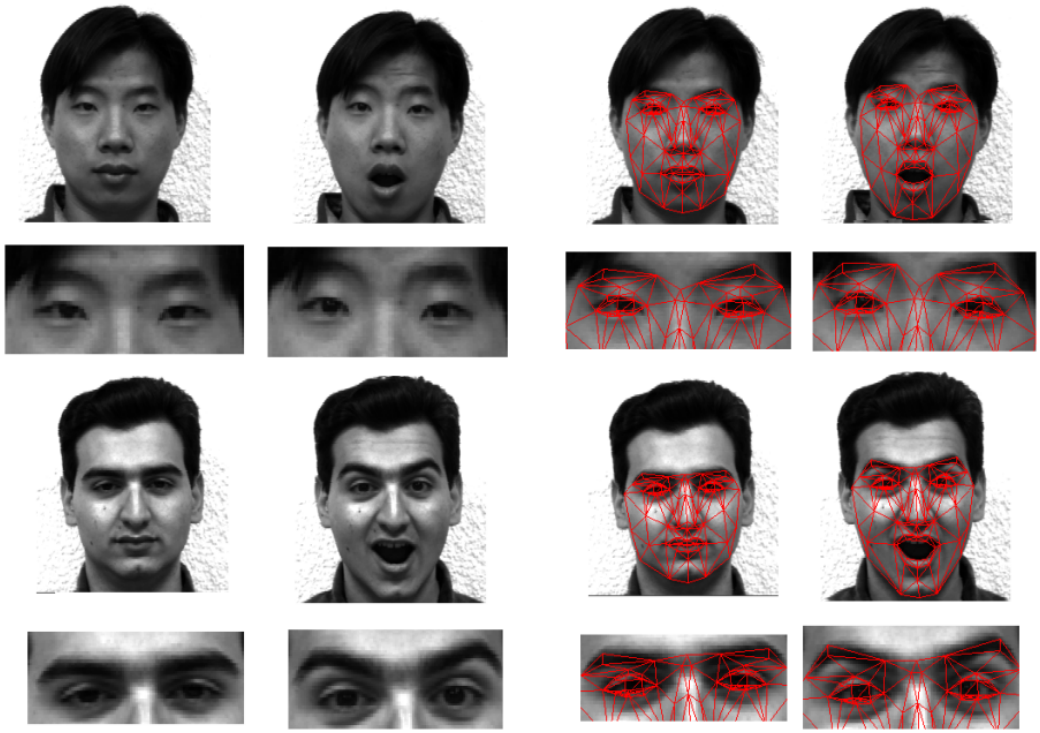


Figure 6.8: Two persons raising their eyebrows and the result of the model fitting (images taken from the Yale database [3])

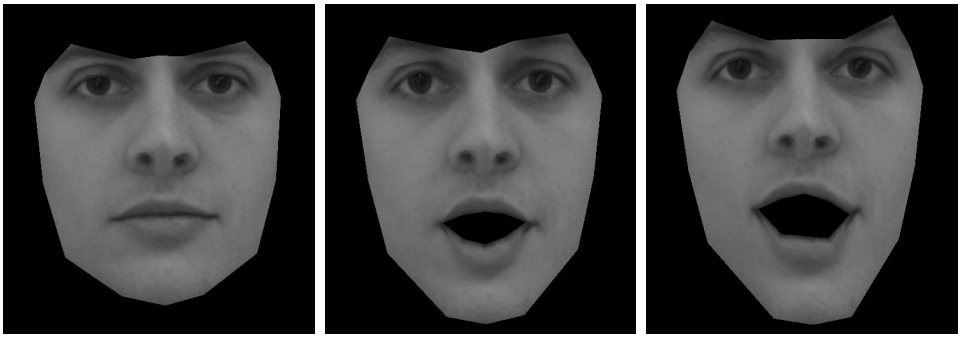


Figure 6.9: The two facial expressions extracted from Fig. 6.8 applied to an average face. From left to right: the average face in neutral expression, the first expression from Fig. 6.8, the second expression from Fig. 6.8

fit the model in the expressive image.

Once the model is fitted on the neutral image and on the expressive one,

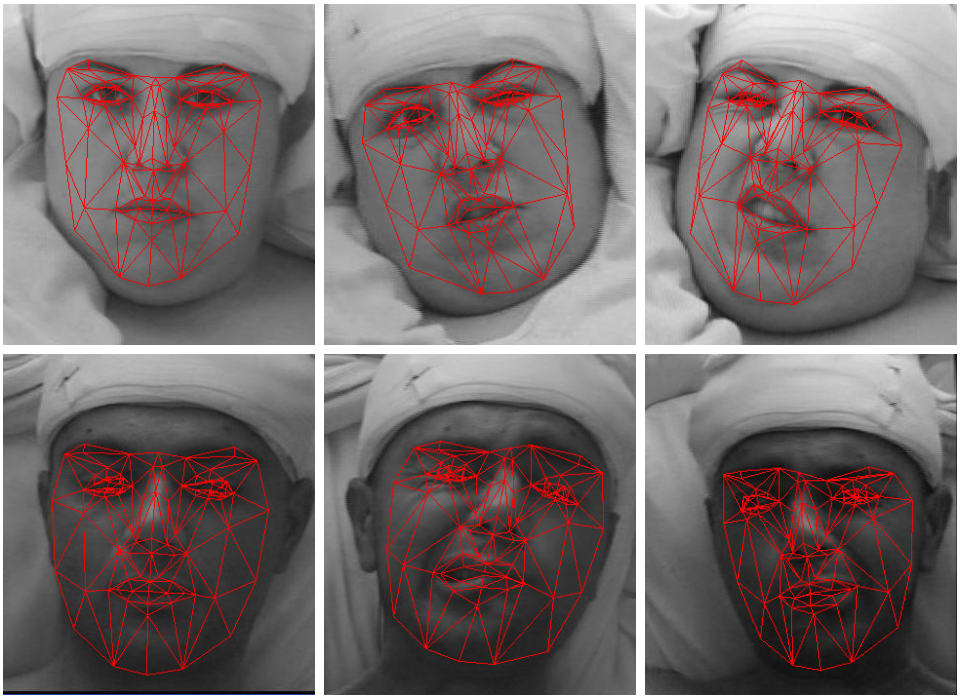


Figure 6.10: Some results: left column shows a neutral view and other columns are images taken during the seizure

we can compute the variation of the animation parameters from the first image to the second one and for example apply that variation to another position of the model (Fig. 6.12). The process of fitting the model in a new image took in the order of a few seconds.

Our idea was to apply this process in a sufficient number of seizures, to extract one or several facial expressions (as the variation of the animation parameters) by seizure, in order to be able to do some statistics or classification (like in section 1.5). Unfortunately the data set was not large enough and that is why we focused our attention on the treatment of several videos and the correlation between the facial parameters with the electrical activity produced by the brain, see next chapter.

6.6 Conclusion

We have proposed a method to fit a 3D animated model in a monocular single image under uncontrolled imaging conditions. Our method is based on an energy defined with a cross-correlation term and an energy minimization

process. We fitted the model on real-world data obtained in a medical framework. The principal drawback of our approach is the time of computation of the optimization process which is quite far from real-time. However real-time processing was not one of our requirements, since our work was motivated by a need for an analysis tool. Further work includes investigating on the potential clinical use of this tool and on a more sophisticated model (such as including more advanced texture statistics than just an average).

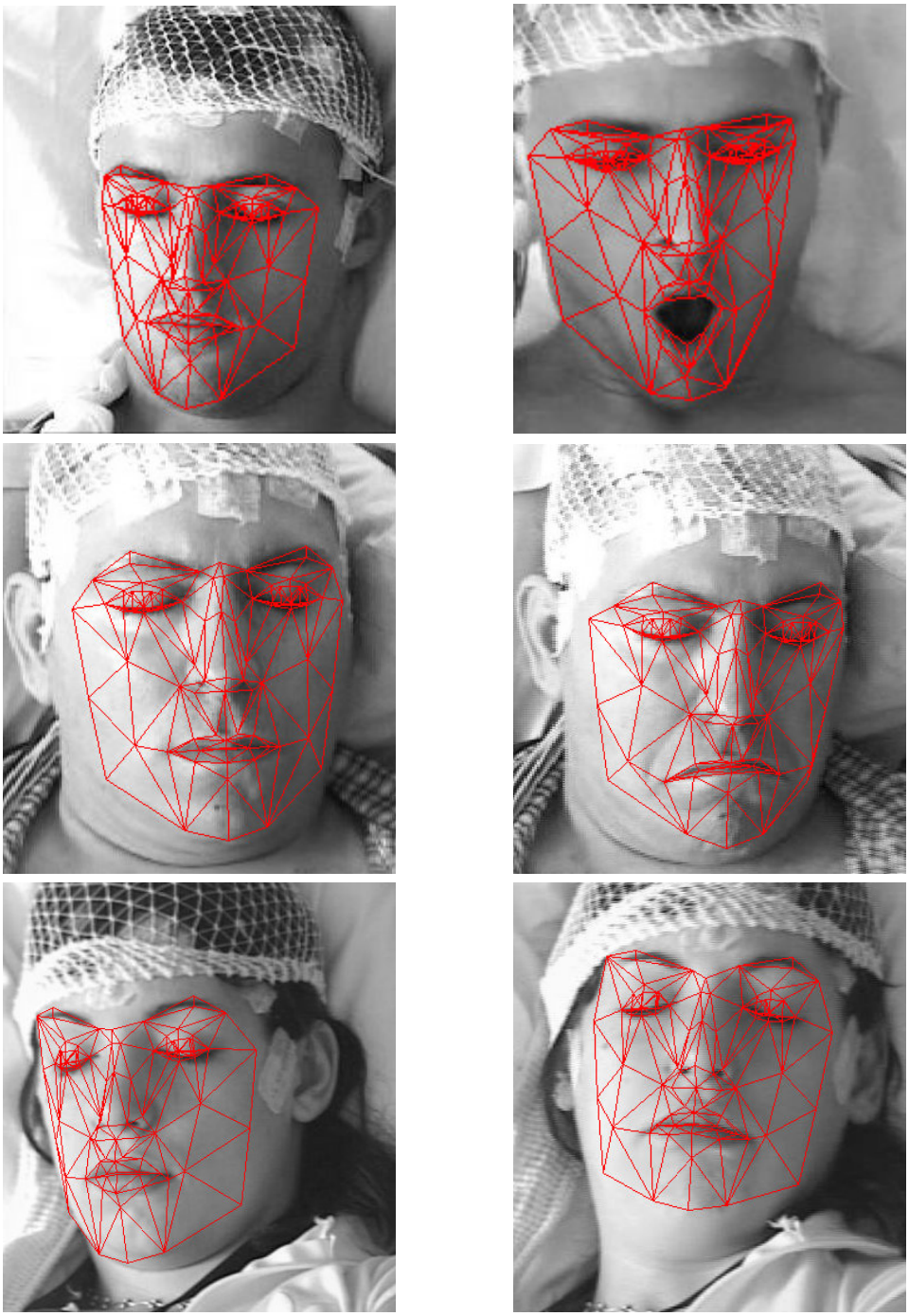


Figure 6.11: Some results: left column shows a neutral view and right column images taken during the seizure

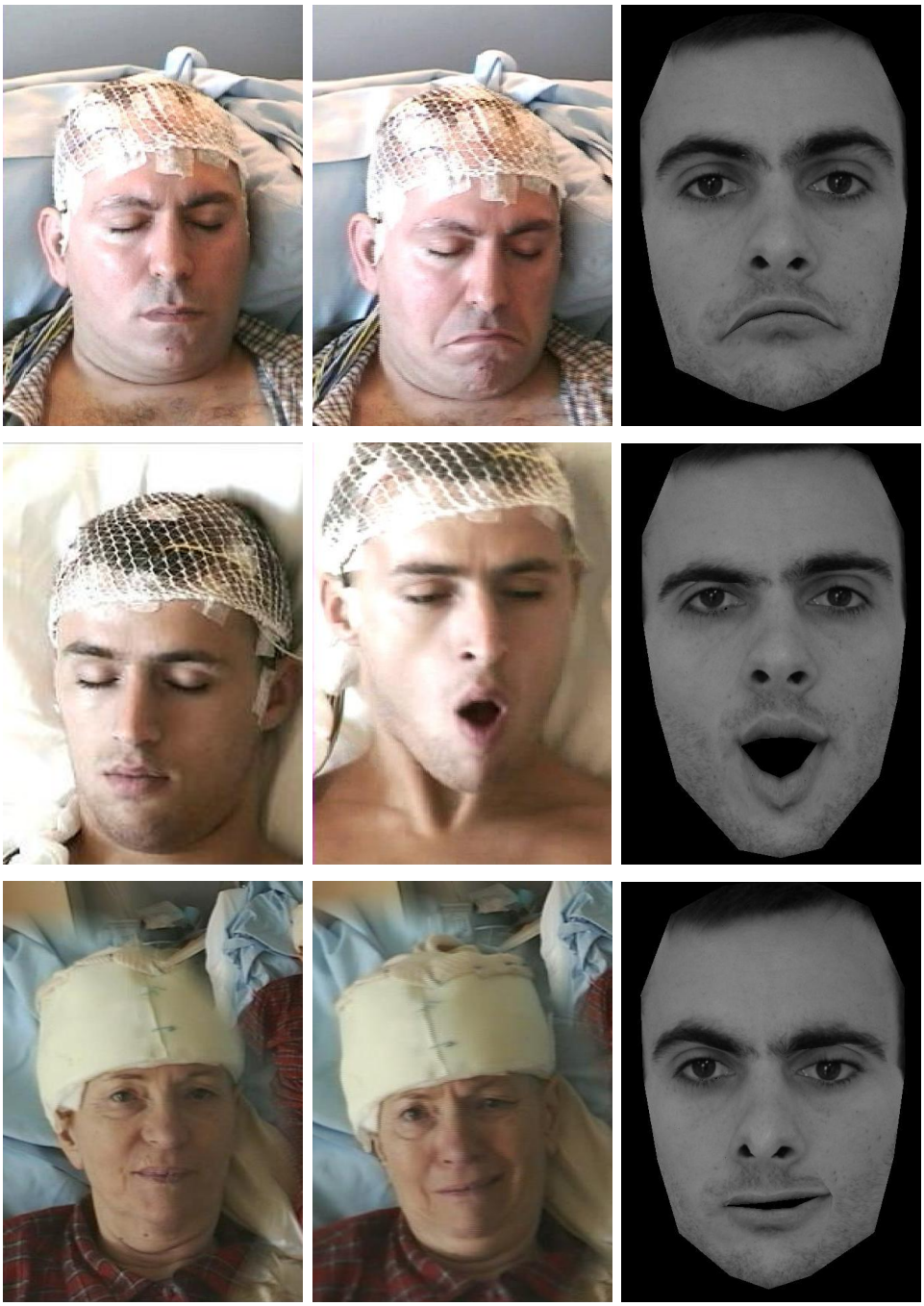


Figure 6.12: first column: a neutral expression, second column: image of the facial expression during an epileptic seizure and last column: the expression mapped on a new image

Chapter 7

Electro-clinical correlation

During pre-surgical exploration, patients are videotaped at the same time as their EEG or SEEG (see chapter 5, section 5.2.1), is recorded. The anatomo-electro-clinical correlation is investigated [35, 129, 11, 122]: the "anatomic" information comes from the knowledge of the electrodes localization on or in the brain , the "electrical" data are given by the EEG or SEEG recordings and the "clinical" information is the object of semiology and obtained by interrogating the patient and observing the video. This close observation of clinical features (including the order in which they occurred) and their relation to the region primarily or secondarily involved in the epileptic discharge, allows a temporal-spatial profile of the seizure's origin and propagation patterns to be established and the corresponding anatomical structures to be defined, thus aiding the decision regarding surgical decision for each patient's particular epilepsy. Our work is an answer to the need of *quantification* of the electro-clinical correlation.

