Analyzing of Facial Paralysis by Shape Analysis of 3D Face Sequences

Paul Audain Desrosiers, Yasmine Ben Bennis, Mohamed Daoudi, Boulbaba Ben Amor, Pierre Guerreschi

Analyzing of Facial Paralysis by Shape Analysis of 3D Face Sequences

Paul Audain Desrosiers[a], Yasmine Bennis[b], Mohamed Daoudi[a,*], Boulbaba Ben Amor[a], Pierre Guerreschi[b]

[a] *IMT Lille Douai, Univ. Lille, CNRS, UMR 9189 - CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France*
[b] *Service de chirurgie plastique reconstructrice et esthétique, Université Lille 2 Droit et Santé, Lille, France*

Abstract

In this paper, we address the problem of quantifying the facial asymmetry from 3D face sequence (4D). We investigate the role of 4D data to reveal the amount of both static and dynamic asymmetry in the clinical case of facial paralysis. The goal is to provide tools to clinicians to evaluate quantitatively facial paralysis treatment based on Botulinum Toxin (BT), which can provide qualitative and quantitative evaluations. To this end, Dense Scalar Fields (DSFs), based on Riemannian analysis of 3D facial shape, is proposed to quantify facial deformations. To assess this approach, a new 3D facial sequences of 16 patients data set is collected, before and after injecting the BT. For each patient, we have collected 8 facial expressions before and after injecting BT. Experimental results obtained on this data set show that the proposed approach allows clinicians to evaluate more accurately the facial asymmetry before and after the treatment.

Keywords: 4D faces, facial asymmetry, facial paralysis, Dense Scalar Fields.

1. INTRODUCTION

Human face is the most important non-verbal channel which allows to human beings to communicate with their peers. Facials muscles dysfunction may

[☆] This paper was presented in part at Visapp 2016 [1]
[*] Corresponding author

causes discomfort among patients affected by facial paralysis. It occurs in the case of partial or complete loss of functioning of certain facial muscles. They are caused by a lesion of the facial nerve for different reasons – acute idiopathic facial paralysis (no cause) called a *Bell's palsy*, trauma, infection, chronic illness, tumor, etc. On the healthy side of the paralyzed face, a compensatory muscular hyperactivity will be established progressively, hence increases the facial asymmetry. That is, to compensate for the dysfunction of the paralyzed side, the patient uses exaggeratedly the non-paralyzed muscles which will have the effect of marking wrinkles and deforming the face since the antagonist balance is no longer assured. In general, Wrinkles are more pronounced (forehead, nasolabial folds, crow's feet) with an elevated eyebrow and deviation of the tip of the nose, and the mouth on the healthy side. On the paralyzed side, involuntary muscle contraction can appear following the voluntary contraction of other muscles. In general, they appear 3 to 6 months after the onset of the disease. The most frequent are the zygomatic-orbicular co-contraction of the eye, zygomatic-occipitofrontalis, zygomatic-platysma and zygomatic-orbicular lip [2, 3]. One common medical treatment, widely used since 1989 in surgery and aesthetic medicine [4, 5], is to inject in the muscles of the face low doses of Botulinum Toxin (BT) to handle compensatory the over-activity of the healthy muscles. The role of BT treatment will be to attenuate muscular hyperactivity of the healthy part of the face to improve the harmony and the facial symmetry. The main difficulty lies in the evaluation of the areas to be treated, the dose to be injected and the results. In general, the results of the medical treatments are evaluated subjectively by the clinicians after few weeks of each BT injection, thus no automated system exists to assist them conducting their assessment. The most conventional procedure consists to ask the patient to convey a specific facial expression like to read a text, or to repeat a specific sentence and record him/her using a 2D camera. The observation of the 3D sequences (recorded before the BT injection) and the current dynamic allows the clinician to evaluate the results and thus monitoring remaining treatments (BT, surgery, etc.).

To our knowledge, no automated tools exists nowadays to allow clinicians accurate evaluation of the BT injection, in terms of facial dynamics.

1.1. Prior work

The recent advances of 3D/4D sensors based on two main technologies – stereo-photogrammetry and structured-light – opened the door to develop approaches for analyzing static (3D) and dynamic (4D) shapes. In particular, targeting the face recognition or expression classification problems.

3D dynamic face analysis: In [6], a 3D facial expression recognition approach is proposed. It is based on a Normal Map and Shape Index Map to describe the geometry attributes of the facial surface. In [7], a 3D expression descriptor namely Multi-Scale Local Normal Patterns is proposed. However, it is not clear how these approaches can be generalized to temporal sequences. In [8], a 3D motion-based features between frames of 3D facial geometry sequences for dynamic facial expression recognition is proposed. These features are then used to train GentleBoost classifiers and build a Hidden Markov Model in order to model the temporal dynamics of the expression. In [9], the 3D faces are represented by collections of radial curves and a Riemannian shape analysis is applied to quantify the deformations induced by the facial expressions, in a given sub-sequence of 3D frames. In [10], an automatic 3D/4D facial expression recognition based on the muscular movement model (MMM). The reader can find other 3D face analysis approaches in [11], [12], [13].