In this chapter we present our work on the treatment of dynamic facial expression in video and on the search of a relation between the facial expression and the electrical activity in the brain during the epileptic seizures.

7.1 Introduction

Facial expressions have been used in clinical research to evaluate deficits in emotional expression and social cognition in neuropsychiatric disorders [102, 92]. It is particularly present in schizophrenia studies. Schizophrenia is a mental disorder characterized by abnormalities in the perception or expression of reality which often imply difficulties in interpreting and understanding facial expressions and emotions. However, most of the current clinical methods are based on subjective ratings and therefore provide qual-

itative measurements. The Facial Action Coding System (FACS) [48] is a widely used method to describe facial expressions using a combination of action units. Each action unit corresponds to a specific muscular activity that produces changes in facial appearance. It usually requires intensive human interaction. Thanks to recent advances in automatic facial expression analysis [70], Peng Wang and his colleagues recently presented an automated video-based facial expression analysis of neuropsychiatric disorders [141, 140]: they used videos of persons which have been asked to mimic "classic" expressions (happiness, sadness, anger and fear) to automatically detect schizophrenia disorder. Nevertheless their data capturing system is quite restrictive: they used six grayscale stereo cameras, one color camera and a video camera and the patients are supposed not to move too much and stay in front of the recording system. In our case, we almost do not have any control on the recording system: one camera is fixed to a top corner of the room and is just recording what is happening in the hospital room. Furthermore, epileptic seizures are often unforeseen and then the conditions in which it happens are uncontrollable.

7.2 Dynamic facial expression

The automatic interpretation of video have been a productive field in the past decade. It has principally been used for video-surveillance [139, 19] consisting in recognizing pre-defined scenarios describing human behaviors from video sequence. But some applications to medical imaging have also been done, see for example [33] in which authors developed a knowledge-based program supervision system for automating the management of medical imaging processing libraries.

In our case we are just interested in the evolution of the facial expression present in the videos. Therefore we first start by selecting videos of epileptic seizures in which the facial expressions were visible during a large part of the seizure. As a matter of fact, in a number of videos just a few frames were showing a workable facial image because of the uncontrolled aspect of the video recording: quality of the images was sometimes not good enough (if the patient was moving too fast or if there was not enough light), the face was sometimes not visible (if an object, e.g. a book, or a person occluded the face).

We then extracted frames from these videos and fitted the face model in these images. Since we did not seek for a real-time video processing, we just applied the fitting procedure presented in the previous chapter on each frame, using the result of the minimization on a frame as the initial position

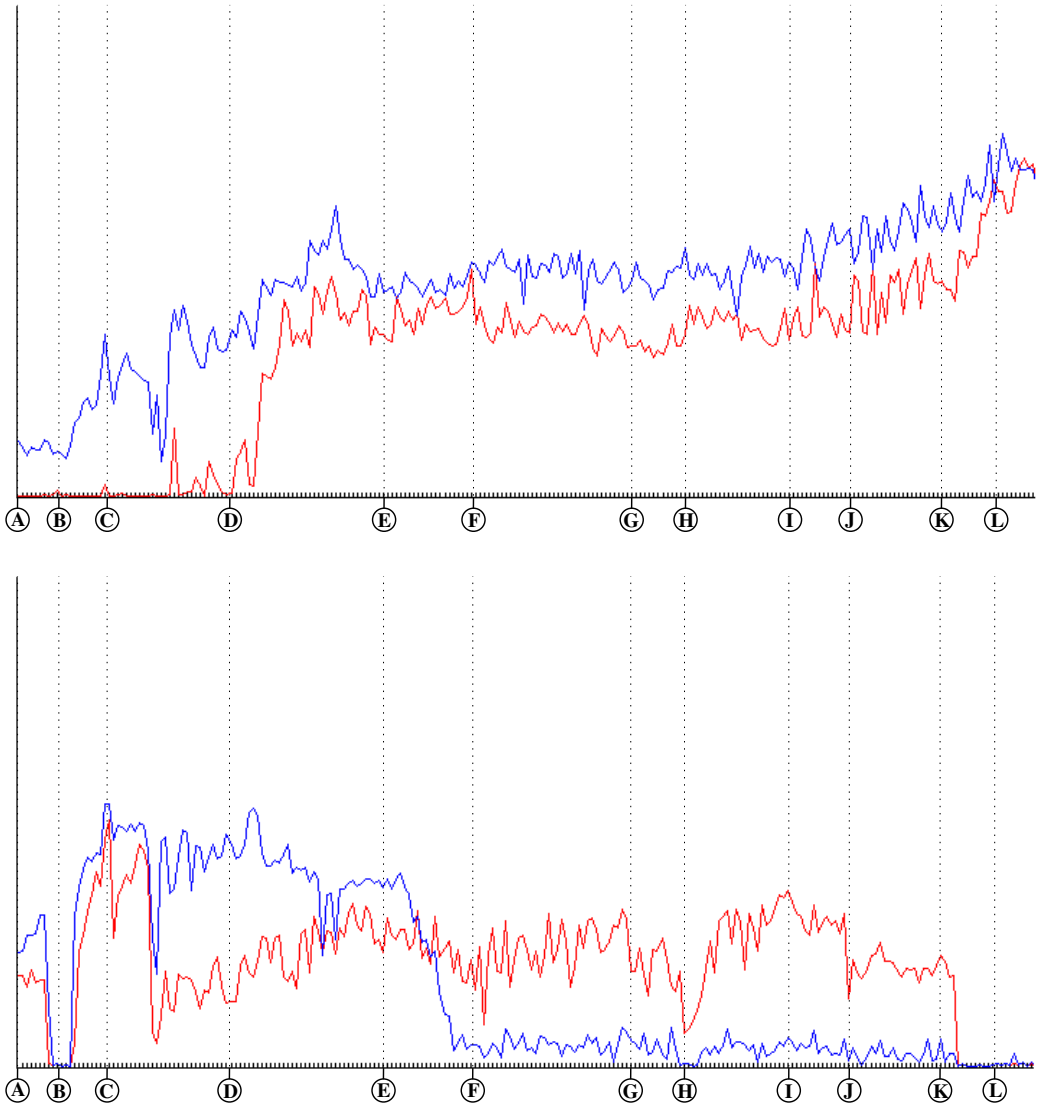


Figure 7.1: Evolution of facial parameters during an epileptic seizure. The time is on the x-axis. Top row: the red curve is the *left* opening of the mouth and the blue one is the *right* opening of the mouth. Bottom row: the red curve is the opening of the *left* eye and the blue one the opening of the *right* eye. Letters mark some instants showed by Fig. 7.2

of the model for the following one and defining our reference texture on the first frame. In fact this method is quite far from real-time since it requires few seconds by frame (and the frame rate was 25 frames by second) but once again since our work was motivated by a need for an analysis tool,

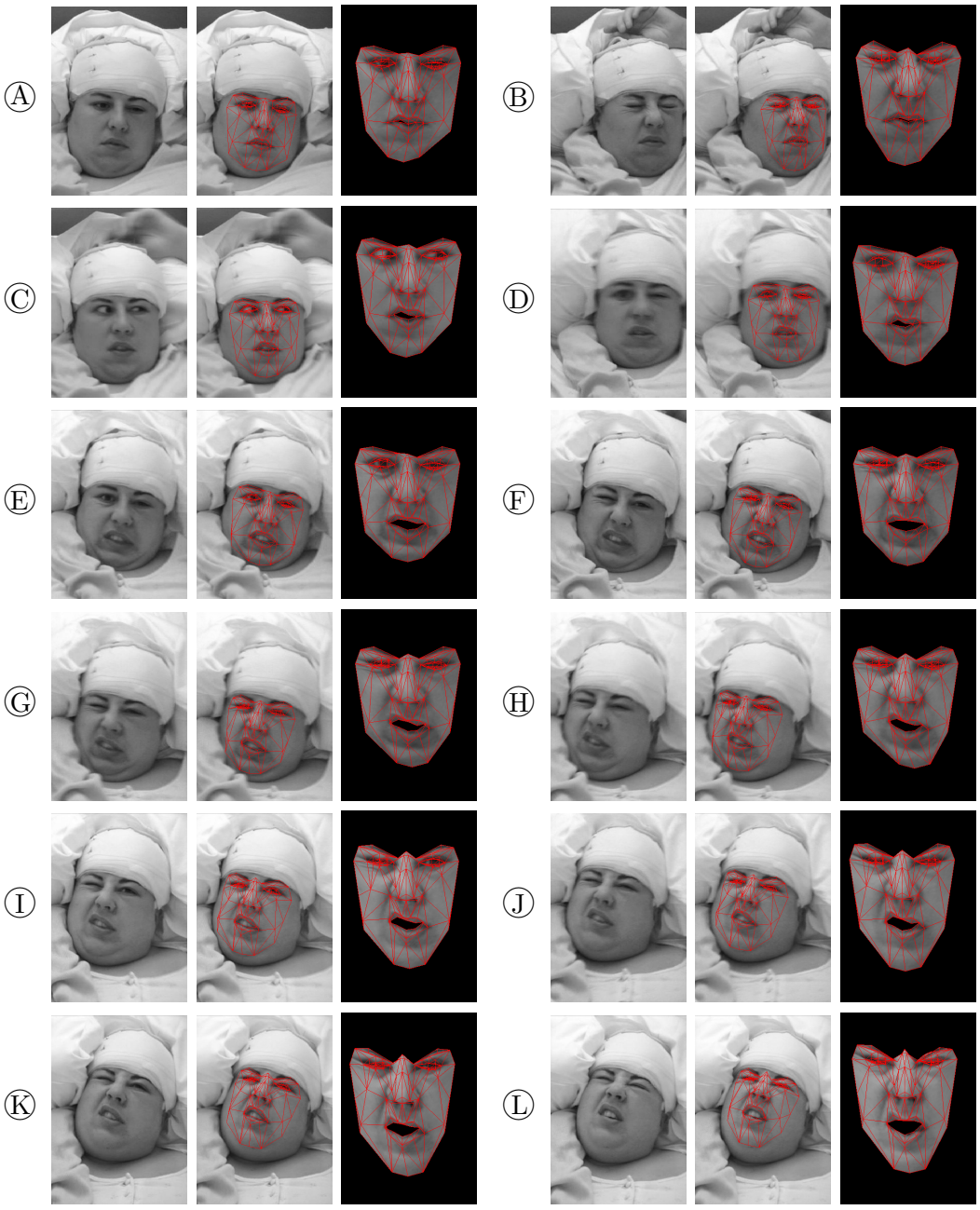


Figure 7.2: Images corresponding to the letters written on Fig. 7.1. For each column, from left to right: the image frame, the result of the minimization, the frontal view of the model

real-time processing was not one of our requirements and we gave priority to robustness.

Once the model is fitted in every frames we can get the positions of the points of the mesh without any global transformation. That is to say, we apply the shape σ and animation α parameters obtained by optimization to the initial points $\overline{\mathbf{g}}$ but not the global transformations $\mathbf{R}$, s and $\mathbf{t}$ (see the previous chapter, section 6.2.1 (b)). It also can be seen as extracting a frontal view from the mesh position in the frame. And on this frontal view we can evaluate any facial quantities we are interested in, for example eye position, the opening of the mouth , mouth position. Therefore we can compute the evolution of the features in order to characterize the evolution of the dynamic facial expression. Fig. 7.1 and Fig. 7.2 show an example of such a treatment. For example we can detect events in the dynamic facial expression, such as : eye blinking (e.g., both eyes: letters A, H and L; only one eye: letter I), asymmetry in the mouth opening (e.g., letter D, the right side of the mouth is more opened than the left side).

7.3 Stereoelectroencephalography signal processing

As presented in chapter 5, section 5.2.1 (c), Stereoelectroencephalography (SEEG) is an investigation technique consisting in recording electrical activities in different brain regions thanks to depth electrodes. Generally, the number of electrodes is less than 10 and on each electrodes about ten sensors (called leads) are positioned, each lead has a diameter of 0.8 mm, a contact length of 2 mm and leads are placed 1.5 mm apart. Usually the signal which is studied is a *bipolar recording*, it means that we consider the difference between one given lead and the consecutive one on the same electrodes. Then a first pre-treatment is often apply in order to remove some of the possible noise: we used the `eegfilt` function in the EEGLAB matlab toolbox. It is a frequency filter that we applied to remove too low (< 1 Hz) and too high (> 120 Hz) frequencies in the signal. Fig. 7.3 shows examples of such bipolar recording during an epileptic seizure.

One of the classical ways to study such SEEG signals is to use a *time-frequency* analysis [144, 81, 143, 123, 8]. While the technique of Fourier transform is useful to obtain the frequency spectrum of a signal whose component frequencies are constant in time, it is not really relevant for analyzing a signal whose component frequencies vary over time which is the case here. Therefore we chose to apply an *analytical wavelet transform* to each bipolar

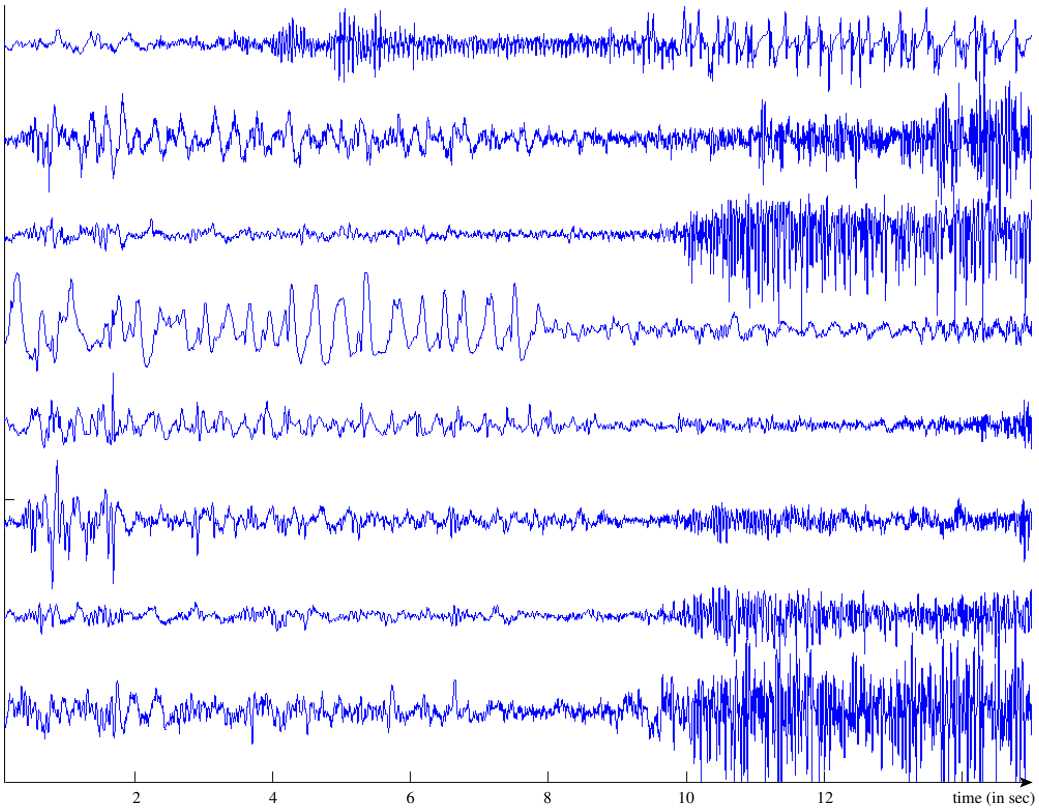


Figure 7.3: Examples of bipolar recordings during an epileptic seizure. The seizure onset is at 2.5 sec.

signal, it give us for each instant and for a given frequency the intensity of this frequency component in the signal at this instant. Fig 7.4 shows a representation of this time-frequency analysis for a SEEG bipolar signal. This analysis has been done using a version of a Matlab function from the WaveLab toolbox (J. Buckheit and D. Donoho [43]) modified by C. Benar and M. Clerc.

This algorithm starts with a Fast Fourier Transform (FFT) on our signal. Then for each frequency we want to study, we construct a wavelet (we chose a Gabor wavelet here). This wavelet is then used to extract the contribution of the considered frequency to our original signal from the result of the FFT. This is done by pointwise multiplying the gabor filter and the FFT result. The result of this operation is the contribution of the considered frequency to the SEEG signal in the frequency domain. Finally an inverse FFT of this result comes back to the temporal domain and give the contribution of the considered frequency to our signal along the time.

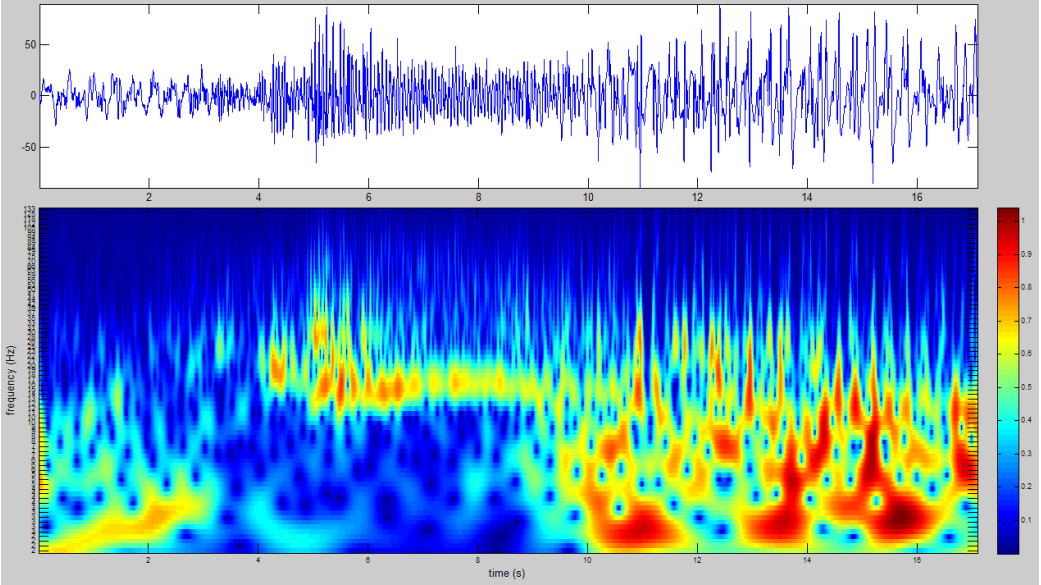


Figure 7.4: Representation of the analytical wavelet transform of a SEEG bipolar signal. Top row: the signal. Bottom row: the time-frequency image, on the x-axis is the time and on the y-axis is the discretized frequency domain $\mathcal{F}$

7.4 Correlation between facial parameters and SEEG signal

Our treatment of the video can be seen as a quantification of some facial features characterizing the evolution of the facial expression. Once these characteristics are quantified, we now have the possibility of designing a way to measure the possible relation between the dynamic facial expression and the electrical recording in the brain during the epileptic seizure.

We will use the time-frequency analysis on the SEEG bipolar signals (see previous section). The result of the time-frequency transform is a function that we note $\omega : I \times \mathcal{F} \to \mathbb{C}$ where $I \subset \mathbb{R}^+$ is the time domain and $\mathcal{F}$ the frequency domain given by this analysis. Typically we have $\mathcal{F} \subset [1 \text{ Hz}, 120 \text{ Hz}]$ since we have applied a band-pass filter to remove too low and too high frequencies. In Fig. 7.4, the image in bottom row is a representation of the complex modulus (magnitude) of $\omega(t, f)$ for $t \in I$ and $f \in \mathcal{F}$.

Let us then define for a given frequency $f \in \mathcal{F}$ the function ψ_f representing the variation along time of the intensity of the signal component of

frequency f:

$$\begin{aligned} \psi_f : \quad & I \to \mathbb{R}^+ \\ & t \mapsto \sum_{g \in \mathcal{F}} \sigma(g-f)\ |\omega(g,t)| \end{aligned}$$

where $\sigma()$ is a one-dimensional centered Gaussian function of variance a:

$$\sigma(x) = \frac{1}{a\sqrt{2\pi}} e^{\frac{-x^2}{2a^2}}$$

Fig. 7.5 shows ψ_f for $f = 5$ and $a = 2$ computed on the signal represented on Fig 7.4.

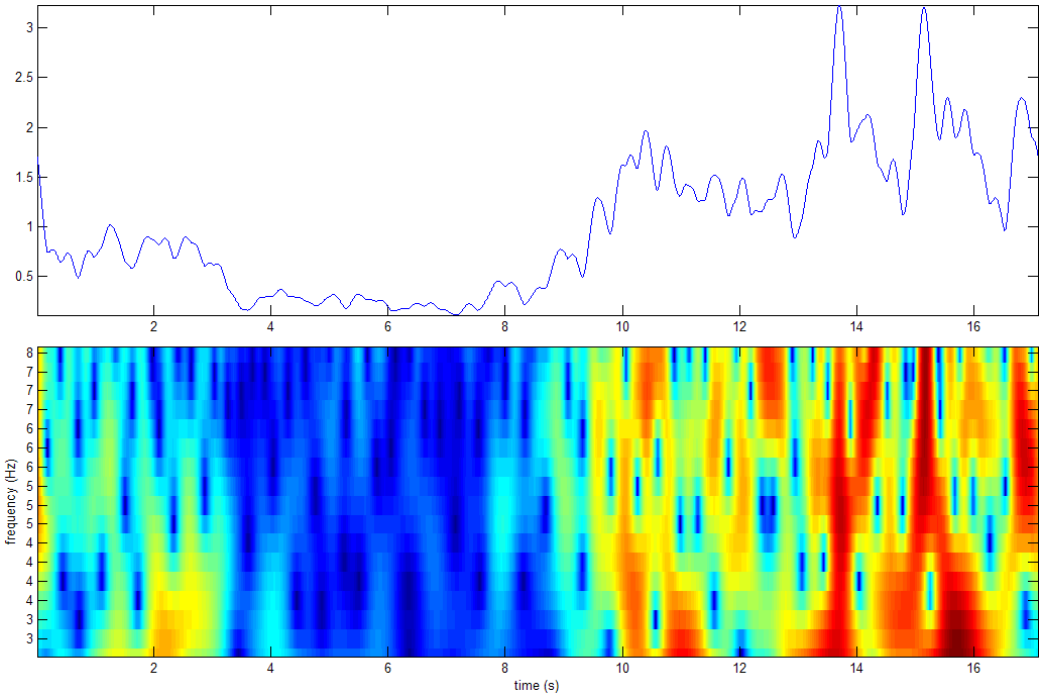


Figure 7.5: Representation of ψ_f (top row) for $f = 5$ and $a = 2$. On bottom row representation of the time-frequency image for a restrained portion of $\mathcal{F}$ centered on $f = 5$ Hz

And finally, for a frequency $f \in \mathcal{F}$ and a facial feature $p : t \in I \mapsto p(t) \in \mathbb{R}$, we can define the correlation between these two real signals by:

$$\rho(f,p) = \frac{\sum_{t \in I} (\psi_f(t) - \overline{\psi_f})(p(t) - \overline{p})}{\sqrt{\sum_{t \in I} (\psi_f(t) - \overline{\psi_f})^2 \sum_{t \in I} (p(t) - \overline{p})^2}}$$

where:

$$\overline{\psi_f} = \frac{1}{|I|}\sum_{t\in I}\psi_f(t) \quad \text{and} \quad \overline{p} = \frac{1}{|I|}\sum_{t\in I}p(t)$$

are the mean of the functions ψ_f and p on the time domain I. This correlation function $\rho(f,p)$ has its values in $[-1,1]$ and is equal to 1 or -1 if one signal is an affine transformation of the other signal (1 if it is an increasing affine transformation and -1 if it is a decreasing one). The signals are told to be *strongly correlated* if the correlation between them is close to the extreme values -1 or 1 and *linearly independents* if the correlation is close to 0.

7.5 Application to a first case

Now that we have this measure of similarity between two signals, we can search for any frequency in $\mathcal{F}$, which one will best correlate with a given facial parameter p, and finally compare for different bipolar signals this value of correlation. This will give us our *electro-clinical correlation* concerning the chosen facial parameter p. We applied this framework on a first "simple" case in order to validate our method.

The epileptic seizure we chose is clinically characterized by a deviation of the eyes and afterward a deviation of the head. We first applied our dynamic facial expression analysis method on the video and chose to focus on the position of the eyes and the rotation of the head. We extracted 435 frames (around 17 seconds, at 25 frames by second) and fitted the model in every frame and finally computed our facial quantities. Results are shown on Fig. 7.6 and Fig. 7.7.

SEEG-electrodes have been implanted in the brain of this same patient. The localization of these electrodes have been done thanks to MRI scans of the brain, see Fig. 7.8. 6 electrodes were considered and a total of 46 channels have been recorded. We focused at the *bipolar signals* (obtained from subtraction of signals recorded on two adjacent leads) as explained earlier which gave us 40 signals ($46-6$ number of channels minus number of electrodes). The sampling frequency of the recordings was 512 Hz.