These approaches have demonstrated the role of 3D dynamic face analysis to reveal deformations hidden in 2D videos. In fact, 2D images resulted from the projection of the observed scene into a plane (X-Y) which conducts to loose the depth coordinate (Z). Using the 3D acquisition sensors allows to recover the Z-coordinates and thus the use of a the complete shape of the 3D face. Current literature is rich of approaches for shape analysis of 3D and 4D faces. However, only few studies have addressed the fundamental question of – How to quantify the amount of bilateral facial asymmetry from the temporal evolution of 3D faces?

3D facial asymmetry analysis: In this context, earlier face analysis approaches are based on the detection and tracking of 2D and 3D facial landmarks in facial sequences. For instance, Al-Anezi et al. [14] have proposed a new method for automatic tracking of facial landmarks in 3D sequences. A set of 23 anthropometrics landmarks with no permanent ink are marked on the faces of 32 subjects aged 18 to 35 years. The subjects are asked to perform a facial animation like maximal smile, lip purse, cheek puff. Al-Anezi et al. [14] have concluded that an accurate tracking solution facilitate the analysis of the dynamic motion. Shujaat et al. [15] have pursued the study conducted in [14] and have developed a new method to quantify dynamic 3D facial animations, in order to characterize the dynamics of 3D lips movement in the head, and neck oncology patients before and after lower lip split mandibulotomy. A data set of 7 subjects aged 42-80 years old is collected. Moreover, a set of nine facials soft tissue landmarks are manually annotated on the first frame of each 3D sequence by an operator, and then tracked automatically. Six landmarks are used to analyze the lip motions, assessing of the magnitude, and the three others to track the effect of the head motion, and to align all the frames into a common reference frame. They measured the feasibility of the change in the magnitude, speed, and motion similarity of facial animation. The results show that the magnitude and the speed difference decreased after the surgical operation for the smile and the lip purse, while motion similarity shows a high score in the case of lip purse animation. Wei et al. [16] proposed a novel approach to analyze the asymmetry of the face on 3D dynamic scans. A data set which comes from the Hi4D-ADSIP database is used, and allow them to make a comparison between stroke patients and healthy individuals. At first sight, in each frame, the nose tip is determined, and a sphere of radius r = 130 mm is centered at the nose tip for segment of the 3D face. Then, with their proposed method, they extracted the symmetry plan of the 3D face. Thus, the asymmetry can be extracted directly from the symmetry plan by using the original data and the mirrored data obtained by reflection across the symmetry plan. After matching the original 3D face and its mirror using the well-known Iterative Closest Point algorithm,

they retained the Euclidean distance between each point of the 3D original face and its mirror. From this, they measured the higher asymmetry level of the 3D face which based on the bigger value of the average closest distance.

More recently, Gaber et al. [17] used the Microsoft Kinect v2 to quantify the facial paralysis of the human face. In their proposed framework, they used 16 landmarks annotations, 2 landmarks for eyebrows, 8 landmarks for the eyes, 1 landmark for the nose tip, 4 landmarks for the mouth, and 1 landmark for the chin. From the landmark positions, the symmetry plan of the face can be calculated through the landmark in the nose tip. Then, they defined a Symmetry Index that is based on the distance between the two identical parts of the face. The Symmetry Index (SI) between both eyebrows, both eyes, and both sides of the mouth is computed as the ratio between the two related distances. The SI can be defined as a single value to quantify the symmetry of each part of the selected features. Thus, for a perfect face, the SI is equal to 100%. Ten subjects were invited to the experiments, all the subjects are between 22-28 years old. The mean of the SI for all the subjects is varied from 95.6% to 99.63%. However, all the results were obtained on the healthy faces without facial paralysis. A paralyzed faces are less symmetric, thus the localization of the landmarks and the detection of the plan of the symmetry are challenging issues.

1.2. Methodology and Contributions

The main goal of this work is to propose a comprehensive and automated pipeline for efficient facial asymmetry evaluation. Using a 3D imaging sensor, several 3D sequences acquisitions are recorded before and after the clinician's intervention. As illustrated in Figure 1, a preprocessing step is then applied to each frame of the 3D videos including – holes filling, smoothing, pose normalization, and nose tip detection. In the Figure 1 and for each frame, each 3D original face is rotated to 180° with respect to the YZ-plane. As result, we obtain a new 3D face which represents the mirrored of the 3D original face (i.e the left half on the 3D original face corresponds to the right half on the mirrored face). The noses tip of the 3D original face and its mirror are well detected.

Then we used ICP algorithm to align the 3D original face and its mirror. At

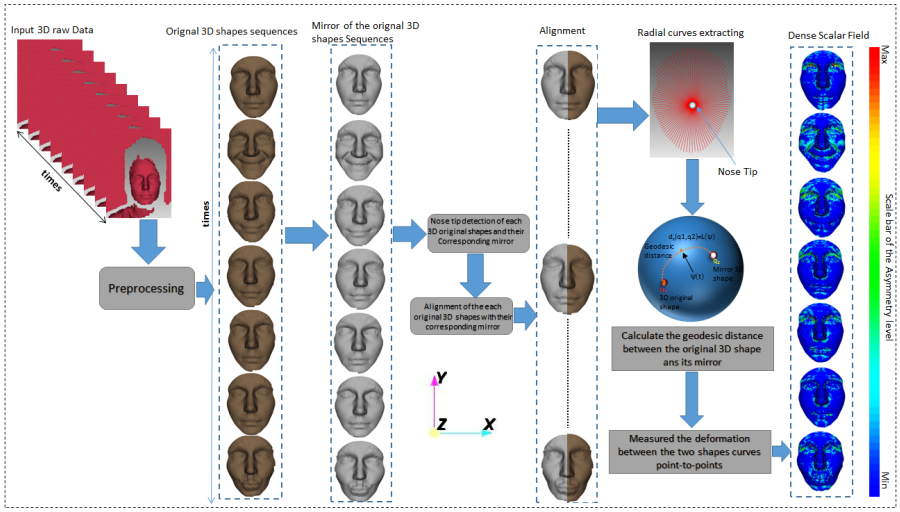

Figure 1: Overview of the method.

a fixed time t, shape analysis of the obtained faces and mirrored faces allows dense and accurate registration. In particular, using the initial velocity vector along the geodesic path connecting their shapes [9], one can quantify accurately the amount of differences separating their shapes. As a result, at each instant of the 3D sequence, Dense Scalar Fields which reflects the facial asymmetry is computed. Finally, the extension of the previous step to the temporal domain allows quantifying the evolution of facial asymmetry across the 3D sequence. An experimental illustration of the above-mentioned idea is reported where a new data set of the patient's face before and after BT injection is recorded.

In summary, the main contributions and novelties of this work are:

- We propose a new geometrical approach to quantify densely the dynamic evolution of 3D facial symmetry before and after the injection of BT.
- We propose the first dynamic 3D facial data set of patients with face paralysis and the first comprehensive protocol for analyzing the paralysis of 3D face. Moreover this will help the clinician to analyze point-to-point the face of the subject.

The rest of the paper is organized as follows. In section 2, we describe the clinical data used in this study, which consists of 3D facial sequences of patients recorded before and after the clinician's intervention, and the preprocessing step is designed. The Riemannian shape analysis approach and its capability to quantify the amount of facial asymmetry is presented in section 3. An experimental illustration involving the clinical data set is discussed, and the experimental results in section 4. Section 5 draw some concluding remarks and open some future investigations.

2. DATA SET DESCRIPTION AND PREPROCESSING

Due to the novelty of the topic, we have conducted with clinicians the acquisition of patients faces who present facial paralysis. The patients are asked to perform a set of specific facial expressions in front of the 3D camera. These expressions have been approved by the clinicians.

2.1. 4D facial data acquisition

In total we have collected facial sequences of 16 subjects with facial paralysis. Their ages vary from 14 to 66 years old. The subjects are invited for two acquisitions sessions, one time before the BT injection and a second time after two weeks of the injection. In addition, facial sequences of 11 healthy subjects have been also recorded. They will help to compare facial asymmetry results to partially paralyzed faces. The participants have been asked to convey 7 different expressions (1) normal smile, (2) forced smile, (3) close the eyes softly, (4) close the eyes very tight, (5) make a kiss, (6) raising-up the eyebrows, (7) and say this specific sentence in French: "Une jolie avenue bordée d'arbre et illuminer". According to the clinician, this sentence allows to observe several facial activities. The average duration of the recorded 3D sequences is 4-6 seconds with a total of 29 to 220 3D frames (meshes). In our acquisition process, we have used the structured light-based ARTEC MHT 3D scanner[1]. This single-