Fig. 7.9 and Fig. 7.10 show the result of our correlation computation: for each bipolar signal, we compute the correlation between one facial parameter and this signal for every frequency in the frequency domain $\mathcal{F}$ (with a variance for the gaussian function of $a = 5$, but let us note that according to our experiments the choice of the variance does not change significantly the results) and keep the maximum correlation. Fig. 7.11 and Fig. 7.12 indicate for each bipolar signal which was the frequency implying the maximum correlation. We focused only on the parameters concerning the eyes (opening

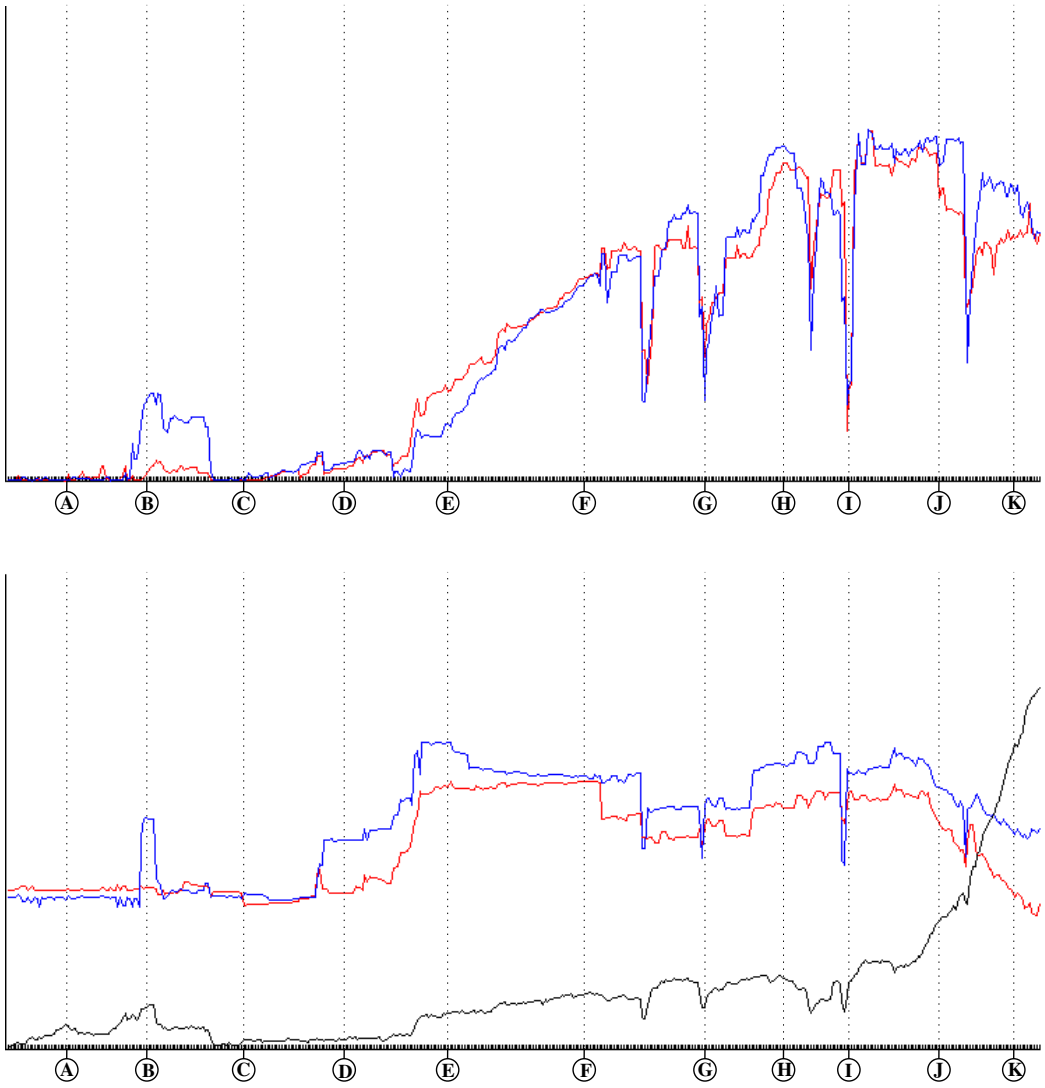


Figure 7.6: Evolution of facial parameters during the studied epileptic seizure. The time is on the x-axis. Top row: the red curve is the opening of the *left* eye and the blue one is the opening of the *right* eye. Bottom row: the red curve is the horizontal deviation of the *left* eye and the blue one the deviation of the *right* eye, the black curve is the head rotation. Letters mark some instants showed by Fig. 7.7

and position), because unfortunately the parameter concerning the rotation of the head did not produced interesting correlations. This might be because this parameter was not varying enough to contain enough characteristic in-

formation to differentiate correlations between different electrodes.

The signal in which we found a frequency component who best correlates with the eye facial parameters (opening and position) is thus located on the electrode called "SA" (bipolar signals number 22 or 23). What is very interesting is that this electrodes was pointed out *before our processing* by the medical team as the one with the most worthwhile potential regarding the information of eye movement and position. This electrode was indeed exploring the premotor cortex which is part of the frontal cortex of the brain. And a part of this region is an area of the brain called the *frontal eye field* (FEF), known to be activated during eye movement. The reader interested in the approximative location of FEF will look back at the position of the electrode SA in Fig. 7.8.

Although we are aware that this first case is not sufficient to draw definitive conclusions about the validity of our method, this result is a preliminary validation and indicates that our facial parameters contain informations that can be correlated with SEEG signals through a time-frequency decomposition of these signals through the presented method. More work still remains to be done and in particular investigating the significance of our correlation. For instance in our first study, the electrode "C" presented also a good correlation with our facial parameters. This electrode was implanted in the posterior temporal area of the brain which is not known as a primary source for eyes movements. Nevertheless the activation of this region seems to be able the provoke such movements under specific circumstances. This could be the subject of further investigations.

7.6 Conclusion

We defined a procedure for measuring the potential interrelation existing between the facial expression that may occur during an epileptic seizure and the region in the brain at the origin of the seizure. This first step in the direction of applying techniques coming from domain of computer vision to a problem at the cutting edge of epilepsy research domain calls for further investigations on this question of the above-mentioned interrelation, including improving our 3D face model (e.g. building a more sophisticated statistical model of the facial texture) for the computer vision side and examining more seizures and their SEEG-recordings, for the medical side. Furthermore, the automation of the image selection step (to keep only frames with non-occluded faces) would enable the treatment of a larger number of videos and facilitate a statistical study of these facial expressions. One can also imagine an extension of this method which not only considers the evolution of the facial expressions

during the seizure but also takes into account the deformation of body parts such as hands or legs.

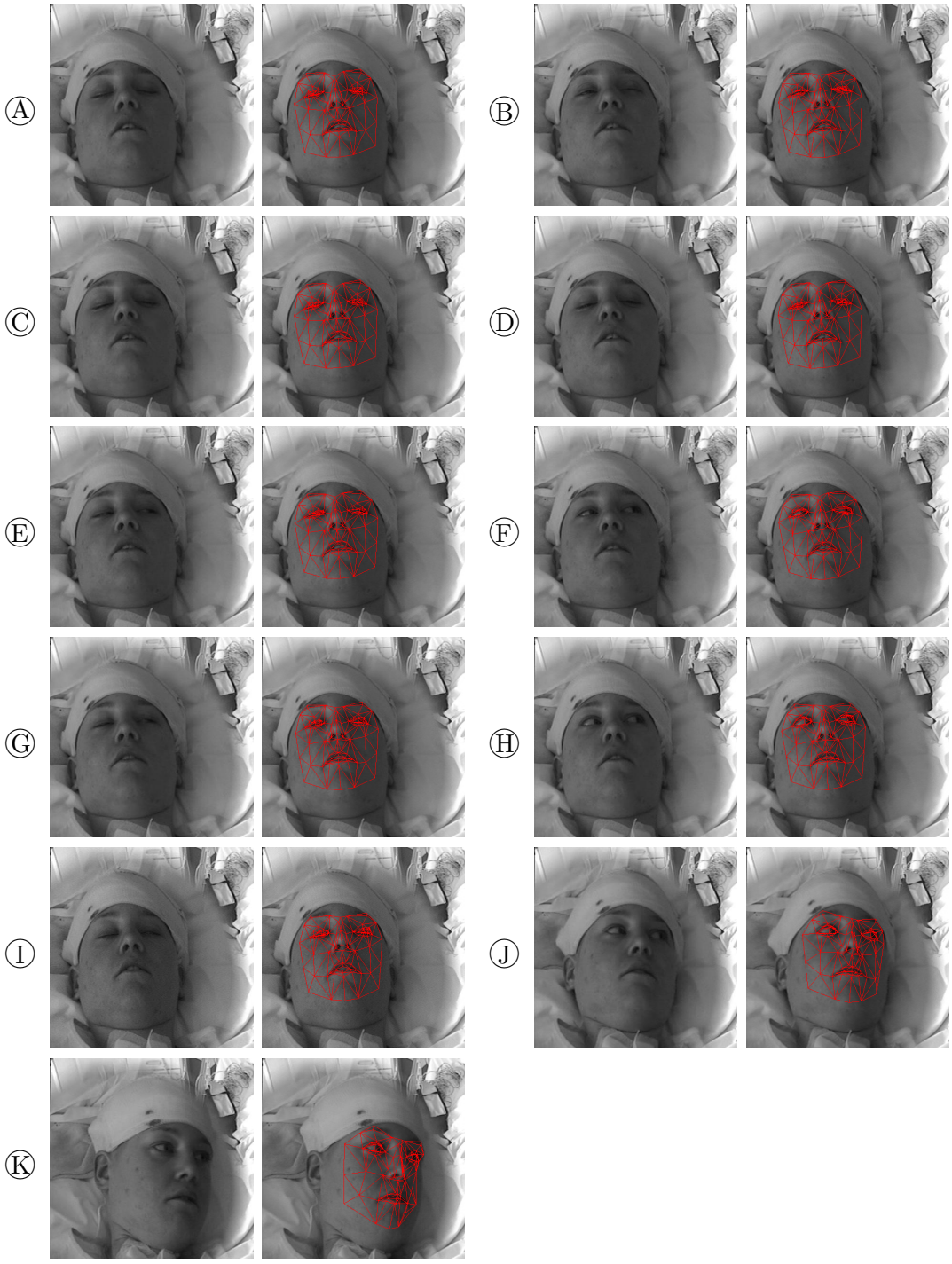
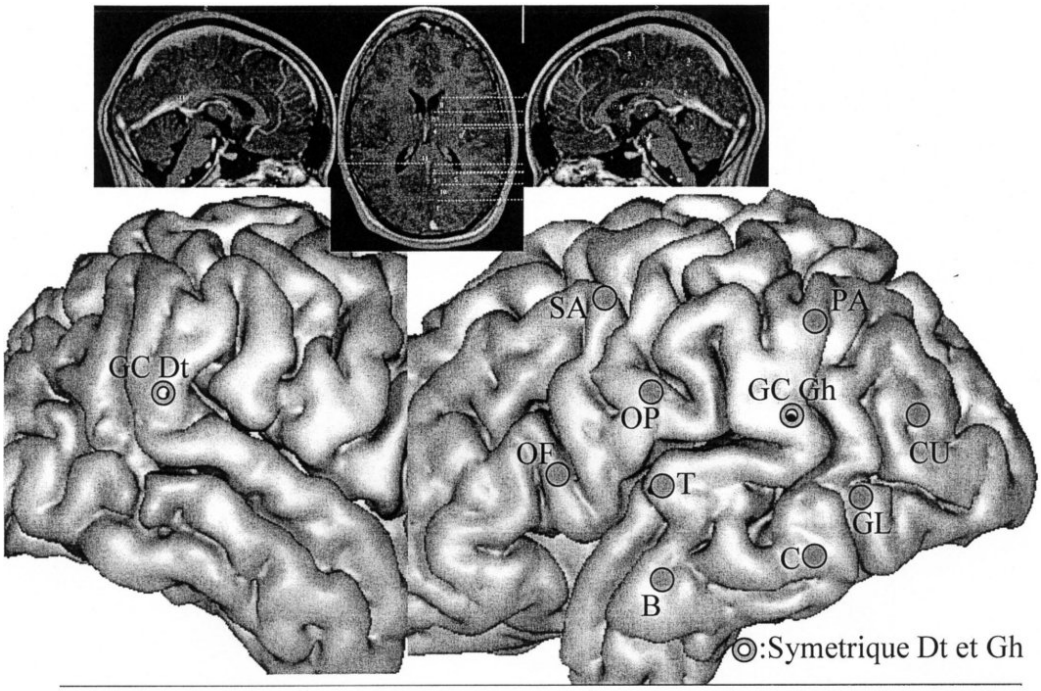


Figure 7.7: Images corresponding to the letters written on Fig. 7.6. For each column, the image frame and the result of the minimization



Implantation 10 à Gauche et 1 à Droit

1-CU	Cuneus/CuneusSupExt	15
2-GC-Gh	IsthGyrCing	15
3-PA	Area7A/ LobeParietale	15
4-T	T1 post/ Thalamus	15
5-C	ParahippoPostérieur	10
6-B	HippoCepha/ T2	10
7-OP	OperParietal	10
8-OF	OperFrontal/ N.Caudate	15
9-SA	FEF/AMS	15
10-GL	GyrusLingularis	15
11-GC-Dt	IsthGyrCing/ PeriV	10

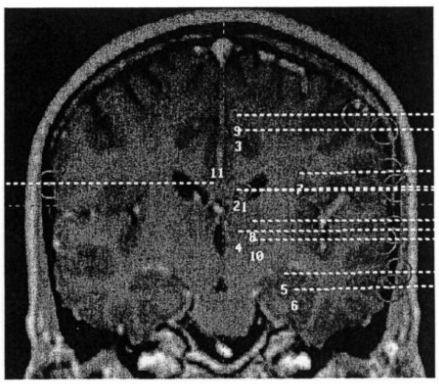


Figure 7.8: Localization of the implanted SEEG-electrodes thanks to MRI scans of the brain.

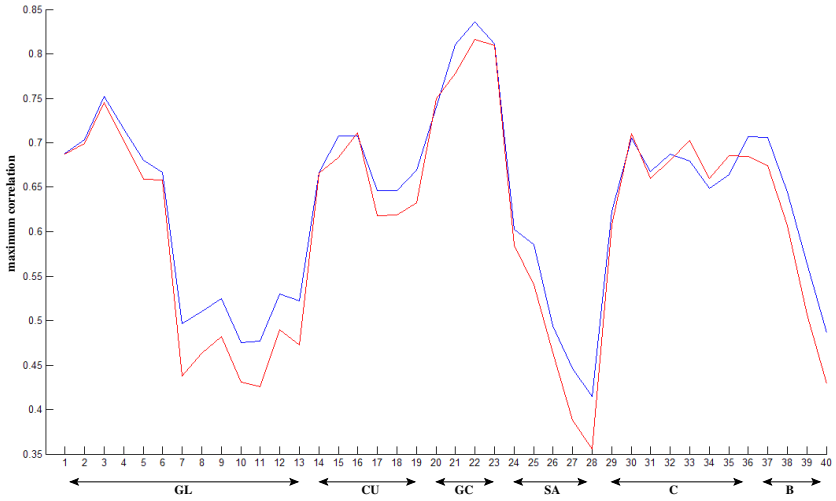


Figure 7.9: Maximum correlations between each signals and the facial parameter "opening of the eyes": in red the left eye, in blue the right eye. On the y-axis is the maximum correlation for every frequency in $\mathcal{F}$ and on the x-axis are the 40 bipolar signals. Letters under x-axis are the 6 electrodes names.

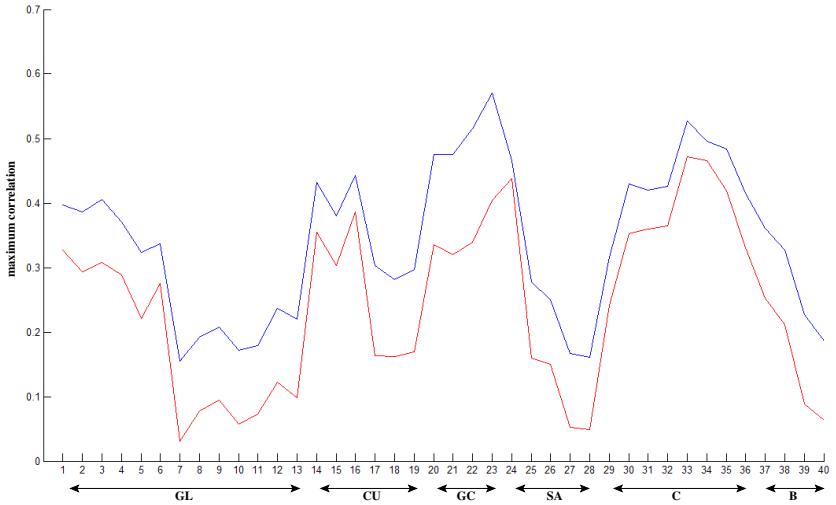


Figure 7.10: Maximum correlations between each signals and the facial parameter "position of the eyes": in red the left eye, in blue the right eye. On the y-axis is the maximum correlation for every frequency in $\mathcal{F}$ and on the x-axis are the 40 bipolar signals. Letters under x-axis are the 6 electrodes names.

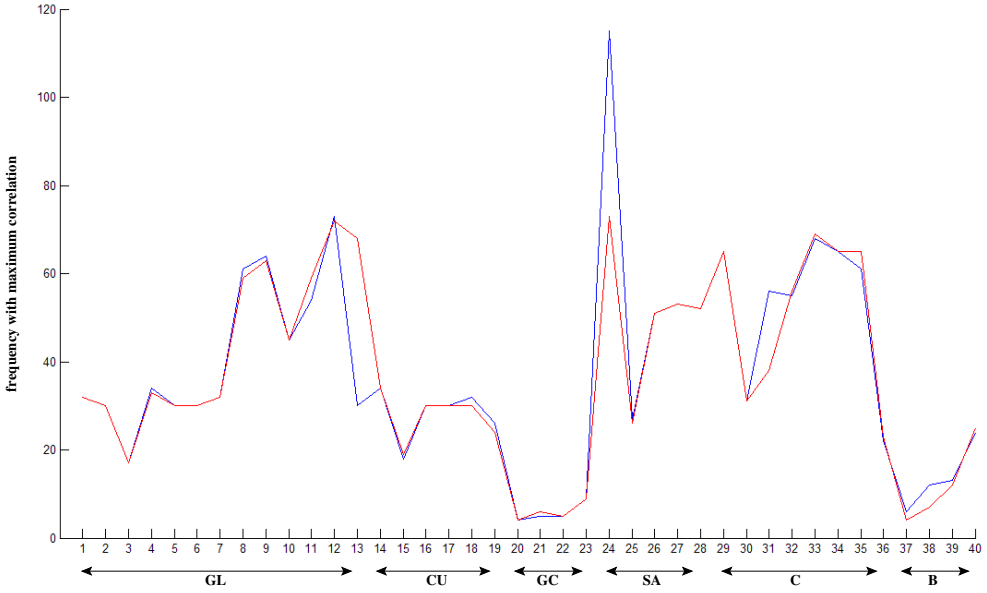


Figure 7.11: Frequencies which give the maximum correlations between each signals and the facial parameter "opening of the eyes": in red the left eye, in blue the right eye.

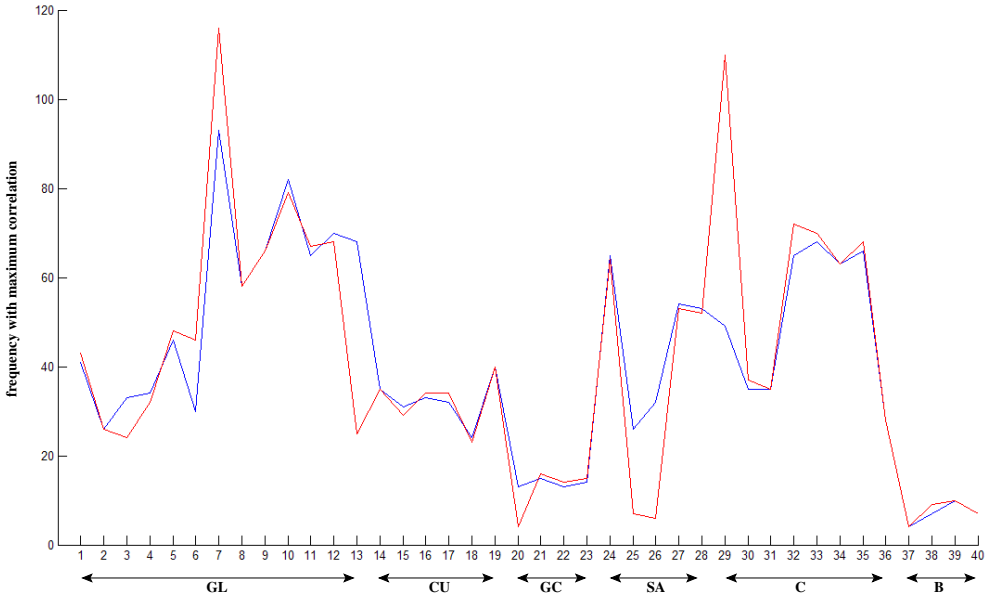


Figure 7.12: Frequencies which give the maximum correlations between each signals and the facial parameter "position of the eyes": in red the left eye, in blue the right eye.

General conclusion

From a theoretical point of view, this work consists in several investigations in the domain of shape statistics and on shape-based energy optimization. The shape warping problem is an important and difficult issue. A rigorous study on this question of deforming a shape onto another one is potentially interesting for every problem of matching (point-to-point) between shapes such as for example shape recognition or the tracking of natural moving objects. We have shown that it is possible to build metrics on the set of shapes and to define energy based on these distances. These energy can then be minimized through a gradient descent scheme.

The impact of the inner product structure of the deformation space on the behavior of the active contours method had been overlooked so far in the computer vision community. We have explored several families of inner products, as well as some minimizing flows not deriving from any inner product by extending the notion of gradient. Given an energy, we now have several ways to minimize it, each of the proposed flows being a minimizing flow but leading to different kinds of evolutions. The inner products and the extended gradients should consequently be seen as priors on the deformation fields, that is to say priors on the evolution paths. They can be used to introduce different degrees of spatial coherence (local, semi-local or global) in the evolution of the contour. We have shown, with several numerical experiments, that these evolutions better fit our intuitive notion of deformation cost and that they can mimic the behavior of the objects of interest. As a result, they are at the same time more meaningful and more robust to local minima attraction. An example with semi-locally rigid motions was specially designed. Although the results were satisfying, the particular flow resulting from our definition required to resolve a minimization problem at each step of the shape evolution and this happened to be difficult to minimize in practice by variational method. The design of more stable similar priors or the search for other minimization methods such as graph cuts for the proposed criterion should be investigated. More generally, the incorporation of linear or non-linear flow priors into the gradient descent framework can probably lead to interesting further developments.

Taking advantage of these advances on the issue of modeling prior knowledge about shapes and motivating by the practical difficulties described above, we also proposed a framework for shape warping based on both shape distances and landmarks. Our method is purely geometric and no extrinsic quantity like a space diffeomorphism has to be considered. Thanks to recent advances in the level set techniques, a level set implementation is possible, reconciling landmarks and the level set methods. Moreover, a matching be-

tween shapes is provided at no additional cost. Two- and three-dimensional examples, combining shape distance and landmarks, demonstrate the improvement brought by our approach on both warping and matching, even with a small number of landmarks. Our approach can be seen as a prior on the matching between the shapes (the landmarks points had to be matched) and also as a prior on the deformation field, since knowing which point on the first shape correspond to which point on the second one intuitively give us a preference on the path followed to deform the shape. The first prior was imposed by the modification of the shape energy to minimize and the second prior was induced by the choice of an appropriate inner product structure. One could say that this method implies a loss of autonomy in the warping operation since several landmark points have to be specified. However, in addition of allowing to process more complex shapes, let us stress that only a few number of landmark points is enough to largely improve the results (in our examples, the number of specified landmarks is typically less than 5). Moreover we can imagine a way to stay completely automated by using a detector of characteristic points (e.g. high curvature points) and by taking the result of this detection as potential landmarks points.

On the application side, the collaboration with professor Patrick Chauvel and his team at La Timone Hospital, Marseille, on the task of interrelating facial expressions and epilepsy aetiology, has proved to be very challenging, both from the computer vision point of view (capturing 3D facial expressions during crisis from low quality videos) and from the mutual discovering of a completely different world, as often with multidisciplinary project. We have proposed a method to fit a three-dimensional animated model in a monocular single image under uncontrolled imaging conditions. Our method is again based on an energy defined with a cross-correlation term and an energy minimization process. We fitted the model on real-world data obtained in a medical framework. And finally this tool has been used to investigate on its potential clinical use. We defined a procedure for measuring the potential interrelation existing between the facial expression that may occur during an epileptic seizure and the region in the brain at the origin of the seizure. This first step in the direction of applying techniques coming from domain of computer vision to a problem at the cutting edge of epilepsy research domain calls for further investigations on this question of the above-mentioned interrelation, including improving our 3D face model (e.g. building a more sophisticated statistical model of the facial texture) for the computer vision side and examining more seizures and their SEEG-recordings, for the medical side.

Of course, there is still place for improvement in many parts of this thesis. In the first part, we are convinced that our theoretical work on shape analysis can be used in number of computer vision problems, such as segmentation with a priori knowledge, shape recognition or articulated shape tracking, but a detailed inventory of these problems and a thorough experimental evaluation should be carried out. In the second part of this thesis our first results are promising and make conceivable a multi-sensor monitoring of the patients. One can also imagine an extension of this method which not only considers the evolution of the facial expressions during the seizure but also takes into account the deformation of body parts such as hands or legs. Future works should therefore include investigations in these directions.

Appendix

Some additional results: 3D face model fitting in videos

We will present some results of our 3D fitting method of a face model in several videos of epileptics seizures. For each seizure, we chose to look into several facial parameters (see figures' caption) and we displayed their variations along time and the x-axis represent the time.

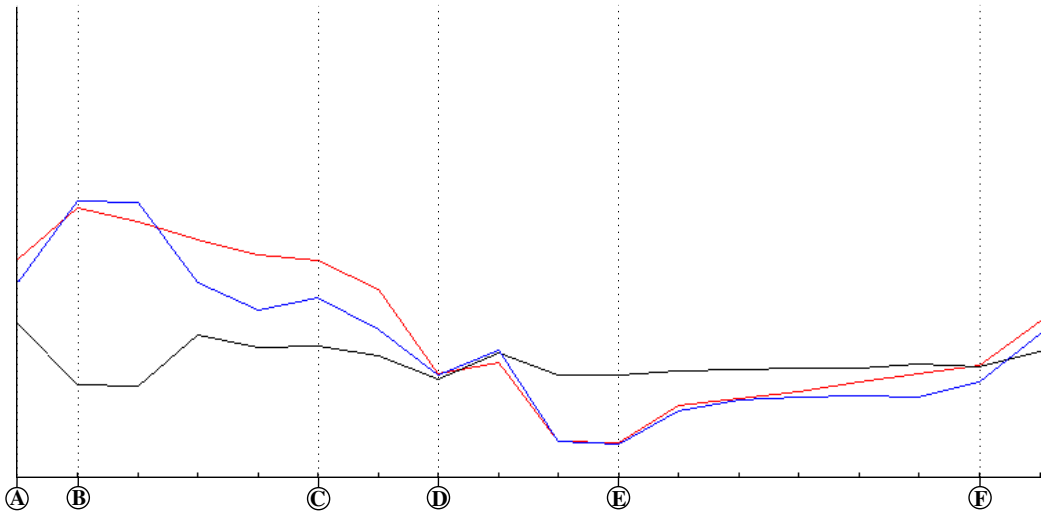


Figure 7.13: The red curve is the distance between the *left* of the mouth and the nose. The black curve is the distance between the *middle* of the mouth and the nose. The blue curve is the distance between the *right* of the mouth and the nose. Letters mark some instants showed by Fig. 7.14

Figure 7.14: Images corresponding to the letters written on Fig. 7.13. For each column, the image frame and the result of the minimization

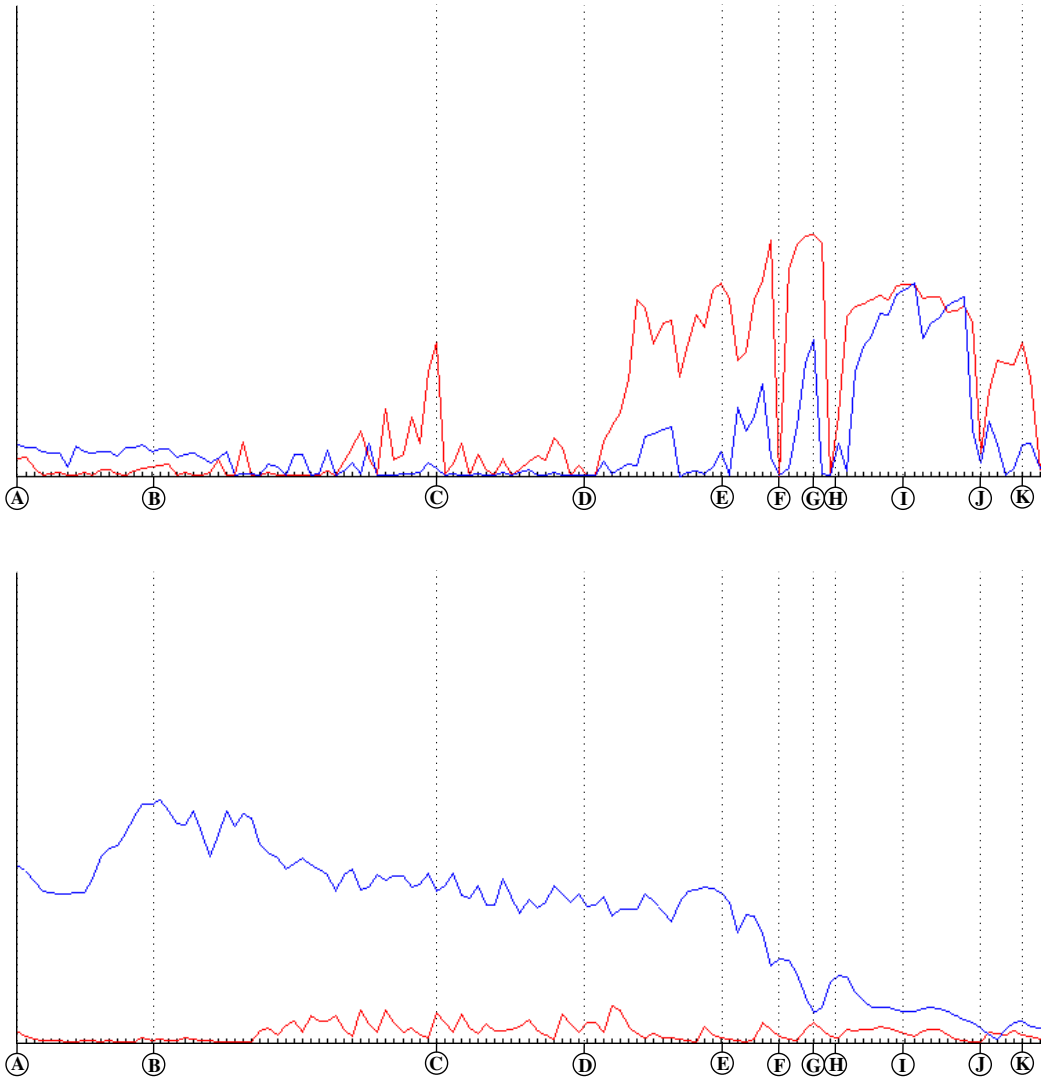


Figure 7.15: Top row: the red curve is the opening of the *left* eye and the blue one is the opening of the *right* eye. Bottom row: the red curve is the *left* opening of the mouth and the blue one the *right* opening of the mouth. Letters mark some instants showed by Fig. 7.16