[1] http://www.artec3d.com/fr/hardware/artec-mht/

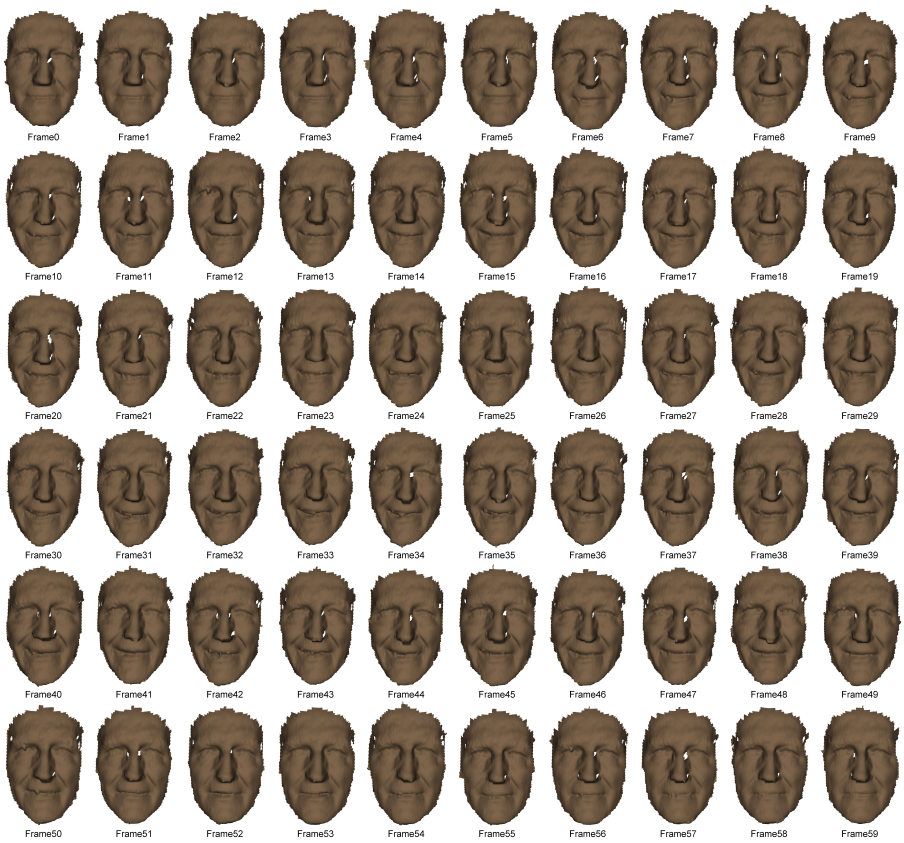

Figure 2: An example of a 3D sequence of a paralyzed face conveying a forced (maximum) smile. The sequence follows the temporal evolution neutral-onset-apex-offset-neutral.

view 3D scanner allows capturing 3D sequences of 15 fps (frame per second), the rate of 15 fps mentioned is the maximum temporal resolution provided by the 3D ARTEC scanner. This rate can vary slightly through the sequences. Each frame consists of about 5,500 vertices, and the 3D resolution is 0.5 mm. To allow an optimal acquisition, the participants are asked to sit in front of the scanner at a distance of about 80 cm and continuously focus on a fixed point during the acquisition see Figure 3. Figure 2 illustrates an example of 3D facial sequences acquired using the 3D sensor of a patient's face asked to exhibit a forced smile.

In the acquisition of 3D face sequences, many patients have been asked to repeat

the same expression 2 or 5 times in front of the 3D scanner, this is due to the difficulty of the patients to convey these expressions correctly. Moreover, most of the young patients are shy to realize these expressions in front of the 3D camera. It appears for smiling and big smiling expression, the clinicians use a funny video that make them smile in order to make them comfortable to convey the asked expression.

2.2. Preprocessing of 3D frames

The 3D sequences presents imperfections such as missing data (i.e holes in the mesh), noise (i.e. spikes) and contain undesirable parts (neck, clothes, hair, etc.) for our analysis, see Figure 3 . Thus, we have applied the following pipeline (see the algorithm 1 part 1) in order to extract the informative part of the face,

- Clean up the meshes by removing identical vertices,
- Fill small holes on the mesh by interpolation,
- Detect the nose tip of each face on each 3D frame,
- Crop the face, and the nose tip is used to define a sphere with a fix radius (r = 80), which serves to crop the face region,
- Smooth the meshes using a Laplacian filter to reduce the noise,

Now, in order to allow quantifying the bilateral symmetry of the face, we added the following additional steps,

- Apply a mirror operator on the 3D faces with respect to the YZ-plan,
- Apply the ICP (Iterative Closest Point) algorithm between the original 3D frame and the mirrored 3D frame for pose normalization.