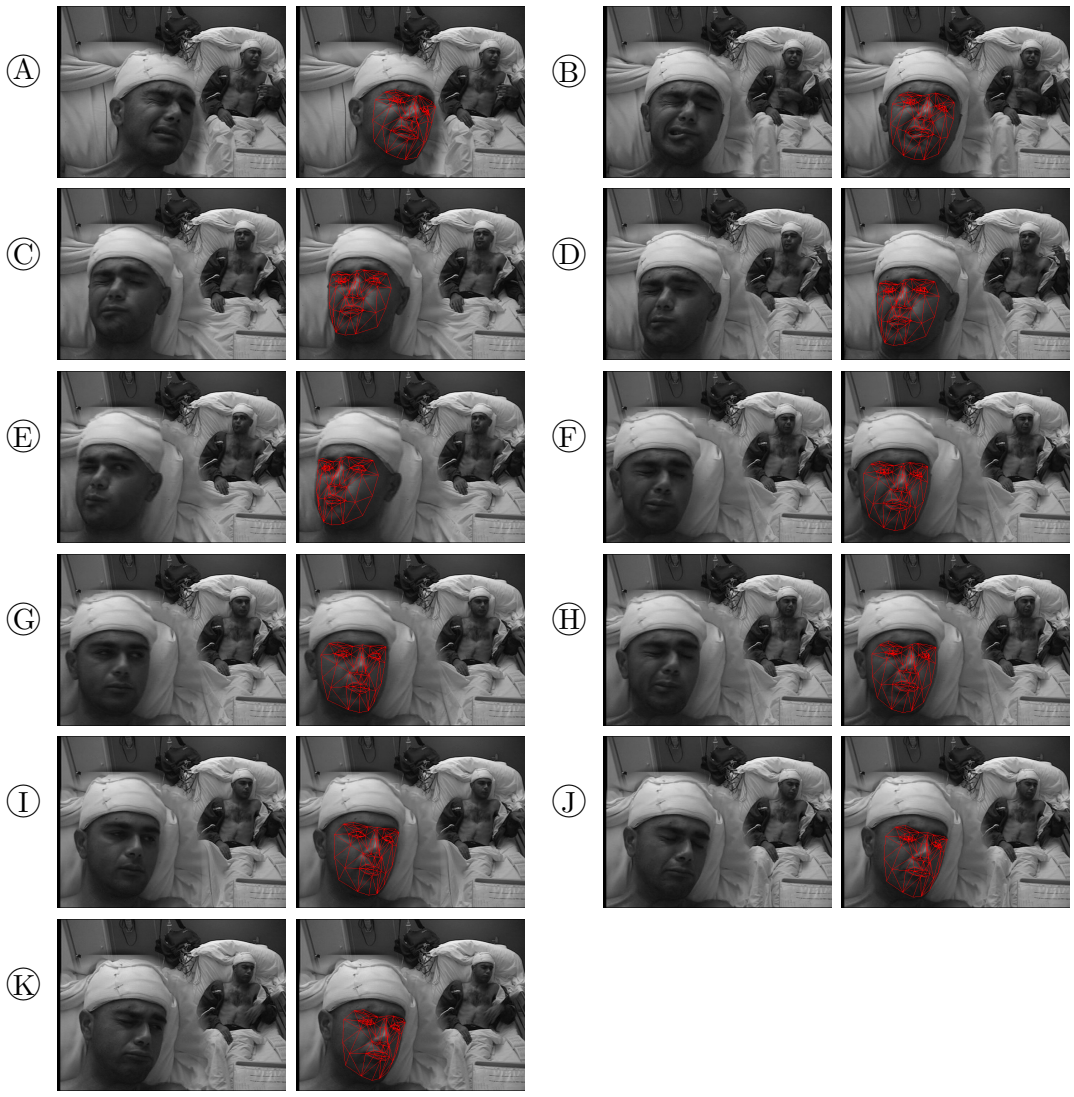


Figure 7.16: Images corresponding to the letters written on Fig. 7.15. For each column, the image frame and the result of the minimization

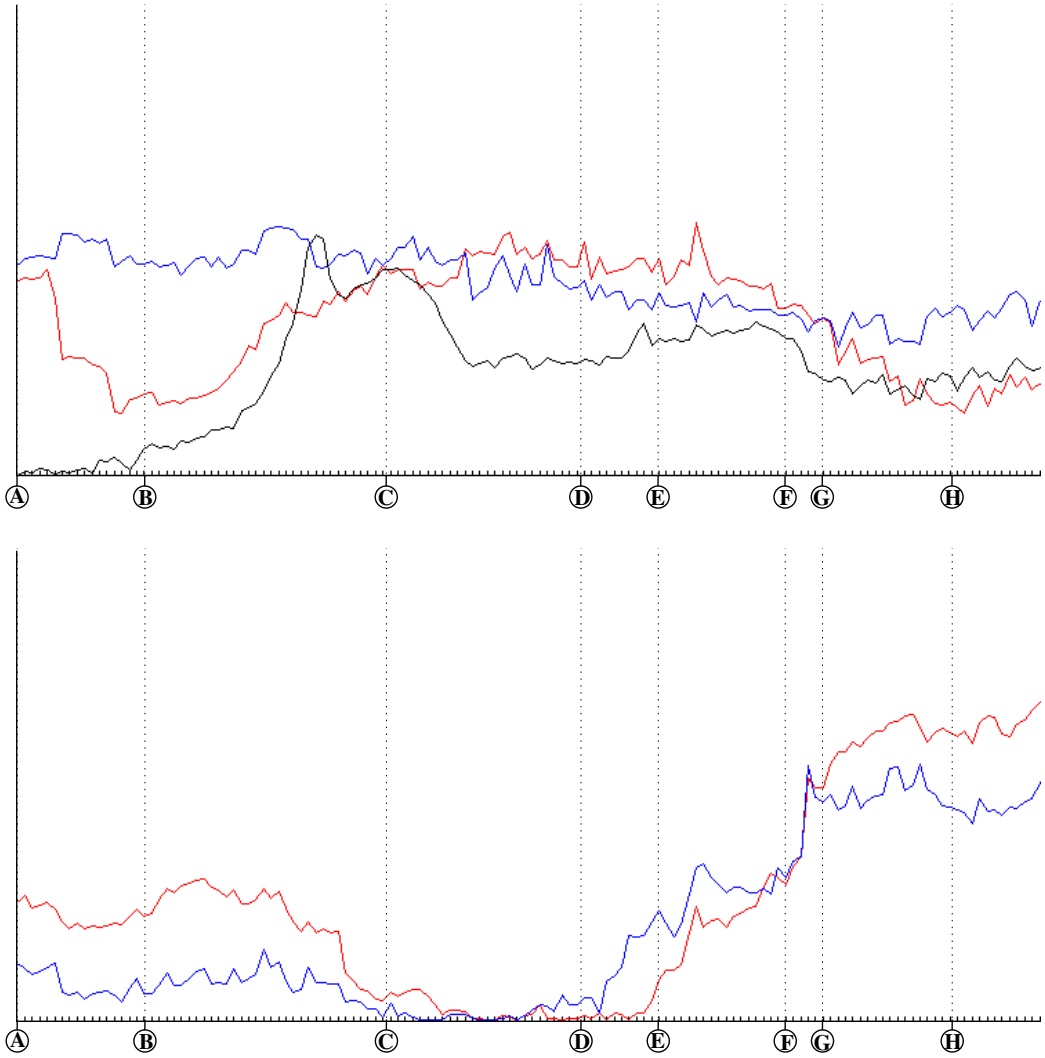


Figure 7.17: Top row: the back curve is the rotation of the head (zero for initial image), the red curve is the opening of the *left* eye and the blue one is the opening of the *right* eye. Bottom row: the red curve is the *left* opening of the mouth and the blue one the *right* opening of the mouth. Letters mark some instants showed by Fig. 7.18

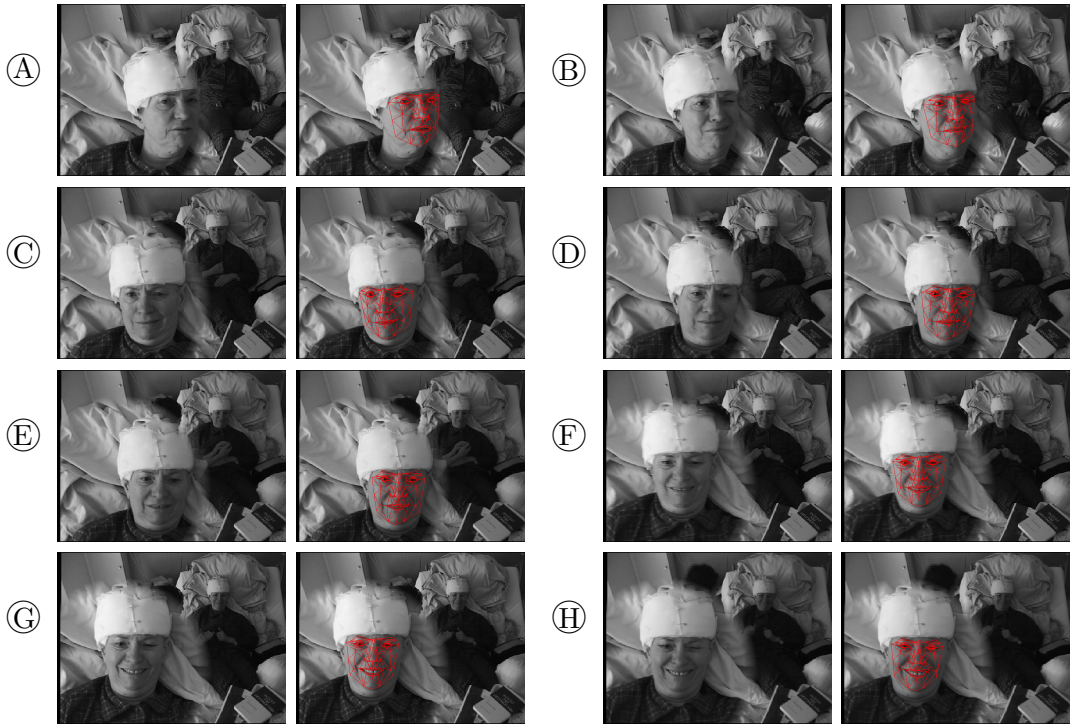


Figure 7.18: Images corresponding to the letters written on Fig. 7.17. For each column, the image frame and the result of the minimization

List of Figures

Bibliography

[1] Rafeef Abugharbieh, Clemens Kaminski, Tomas Gustavsson, and Ghassan Hamarneh. Flame front matching and tracking in plif images using geodesic paths and level sets. In *IEEE Workshop on Variational and Level Set Methods*, Vancouver, Canada, July 2001. (Cited on page 83.)

[2] David Adalsteinsson and James A. Sethian. The fast construction of extension velocities in level set methods. *Journal of Computational Physics*, 148:2–22, 1999. (Cited on page 88.)

[3] Jörgen Ahlberg. Candide 3, an updated parameterised face. (Cited on page 106.)

[4] Naum ll'ich Akhiezer and Izrail' Markovich Glazman. *Theory of Linear Operators in Hilbert Space*. Pitman, 1981. (Cited on page 54.)

[5] Shun'ichi Amari and Hiroshi Nagaoka. *Methods of Information Geometry*. American Mathematical Society, 2001. (Cited on page 19.)

[6] Jean Bancaud and Jean Talairach. Clinical semiology of frontal lobe seizures. *Adv Neurol*, 57(NIL):3–58, 1992. (Cited on page 104.)

[7] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15:1373–1396, 2003. (Cited on page 30.)

[8] Christian Benar, Christophe Grova, Eliane Kobayashi, Andrew Bagshaw, Yahya Aghakhani, Francois Dubeau, and Jean Gotman. EEG-fMRI of epileptic spikes: concordance with EEG source localization and intracranial EEG. *Neuroimage*, 30(4):1161–70, 2006. (Cited on page 129.)

[9] Marcello Bertalmio, Li-Tien Cheng, Stanley Osher, and Guillermo Sapiro. Variational problems and partial differential equations on implicit surfaces. *Journal of Computational Physics*, 174(2):759–780, 2001. (Cited on pages 62, 63, 83, and 89.)

[10] Arnaud Biraben, Delphine Taussig, P Thomas, C Even, Jean-Pierre Vignal, Jean-Marie Scarabin, and Patrick Chauvel. Fear as the main feature of epileptic seizures. *J Neurol Neurosurg Psychiatry*, 70(2):186–91, 2001. (Cited on page 104.)

[11] Arnaud Biraben, Jean-Pierre Vignal, Delphine Taussig, B. Askar, Patrick Chauvel, Xavier Morandi, and Jean-Marie Scarabin. Outcome after temporal resection for intractable epilepsy in 111 patients investigated using the anatomo-electro-clinical correlative method of bancaud and talairach. *Epilepsia 38 (suppl. 8)*, 2(107), 1997. (Cited on page 125.)

[12] Stan Birchfield. Elliptical head tracking using intensity gradients and color histograms. In *Computer Vision and Pattern Recognition*, pages 232–237. IEEE Computer Society, 1998. (Cited on page 105.)

[13] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3D faces. In Alyn Rockwood, editor, *Siggraph 1999*, pages 187–194, Los Angeles, 1999. Addison Wesley Longman. (Cited on page 106.)

[14] Frédéric Bonnans, Jean Charles Gilbert, Claude Lemarechal, and Claudia A. Sagastizabal. *Numerical Optimization: Theoretical and Practical Aspects*. Springer-Verlag, 2002. (Cited on pages 52 and 73.)

[15] Fred L. Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1:181–242, 1986. (Cited on pages 19, 33, and 81.)

[16] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:567–585, 1989. (Cited on page 82.)

[17] Fred L. Bookstein. Landmark methods for forms without landmarks: localizing group differences in outline shapes. *Medical Image Analysis*, 1(3):225–244, 1997. (Cited on page 82.)

[18] Yuri Boykov and Vladimir Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *International Conference on Computer Vision*, volume 1, pages 26–33, 2003. (Cited on page 50.)

[19] François Bremond and Monique Thonnat. Object tracking and scenario recognition for video-surveillance. In *International Joint Conference on Artificial Intelligence*, 1997. (Cited on page 126.)

[20] Vincent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997. (Cited on pages 50, 52, and 66.)

[21] Tony Chan and Luminita Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, February 2001. (Cited on page 50.)

[22] Guillaume Charpiat, Olivier Faugeras, and Renaud Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, February 2005. (Cited on pages 19, 24, 27, 40, 45, 63, and 82.)

[23] Guillaume Charpiat, Olivier Faugeras, Renaud Keriven, and Pierre Maurel. Approximations of shape metrics and application to shape warping and empirical shape statistics. In H. Krim and A. Yezzi, editors, *Statistics and Analysis of Shapes*. Birkhauser, 2006. (Cited on pages 16, 19, 20, 26, 27, 29, 33, and 40.)

[24] Guillaume Charpiat, Renaud Keriven, Jean-Philippe Pons, and Olivier Faugeras. Designing spatially coherent minimizing flows for variational problems based on active contours. In *10th International Conference on Computer Vision*, Beijing, China, 2005. (Cited on pages 49, 52, 56, and 61.)

[25] Guillaume Charpiat, Pierre Maurel, Renaud Keriven, and Olivier Faugeras. Distance-based shape statistics. *International Conference on Acoustics, Speech, and Signal Processing, Special Session: Statistical Inferences on Nonlinear Manifolds with Applications in Signal and Image Processing*, 5, 2006. (Cited on pages 16, 30, and 33.)

[26] Guillaume Charpiat, Pierre Maurel, Jean-Philippe Pons, Renaud Keriven, and Olivier Faugeras. Generalized gradients: Priors on minimization flows. *International Journal of Computer Vision*, 73(3):325–344, 2007. (Cited on pages 16 and 49.)

[27] Laurent D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing. Image Understanding*, 53(2):211–218, March 1991. (Cited on page 50.)

[28] Laurent D. Cohen and Ron Kimmel. Global minimum for active contour models: A minimal path approach. *International Journal of Computer Vision*, 24:57–78, 1997. (Cited on page 50.)

[29] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6):681–685, 2001. (Cited on pages 82 and 106.)

[30] Timothy F. Cootes, Christopher J. Taylor, David H. Cooper, and Jim Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995. (Cited on pages 82 and 106.)

[31] Michael G. Crandall, Lawrence C. Evans, and Pierre-Louis Lions. Some properties of viscosity solutions of Hamilton–Jacobi equations. *Trans. AMS*, 282:487–502, 1984. (Cited on page 85.)

[32] Daniel Cremers, Christoph Schnoerr, and Joachim Weickert. Diffusion-snakes: Combining statistical shape knowledge and image information in a variational framework. In *1st IEEE Workshop on Variational and Level Set Methods in Computer Vision*, 2001. (Cited on page 50.)

[33] Monica Crubézy, Florent Aubry, Sabine Moisan, Virginie Chameroy, Monique Thonnat, and Robert Di Paola. Managing Complex Processing of Medical Image Sequences by Program Supervision Techniques. In *SPIE international symposium Medical Imaging*, volume 3035, pages 614–625, February 1997. (Cited on page 126.)

[34] Charles Darwin. *The Expression of the Emotions in Man and Animals*. Murray, John, London, 1872. (Cited on pages 103 and 105.)

[35] M. David, Jean Talairach, and Jean Bancaud. Post-traumatic epilepsies of multiple cortical origin. *Epilepsia*, 11(1):49–58, 1970. (Cited on page 125.)

[36] Michel C. Delfour and Jean-Paul Zolésio. Shape analysis via oriented distance functions. *Journal of Functional Analysis*, 123(1):129–201, July 1994. (Cited on page 23.)

[37] Michel C. Delfour and Jean-Paul Zolésio. Shape analysis via distance functions: Local theory. In *Boundaries, interfaces and transitions*, volume 13 of *CRM Proc. Lecture Notes*, pages 91–123. AMS, Providence, RI, 1998. (Cited on pages 20, 34, and 82.)

[38] Michel C. Delfour and Jean-Paul Zolésio. *Shapes and geometries*. Advances in Design and Control. Siam, 2001. (Cited on page 23.)

[39] Alain Dervieux and François Thomasset. A finite element method for the simulation of Rayleigh-Taylor instability. *Lecture Notes in Mathematics*, 771:145–159, 1979. (Cited on pages 21, 40, 62, 63, 82, and 87.)

[40] John C. DeToledo and R. Eugene Ramsay. Patterns of involvement of facial muscles during epileptic and nonepileptic events: review of 654 events. *Neurology*, 47(3):621–5, 1996. (Cited on page 104.)

[41] Miodrag Dimitrijevic, Slobodan Ilic, and Pascal Fua. Accurate face models from uncalibrated and ill-lit video sequences. In *Conference on Computer Vision and Pattern Recognition, Washington, DC*, June 2004. (Cited on page 105.)

[42] Manfredo Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976. (Cited on page 56.)

[43] David Donoho, Mark Duncan, Xiaoming Huo, and Ofer Levi. *Wavelab*. Online Book, 1999. (Cited on pages 16 and 130.)

[44] Fadi Dornaika and Jörgen Ahlberg. Fast and reliable active appearance model search for 3-D face tracking. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 34(4):1838–1853, 2004. (Cited on page 105.)

[45] Ian L. Dryden and Kanti V. Mardia. *Statistical Shape Analysis*. John Wiley & Son, 1998. (Cited on pages 19, 33, and 81.)

[46] Ye Duan, Liu Yang, Hong Qin, and Dimitri Samaras. Shape reconstruction from 3D and 2D data using PDE-based deformable surfaces. In *European Conference on Computer Vision*, volume 3, pages 238–251, 2004. (Cited on page 50.)

[47] Paul Dupuis, Ulf Grenander, and Michael Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Math.*, 56:587–600, 1998. (Cited on page 20.)

[48] Paul Ekman and Wallace Friesen. The facial action coding system: A technique for the measurement of facial movement. *Consulting Psychologists Press*, 1978. (Cited on pages 105 and 126.)

[49] Christian E. Elger and Dieter Schmidt. Modern management of epilepsy: a practical approach. *Epilepsy Behavior*, 12(4):501–539, 2008. (Cited on pages 100, 101, and 157.)

[50] Patrick Etyngier, Florent Segonne, and Renaud Keriven. Shape priors using manifold learning techniques. In *International Conference on Computer Vision*, pages 1–8, 2007. (Cited on page 45.)

[51] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. Proceedings of the American Mathematical Society, 1998. (Cited on page 23.)

[52] Mohammed Reza Faghihi, Charles C. Taylor, and Ian L. Dryden. Procrustes shape analysis of triangulations of a two coloured point pattern. *Statistics and Computing*, 9(1):43–53, 1999. (Cited on page 19.)

[53] Beat Fasel and Juergen Luettin. Automatic Facial Expression Analysis: A Survey. *Pattern Recognition*, 36(1):259–275, 2003. IDIAP-RR 99-19. (Cited on page 105.)

[54] Olivier Faugeras and Renaud Keriven. Variational principles, surface evolution, PDE's, level set methods and the stereo problem. *IEEE Transactions on Image Processing*, 7(3):336–344, 1998. (Cited on pages 50 and 52.)

[55] Matthieu Ferrant, Arya Nabavi, Benoit Macq, Ron Kikinis, and Simon K. Warfield. Registration of 3d intraoperative mr images of the brain using a finite element biomechanical model. *IEEE Transactions on Medical Imaging*, 20(12):1384–1387, 2001. (Cited on page 83.)

[56] Robert S. Fisher, Walter van Emde Boas, Warren T. Blume, Christian Elger, Pierre Genton, Phillip Lee, and Jerome Jr Engel. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4):470–2, 2005. (Cited on page 97.)

[57] Maurice Fréchet. L'intégrale abstraite d'une fonction abstraite d'une variable abstraite et son application à la moyenne d'un élément aléatoire de nature quelconque. *Revue Scientifique*, pages 483–512, 1944. 82ème année. (Cited on page 44.)

[58] Laurent Garcin, Anand Rangarajan, and Laurent Younes. Non rigid registration of shapes via diffeomorphic point matching and clustering. In *International Conference on Image Processing*, Singapore, 2004. (Cited on page 34.)

[59] Joan A. Glaunes, Marc Vaillant, and Michael Miller. Landmark matching via large deformation diffeomorphisms on the sphere. *Journal of Mathematical Imaging and Vision*, 20:179–200, 2004. (Cited on page 34.)

[60] Joan A. Glaunes, Laurent Younes, and Alain Trouvé. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In *International Conference on Pattern Recognition*, Cambridge, 2004. (Cited on page 34.)

[61] Bastian Goldlücke and Marcus A. Magnor. Space-time isosurface evolution for temporally coherent 3D reconstruction. In *International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 350–355, 2004. (Cited on pages 50 and 52.)

[62] Bastian Goldlücke and Marcus A. Magnor. Weighted minimal hypersurfaces and their applications in computer vision. In *European Conference on Computer Vision*, volume 2, pages 366–378, 2004. (Cited on page 52.)

[63] Jose Gomes and Olivier Faugeras. Reconciling distance functions and level sets. *Journal of Visual Communication and Image Representation*, 11(2):209–223, 2000. (Cited on pages 82, 88, and 89.)

[64] Colin Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society*, B53:285–339, 1991. (Cited on page 19.)

[65] Ulf Grenander. *Lectures in Pattern Theory*. Springer, Berlin, 1976. (Cited on page 20.)

[66] Ulf Grenander. *General Pattern Theory*. Oxford University Press, 1993. (Cited on page 20.)

[67] Ulf Grenander, Yunshong Chow, and Daniel Macrae Keenan. *HANDS: A Pattern Theoretic Study of Biological Shapes*. Springer-Verlag, 1990. (Cited on page 20.)

[68] Steven Haker, Simon K. Warfield, and Clare Tempany. Landmark-guided surface matching and volumetric warping for improved prostate biopsy targeting and guidance. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI2004*, Lecture Notes in Computer Science. Springer-Verlag, 2004. (Cited on page 83.)

[69] John Raymond Hurley and Raymond B. Cattell. The procrustes program: producing a direct rotation to test an hypothesized factor structure. *Behav. Sci.*, 7:258–262, 1962. (Cited on pages 19 and 82.)

[70] Anil K. Jain and Stan Z. Li. *Handbook of Face Recognition*. Springer, June 2004. (Cited on page 126.)

[71] Hailin Jin, Stefano Soatto, and Anthony J. Yezzi. Multi-view stereo beyond Lambert. In *International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 171–178, 2003. (Cited on pages 50 and 52.)

[72] Barbara C. Jobst, Adrian M. Siegel, Vijay M. Thadani, David W. Roberts, Harker C. Rhodes, and Peter D. Williamson. Intractable seizures of frontal lobe origin: clinical characteristics, localizing signs, and results of surgery. *Epilepsia*, 41(9):1139–52, 2000. (Cited on page 104.)

[73] Hermann Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30:509–669, 1976. (Cited on page 44.)

[74] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. In *First International Conference on Computer Vision*, pages 259–268, London, June 1987. (Cited on page 50.)

[75] David G. Kendall. The diffusion of shape. *Advances in Applied Probability*, 9:428–430, 1977. (Cited on page 19.)

[76] David G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of London Mathematical Society*, 16:81–121, 1984. (Cited on pages 19 and 82.)

[77] Satyanad Kichenassamy, Arun Kumar, Peter J. Olver, Allen Tannenbaum, and Anthony J. Yezzi. Gradient flows and geometric active contour models. In *ICCV '95: Proceedings of the Fifth International Conference on Computer Vision*, page 810, Washington, DC, USA, 1995. IEEE Computer Society. (Cited on page 50.)

[78] Junmo Kim, John Fisher, Anthony J. Yezzi, Mujdat Cetin, and Alan Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *IEEE Internation Conference on Image Processing*, pages 797–800, September 2002. (Cited on page 50.)

[79] Vladimir Kolmogorov and Ramin Zabih. Multi-camera scene reconstruction via graph cuts. In *European Conference on Computer Vision*, volume 3, pages 82–96, 2002. (Cited on page 50.)

[80] Vladimir Kolmogorov and Ramin Zabih. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004. (Cited on page 51.)

[81] Jean-Philippe Lachaux, Eugenio Rodriguez, Jacques Martinerie, and Francisco J. Varela. Measuring phase synchrony in brain signals. *Hum Brain Mapp*, 8(4):194–208, 1999. (Cited on page 129.)

[82] Rasmus Larsen. Functional 2D procrustes shape analysis. In Jussi Parkkinen Heikki Kalviainen and Arto Kaarna, editors, *14th Scandinavian Conference on Image Analysis*, volume 3540 of *LNCS*, pages 205–213, Berlin Heidelberg, jun 2005. Springer Verlag. (Cited on page 19.)

[83] Rasmus Larsen. L1 generalized procrustes 2D shape alignment. *Journal of Mathematical Imaging and Vision*, 2008. (Cited on page 19.)