An efficient implementation of this pipeline is designed using existing filters available in the VTK[2] library. This preprocessing pipeline is applied to each 3D frame of the sequence separately. The next step is to extract radial curves and to quantify the amount of asymmetry by comparing each 3D shapes to their

[2] `http://www.vtk.org`

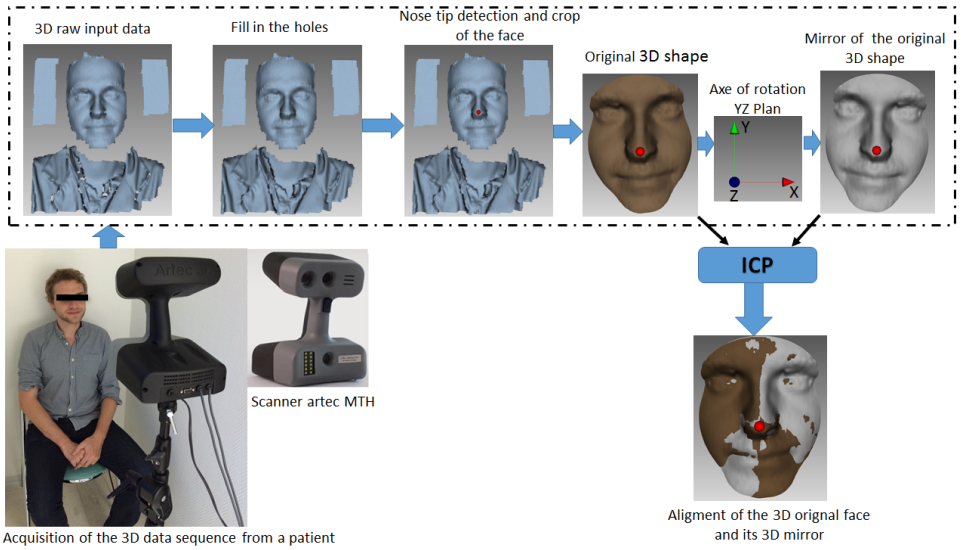

Figure 3: Preprocessing step, a red sphere is centered in the nose tip.

reflections using our Riemannian approach [9]. Thus, the problem is turned to a comparison of 3D (static) facial shapes, i.e. at each instant t, we preprocessed the 3D face with its mirrored. To achieve such comparison, one needs for an accurate registration algorithm which can handle the non-rigid deformations (bending, stretching/compression) of the facial surfaces. The registration of vertices across 3D faces is an important ingredient when studying their shapes (we assume here an invariance to rigid motions – translation, scaling and rotation). Specifically, in comparing shapes of faces, it is important that similar biological parts are registered to each other, in particular the left and right halves of the face, when studying the face asymmetry. Several methods have been proposed in the literature as discussed above such as the Non-rigid ICP algorithm [18], the Free Form Deformation (FFD) algorithm [8] and the Thin-plate Spline (TPS) algorithm [19]. Most of these solutions try to find an optimal registration between two 3D faces, however, their cost functions which minimize the distance between 3D meshes is not a proper metric; it is not even symmetric. That is,

the optimal registration of a 3D face F^1 to another 3D face F^2 may not be the same as the registration of F^2 to F^1. This makes difficult to interpret the results (i.e. quantifying the divergence between the compared 3D faces). The Riemannian framework used in [9] for 4D facial expression analysis grounding on elastic radial curves (i.e. an elastic metric is used [20]) provides a nice physical interpretation of measuring the facial deformations between curves using a combination of stretching and bending. These elastic deformations are captured by the Dense Scalar Field (DSFs) features. Hence, the main motivation of using a Riemannian approach is to perform registration that matches corresponding anatomical features, and obtain deformation fields that are physically interpretable. In the next section, we recall essential materials to compute the Dense Scalar Fields (DSFs) and illustrates their used to quantify the amount of facial asymmetry. Finally, we will make their extensions to study dynamic 3D faces.

3. BILATERAL FACIAL ASYMMETRY

Given a 3D facial sequence, our aim is to measure the amount of bilateral symmetry at each time and extend it to the temporal dimension. We will start by the first case (i.e. static 3D face), the most common way to quantify the asymmetry is to first detect the plane of symmetry of the face (called also the mirror plane), then apply registration algorithms as the well-know Iterative Closest Point as proposed in [16]. Here, the most problematic step is to detect accurately the mirror plane before comparing the shape divergence between the two halves of the facial surface. Quan et al. [16] consider the 3D vertices of the face as a cloud of points in $\mathbb{R}^3$, they apply a Principal Component Analysis (PCA) to extract its dominant direction, which corresponds to the mirror plane. When this approach achieved good results on healthy faces, it is problematic to be used on faces with significant paralysis. That is, in the latter, the plane location may be far from the the nose tip, for example. A second problem of this approach, is the use of the ICP algorithm, which considers its cost function

only rigid transformations (translation, scaling and rotation) and a very basic optimization criteria based on the point-to-point (or point-to-plane) distance. This makes difficult to achieve any physical interpretation of the results. In the case of facial paralysis, the registration algorithm may register different anatomical parts on the two halves of the 3D face.

According to the above-mentioned issues, we present in this paper a different approach to measure the bilateral facial asymmetry, when first avoiding the mirror plane detection, and use a Riemannian shape analysis framework developed recently in [9] to achieve accurate registration. The basic idea is to approximate the facial surface by an indexed (ordered) collection of radial curves, then consider pairwise curves comparison, using a Riemannian elastic metric which account for the stretching and shrinking of the face. Before reaching this step, we avoid the step of mirror plane detection by applying a simple mirroring to the original face according to the YZ-plane, then align the two surfaces using the ICP algorithm, and finally use the Riemannian approach to quantify the optimal deformation between the 3D face and the mirrored face.

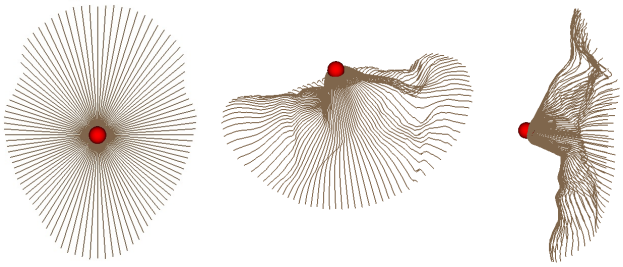

Figure 4: Three views of the same collection of curves extracted to approximate a 3D facial shape. The point in red represent the detected nose tip (origin of the radial curves).

Let's first note the facial curves emanating from the nose tip and tacking equally-spaced directions dictated by an angle $\alpha \in [0, 2\pi[$, which is the correct notation to cover the whole face with the collection of radial curves. The obtained collection of curves provides in Figure 4, are not only an approximation of the face, but also impose on it a parametrization. That is, considering the nose tip as an origin, each point on the face is identified in a unique way by α

the index of the curve, and k the index of the point on the curve. Based on this radial curves representation, the problem of face-to-face registration is turned to a pairwise-curves registration. Elastic shape analysis of open and closed curves is a well-explored problem. An emerging solution using Riemannian geometry have demonstrated convincing results, when the problems of alignment and registration are leaded in the same and unique step.

Recently, the geometric framework of curves proposed by [20] has been extended to 4D facial expression recognition [9]. They analyzed the facial shapes along a sequence in order to extract the deformations, then learn them using HMM or Random Forest for the purpose of expression classification. In particular, using the square-root velocity function (SRVF) representation $q^{\alpha}(t)$ of the facial curves $\beta^{\alpha}(t)$, where α is the curve index on the face, to form the shape space $\mathcal{S}^{[0,2\pi[}$. As the shape space is a quotient space of an hypersphere of the Hilbert space $L^2([0,1],\mathbb{R}^3)$, the geodesic path between two facial shapes is given by the minor arc of the great circle of the hypersphere. A facial surface is represented by a collection of radial curves and a Riemannian framework is used to study shapes of these curves. We start by representing facial curves as absolutely continuous maps from $\beta : [0,1] \to \mathbb{R}^3$ and our goal is to analyze shapes represented by these maps. The problem in studying shapes using these maps directly is that they change with re-parameterizations of curves. If γ is a re-parameterization function (typically a diffeomorphism from $[0,1]$ to itself), then under the standard $\mathbb{L}^2$ norm, the quantity $\|\beta_1 - \beta_2\| \neq \|\beta_1 \circ \gamma - \beta_2 \circ \gamma\|$, which is problematic. The solution comes from choosing a Riemannian metric under which this inequality becomes equality and the ensuing analysis simplifies. As described in [20], we represent the facial curves using a new function q(see Eq. (1)). The advantage of using SRVF representation is that under this representation the elastic metric becomes the standard $\mathbb{L}^2$ metric and an identical re-parameterization of curves preserves the $\mathbb{L}^2$ norm of between their SRVFs. The mapping from a curve β to q is a bijection (up to a translation) and the space of all SRVFs is the Hilbert space of all square-integrable maps of the type $q : [0,1] \to \mathbb{R}^3$. This space under the natural $\mathbb{L}^2$ inner product is actually a

vector space and geodesics between points in this space are straight lines.

With the proposed representation, a 3D facial surface is approximated by an indexed collection of radial curves β_α, where the index α denotes the angle formed by the curve with respect to a *reference* radial curve. In particular, the reference radial curve (i.e., the curve with $\alpha = 0$) is chosen as oriented along the vertical axis, while the other radial curves are separated each other by a fixed angle and are ordered in a clockwise manner.