[84] Craig T. Lawrence and André L. Tits. A computationally efficient feasible sequential quadratic programming algorithm. *SIAM J. on Optimization*, 11(4):1092–1118, 2000. (Cited on pages 16 and 117.)

[85] Alex Leow, Ming-Chang Chiang, Hillary Protas, Paul Thompson, Luminata Vese, and Sung-Cheng Huang. Linear and non-linear geometric object matching with implicit representation. In *International Conference on Pattern Recognition*, Cambridge, 2004. (Cited on page 82.)

[86] Alex Leow, Andrew Y. Wang, Hillary Protas, and Paul Thompson. Brain warping via landmark points and curves with a level set representation. Technical Report 4-14, UCLA CAM Report, 2004. (Cited on page 82.)

[87] Michael E. Leventon, Olivier Faugeras, W. Eric L. Grimson, and William Wells. Level Set Based Segmentation with Intensity and Curvature Priors. In *Mathematical Method in Biomedical Image Analysis*, 2000. (Cited on page 82.)

[88] Michael E. Leventon, W. Eric L. Grimson, and Olivier Faugeras. Statistical shape influence in geodesic active contours. *International Conference on Computer Vision and Pattern Recognition*, 01:1316, 2000. (Cited on page 50.)

[89] Wei-Hsun Liao, A. Khuu, Marvin Bergsneider, Luminita A Vese, Sung-Cheng Huang, and Stanley Osher. From landmark matching to shape and open curve matching: a level set approach. In *IEEE Conference on Vision and Pattern Recognition*, Madison, 2003. (Cited on page 82.)

[90] Wei-Hsun Liao, Hillary Protas, Marvin Bergsneider, Luminita A. Vese, Sung-Cheng Huang, and Stanley Osher. A new framework for object warping: a semi-lagrangian level set approach. In *International Conference on Image Processing*, Barcelona, 2003. (Cited on page 82.)

[91] Ravikanth Malladi, James A. Sethian, and Baba C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, February 1995. (Cited on page 50.)

[92] M. Mandal, R. Pandey, and A. Prasad. Facial expressions of emotions and schizophrenia: a review. *Schizophrenia Bull*, 24(3):399–412, 1998. (Cited on page 125.)

[93] Pierre Maurel, Renaud Keriven, and Olivier Faugeras. Reconciling landmarks and level sets. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 69–72, Washington, DC, USA, 2006. IEEE Computer Society. (Cited on pages 16, 52, and 81.)

[94] Pierre Maurel, Aileen McGonigal, Patrick Chauvel, and Renaud Keriven. 3d model fitting for facial expression analysis under uncontrolled imaging conditions. In *19th International Conference on Pattern Recognition*, Tampa, US, Dec 2008. (Cited on pages 16 and 105.)

[95] Pierre Maurel and Guillermo Sapiro. Dynamic shapes average. In *IEEE 2nd Int. Workshop on Variational, Geometric and Level Set Methods, ICCV*, Nice, 2003. (Cited on page 45.)

[96] Stephen J. Maybank. The fisher-rao metric for projective transformations of the line. *International Journal of Computer Vision*, 63(3):191–206, 2005. (Cited on page 19.)

[97] Stephen J. McKenna, Shaogang Gong, Rolf P. Würtz, Jonathan Tanner, and Daniel Banin. Tracking facial feature points with Gabor wavelets and shape models. In *Int. Conference on Audio- and Video-based Biometric Person Authentication*, 1997. (Cited on page 105.)

[98] Peter W. Michor and David Mumford. Riemannian geometries on spaces of plane curves. *J.EUR.MATH.SOC*, 8:1, 2006. (Cited on page 52.)

[99] Michael Miller and Laurent Younes. Group actions, homeomorphisms, and matching : A general framework. *International Journal of Computer Vision*, 41(1/2):61–84, 2001. (Cited on pages 19, 20, and 34.)

[100] Facundo Mémoli. On the use of Gromov-Hausdorff Distances for Shape Comparison. In M. Botsch, R. Pajarola, B. Chen, and M. Zwicker, editors, *Symposium on Point Based Graphics*, pages 81–90, Prague, Czech Republic, 2007. Eurographics Association. (Cited on page 26.)

[101] Facundo Mémoli. Gromov-hausdorff distances in euclidean spaces. In *Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (CVPR workshop, NORDIA'08)*, june 2008. (Cited on page 26.)

[102] R. L. Morrison, Bellack A. S., and K. T. Mueser. Deficits in facial-affect recognition and schizophrenia. *Schizophrenia Bull*, 14(1):67–83, 1988. (Cited on page 125.)

[103] David Mumford and Jayant Shah. Boundary detection by minimizing functionals, I. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 22–26, 1985. (Cited on page 50.)

[104] Delphine Nain, Anthony J. Yezzi, and Greg Turk. Vessel segmentation using a shape driven flow. In *In Medical Imaging Copmuting and Computer-Assisted Intervention*, pages 51–59, 2004. (Cited on page 50.)

[105] Stanley Osher and Ronald P. Fedkiw. Level set methods: An overview and some recent results. *J. Comput. Phys*, 169:463–502, 2001. (Cited on page 88.)

[106] Stanley Osher and Ronald P. Fedkiw. *The Level Set Method and Dynamic Implicit Surfaces*. Springer-Verlag, 2002. (Cited on page 63.)

[107] Stanley Osher and Nikos Paragios, editors. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003. (Cited on page 63.)

[108] Stanley Osher and James Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988. (Cited on pages 21, 40, 62, 63, 82, 87, and 88.)

[109] Niels Chr. Overgaard and Jan Erik Solem. An analysis of variational alignment of curves in images. In *The 5th International Conference on Scale Space and PDE methods in Computer Vision, Scale Space 2005, Hofgeismar, Germany*. Springer, 2005. (Cited on page 53.)

[110] Maja Pantic and Leon J. M. Rothkrantz. Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1424–1445, 2000. (Cited on page 105.)

[111] Nikos Paragios and Rachid Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:266–280, March 2000. (Cited on page 50.)

[112] Nikos Paragios and Rachid Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. Comput. Vision*, 46(3):223–247, 2002. (Cited on page 50.)

[113] Nikos Paragios and Rachid Deriche. Geodesic active regions and level set methods for motion estimation and tracking. *Comput. Vis. Image Underst.*, 97(3):259–282, 2005. (Cited on page 66.)

[114] Nikos Paragios, Mickael Rousson, and Visvanathan Ramesh. Matching distance functions: A shape-to-area variational approach for global-to-local registration. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part II*, pages 775–789, London, UK, 2002. Springer-Verlag. (Cited on page 82.)

[115] Danping Peng, Barry Merriman, Stanley Osher, Hong-Kai Zhao, and Myungjoo Kang. A PDE-based fast local level set method. *Journal of Computational Physics*, 155(2):410–438, 1999. (Cited on pages 63, 82, and 88.)

[116] Adrian Peter and Anand Rangarajan. Shape analysis using the fisher-rao riemannian metric: unifying shape representation and deformation. In *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*, pages 1164–1167, 2006. (Cited on page 19.)

[117] Jean-Philippe Pons, Gerardo Hermosillo, Renaud Keriven, and Olivier Faugeras. How to deal with point correspondences and tangential velocities in the level set framework. In *In International Conference on Computer Vision*, volume 2, pages 894–899, 2003. (Cited on pages 66, 83, 85, and 90.)

[118] Calyampudi Radhakrishna Rao. Information and accuracy attainable in estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–91, 1945. (Cited on page 19.)

[119] Tammy Riklin-raviv, Nahum Kiryati, and Nir Sochen. Unlevel-sets: Geometry and prior-based segmentation. In *In ECCV*, pages 50–61. Springer, 2004. (Cited on page 50.)

[120] Mikael Rousson and Nikos Paragios. Shape priors for level set representations. In *In ECCV*, pages 78–92. Springer, 2002. (Cited on page 50.)

[121] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, 1966. (Cited on pages 27, 28, 36, and 52.)

[122] Jean-Marie Scarabin, Jean-Pierre Vignal, Delphine Taussig, B. Askar, Xavier Morandi, Patrick Chauvel, and Arnaud Biraben. Outcome after extratemporal cortectomies for intractable epilepsy in 31 patients investigated using the anatomo-electro-clinical correlative method of bancaud and talairach. *Epilepsia 38 (suppl. 8)*, 2(108), 1997. (Cited on page 125.)

[123] Lofti Senhadji, Mohammad Bagher Shamsollahi, and Régine Le Bouquin-Jeannes. Representation of seeg signals using time-frequency signatures. In *20th International Conference in Engineering in Medicine and Biology Society*, volume 3, pages 1454–1457, Oct. 1998. (Cited on page 129.)

[124] James A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Sciences*. Cambridge Monograph on Applied and Computational Mathematics. Cambridge University Press, 1999. (Cited on pages 63 and 85.)

[125] Daniel Sharvit, Jacky Chan, Hüseyin Tek, and Benjamin B. Kimia. Symmetry-based indexing of image databases. *Journal of Visual Communication and Image Representation*, 9(4):366–380, December 1998. (Cited on pages 42, 44, and 155.)

[126] Simon D. Shorvon. *Handbook of epilepsy treatment*. Blackwell Science, 2000. (Cited on pages 99, 103, and 157.)

[127] Christopher G. Small. *The Statistical Theory of Shapes*. Springer-Verlag, 1996. (Cited on page 19.)

[128] Jan Erik Solem and Niels Chr. Overgaard. A geometric formulation of gradient descent for variational problems with moving surfaces. In *International Conference on Scale Space and PDE Methods in Computer Vision*, pages 419–430, 2005. (Cited on pages 52 and 53.)

[129] Ivan Soltesz and Staley Kevin. *Computational Neuroscience in Epilepsy*. Academic Press, 2008. (Cited on page 125.)

[130] Ganesh Sundaramoorthi, Anthony J. Yezzi, and Andrea C. Mennucci. Sobolev active contours. In *IEEE Workshop on Variational and Level Set Methods*, pages 109–120, Beijing, China, 2005. (Cited on pages 52 and 61.)

[131] Ganesh Sundaramoorthi, Anthony J. Yezzi, and Andrea C. Mennucci. Sobolev active contours. *Int. J. Comput. Vision*, 73(3):345–366, 2007. (Cited on page 52.)

[132] Ganesh Sundaramoorthi, Anthony J. Yezzi, Andrea C. Mennucci, and Guillermo Sapiro. New possibilities with sobolev active contours. (Cited on page 52.)

[133] Mark Sussman, Peter Smereka, and Stanley Osher. A level set approach for computing solutions to incompressible two-phase flow. *Journal of Computational Physics*, 114(1):146–159, 1994. (Cited on page 42.)

[134] Motoi Suwa, Noboru Sugie, and Keisuke Fujimora. A preliminary note on pattern recognition of human emotional expression. *4th International Joint Conference on Pattern Recognition*, pages 408–410, 1978. (Cited on page 105.)

[135] Arthur W. Toga, editor. *Brain Warping*. Academic Press, 1998. (Cited on pages 19 and 34.)

[136] Arthur W. Toga and Paul Thompson. The role of image registration in brain mapping. *Image and Vision Computing*, 19(1-2):3–24, 2001. (Cited on page 34.)

[137] Alain Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–21, 1998. (Cited on pages 20 and 52.)

[138] Alain Trouvé and Laurent Younes. Diffeomorphic matching problems in one dimension: Designing and minimizing matching functionals. In *ECCV '00: Proceedings of the 6th European Conference on Computer*

Vision-Part I, pages 573–587, London, UK, 2000. Springer-Verlag. (Cited on pages 20 and 34.)

[139] Van-Thinh Vu, François Brémond, and Monique Thonnat. Human behaviour visualisation and simulation for automatic video understanding. In *International Conference in Central Europe on Computer Graphics and Visualization*, pages 485–492, 2002. (Cited on page 126.)

[140] Peng Wang, Fred Barrett, Elizabeth Martin, M. Milanova, Raquel Gur, Ruben Gur, Christian Kohler, and Ragini Verma. Automated video based facial expression analysis of neuropsychiatric disorders. *Journal of Neuroscience Methods*, 168, February 2008. (Cited on page 126.)

[141] Peng Wang, Christian Kohler, Fred Barrett, Raquel Gur, Ruben Gur, and Ragini Verma. Quantifying facial expression abnormality in schizophrenia by combining 2d and 3d features. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007. (Cited on page 126.)

[142] Song Wang, Jim Xiuquan Ji, and Zhi-Pei Liang. Landmark-based shape deformation with topology preserving constraints. In *9th International Conference on Computer Vision*, Nice, France, 2003. (Cited on page 82.)

[143] Fabrice Wendling, Fabrice Bartolomei, Jean-Jacques Bellanger, Jérome Bourien, and Patrick Chauvel. Epileptic fast intracerebral EEG activity: evidence for spatial decorrelation at seizure onset. *Brain*, 126(Pt 6):1449–59, 2003. (Cited on page 129.)

[144] Fabrice Wendling, Mohammad Shamsollahi, Jean-Michel Badier, and Jean-Jacques Bellanger. Time-frequency matching of warped depth-EEG seizure observations. *IEEE Trans Biomed Eng*, 46(5):601–5, 1999. (Cited on page 129.)

[145] Anthony J. Yezzi and Stefano Soatto. Deformotion: Deforming motion, shape average and the joint registration and approximation of structures in images. *Int. J. Comput. Vision*, 53(2):153–167, 2003. (Cited on pages 20, 24, 63, and 82.)

[146] Anthony J. Yezzi, Andy Tsai, and Alan Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *ICCV '99: Proceedings of the International Conference on Computer Vision-Volume 2*, page 898, Washington, DC, USA, 1999. IEEE Computer Society. (Cited on page 50.)

[147] Laurent Younes. Computable elastic distances between shapes. *SIAM Journal of Applied Mathematics*, 58(2):565–586, 1998. (Cited on pages 20, 34, and 63.)

[148] Laurent Younes. Optimal matching between shapes via elastic deformations. *Image and vision computing journal*, 17(5/6):381–389, April 1999. (Cited on page 20.)

[149] Hong-Kai Zhao, Stanley Osher, Barry Merriman, and Myungjoo Kang. Implicit and non-parametric shape reconstruction from unorganized points using a variational level set method. *Computer Vision and Image Understanding*, 80(3):295–314, 2000. (Cited on page 50.)

[150] Song Chun Zhu and Alan Yuille. Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, September 1996. (Cited on page 50.)