Considering an arbitrary radial curve β of the face, it can be parameterized as β: $I \rightarrow \mathbb{R}^3$, with $I = [0, 1]$, and mathematically represented through the *square-root velocity function* (SRVF) [21, 20], denoted by $q(t)$, according to:

$$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}, \;\; t \in [0, 1]. \tag{1}$$

This specific representation has the advantage of capturing the shape of the curve and makes the calculus simpler. Let us define the space of the SRVFs as $\mathcal{C} = \{q : I \rightarrow \mathbb{R}^3, \|q\| = 1\} \subset \mathbb{L}^2(I, \mathbb{R}^3)$, with $\|\cdot\|$ indicating the $\mathbb{L}^2$ norm. With the $\mathbb{L}^2$ metric on its tangent space, $\mathcal{C}$ becomes a Riemannian manifold. Basically, with this parametrization each radial curve is represented on the manifold $\mathcal{C}$ by its SRVF. According to this, given the SRVFs q_1 and q_2 of two radial curves, the shortest path ψ^* on the manifold $\mathcal{C}$ between q_1 and q_2 (called *geodesic path*) is a critical point of the following energy function:

$$E(\psi) = \frac{1}{2} \int ||\dot{\psi}(\tau)||^2 d\tau \;, \tag{2}$$

where ψ denotes a path on the manifold $\mathcal{C}$ between q_1 and q_2, τ is the parameter for traveling along the path ψ, $\dot{\psi} \in T_\psi(\mathcal{C})$ is the tangent vector field on the curve $\psi \in \mathcal{C}$, and $||.||$ denotes the $\mathbb{L}^2$ norm on the tangent space.

Since elements of $\mathcal{C}$ have a unit $\mathbb{L}^2$ norm, $\mathcal{C}$ is a hypersphere in the Hilbert space $\mathbb{L}^2(I, \mathbb{R}^3)$. As a consequence, the geodesic path between any two points $q_1, q_2 \in \mathcal{C}$ is simply given by the minor arc of the great circle connecting them on this hypersphere, $\psi^* : [0, 1] \rightarrow \mathcal{C}$. This is given by:

$$\psi^*(\tau) = \frac{1}{\sin(\theta)} \left(\sin((1 - \tau)\theta)q_1 + \sin(\theta\tau)q_2\right) \;, \tag{3}$$

where $\theta = d_{\mathcal{C}}(q_1, q_2) = cos^{-1}(\langle q_1, q_2 \rangle)$. We point out that $sin(\theta) = 0$, if the distance between the two curves is zero, in other words $q_1 = q_2$. In this case, for each τ, $\psi^*(\tau) = q_1 = q_2$.

The tangent vector field on this geodesic is then written as $\frac{d\psi^*}{d\tau} : [0,1] \rightarrow T_\psi(\mathcal{C})$, and is obtained by the following equation:

$$\frac{d\psi^*}{d\tau} = \frac{-\theta}{\sin(\theta)} \left(\cos((1-\tau)\theta)q_1 - \cos(\theta\tau)q_2\right) \ . \tag{4}$$

Knowing that on geodesic path, the covariant derivative of its tangent vector field is equal to 0, $\frac{d\psi^*}{d\tau}$ is parallel along the geodesic ψ^* and one can represent it with $\frac{d\psi^*}{d\tau}|_{\tau=0}$ without any loss of information. Accordingly, Eq. (4) becomes:

$$\frac{d\psi^*}{d\tau}|_{\tau=0} = \frac{\theta}{\sin(\theta)} \left(q_2 - \cos(\theta)q_1\right) \ (\theta \neq 0). \tag{5}$$

Based on the above representation, we define a *Dense Scalar Field* capable to capture deformations between two corresponding radial curves β^1_α and β^2_α of two 3D faces approximated by a collection of radial curves.

Now, let q and q^r the SRVFs of the curves β and its reflection β^r, respectively, and ψ^* is the optimal path (the geodesic) connecting q and q^r and the angle θ is the length of this geodesic which represents the geodesic distance between q and q^r. Integrating θ over all $\alpha \in [0, 2\pi[$ gives rise to a scalar value which measures the divergence between each facial curve and its mirror. We will use this measure later on in section 4 in our quantitative analysis. To quantify densely the divergence between faces across a 3D sequence, Ben Amor et al. [9] proposed the Dense Scalar Fields. Roughly speaking, it consists to have computed the initial velocity vector $\frac{d\psi^*}{d\tau}|_{\tau=0}$ along the geodesic ψ^*, as the geodesic is by definition a constant speed curve on the shape manifold in Eq. (5).

As described above, the Dense Scalar Fields (between a curve and its mirror) denotes the magnitude of $\frac{d\psi^*}{d\tau}|_{\tau=0}$ at each point of the curves. In 1(part 2), we provided the algorithm of the DSF.

Algorithm 1 – Facial Paralysis of 3D Face Sequences based on Shape Analysis of 3D Face **(Part 1 - Preprocessing Step)**

Ensure: A 3D video set is giving, and nf the number of frame f in the 3D video sequence.

Require: Method of NoseTip-Detection, Crop-Face-From-NoseTip, Align-Frame, and the methods of Fill-hole and smoothing in a 3D mesh are available in [3]

procedure PREPROCESSING-STEP(Input a raw 3D video sequence)
  for $k \leftarrow 0, nf$ **do**
    $f_1(k) \leftarrow$ Fill-hole($f_{(k)}$)
    $f_2(k) \leftarrow$ NoseTip-Detection($f_1(k)$)
    $f_3(k) \leftarrow$ Crop-Face-From-NoseTip($f_2(k)$)
    $f_4(k) \leftarrow$ Smooth($f_3(k)$)
    $f_5(k) \leftarrow$ Align-Frame($f_{(0)}$,$f_4(k)$)
  end for
  It is very important that all the frame is aligned to the first one. The nose tip is used as the reference point of the alignment.
end procedure

Require: Preprocessing-Step

procedure MIRROR CALCULATION(3D video from preprocessing-step)
  for $l \leftarrow 0, nf$ **do**
    $f_6(l) \leftarrow$ RotateYZ180($f_5(l)$) see 3 for the axis orientation.
    $f_7(l) \leftarrow$ ICP ($f_6(l)$, $f_5(l)$))
  end for
  Let be the 3D original face $f_5(l)$ as F^o, and the mirror of the 3D original face $f_6(l)$ as F^r, then $f_6(l)$ represents the ICP result (pose normalization) of the 3D original face and its mirror)
end procedure

Figure 5 illustrates this idea, on the left the original static 3D face, in the middle its mirror and on the right the DSFs mapped into the mirrored face. In

Algorithm 1 – Facial Paralysis of 3D Face Sequences based on Shape Analysis of 3D Face **(Part 2 - Dense Scalar Field)**

Require: 3D original face F^o and mirror of the original 3D face F^r; T, number of sample points on a curve; $\Delta\alpha$, angle between successive radial curves; $|\Lambda|$, number of curves per face

Ensure: $DSF(F^o, F^r)$, the DSF between the two faces

procedure ComputeDSF($F^o, F^r, T, \Delta\alpha, |\Lambda|$)
  $n \leftarrow 0$
  while $n < |\Lambda|$ **do**
    $\alpha = n \cdot \Delta\alpha$
    for $i \leftarrow o, r$ **do**
      extract the curve β^i_α
      compute the SRVF of β^i_α:
        $$q^i_\alpha(t) \doteq \frac{\dot{\beta}^i_\alpha(t)}{\sqrt{\|\dot{\beta}^i_\alpha(t)\|}} \in \mathcal{C},\ t = 1, \ldots, T$$
    end for
    compute the distance between q^o_α and q^r_α:
      $$\theta = d_{\mathcal{C}}(q^o_\alpha, q^r_\alpha) = cos^{-1}(\langle q^o_\alpha, q^r_\alpha \rangle)$$
    compute the deformation vector $\frac{d\psi^*}{d\tau}|_{\tau=0}$ using
    Eq. (5) as:
      $f(q^o_\alpha, q^r_\alpha) = (x_\alpha(o), x_\alpha(r), \ldots, x_\alpha(T)) \in \mathbb{R}^T_+$
      $x_\alpha(t) = |\frac{\theta}{\sin(\theta)}\,(q^r_\alpha - \cos(\theta) q^o_\alpha)\,|,\ t = 1, \ldots, T$
  end while
  compute $DSF(F^o, F^r)$ as the magnitude
  of $\frac{d\psi^*}{d\tau}|_{\tau=0}(k)$:
    $DSF(F^o, F^r) = (f(q^o_0, q^r_0), \ldots, f(q^o_{|\Lambda|}, q^r_{|\Lambda|}))$ **return** DSF
end procedure

this figure, the warmer the color the higher is the DSF value. Consequently, cold colors reflect the most symmetric regions of the 3D face, where warm colors highlight the most asymmetric areas. Compared to the 3D landmarks-based approach adopted in [15], our approach provides dense measurements of the

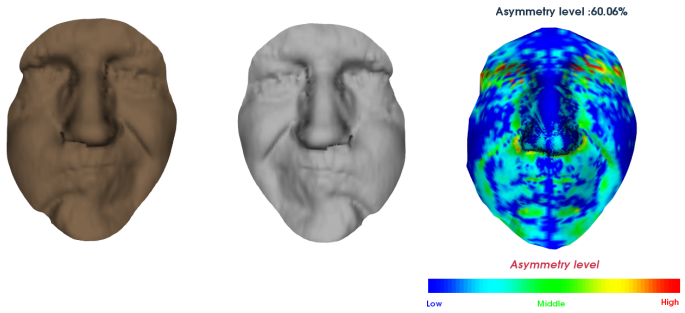

Figure 5: An example of computing the Dense Scalar Fields (DSFs) on a patient's face with paralysis. From left to right (a) original 3D face, (2) 3D face after reflection, and (3) the DSFs shown as a color map mapped on the original 3D face.

bilateral asymmetry on the face. This could be of high importance for clinicians to evaluate locally the bilateral asymmetry of the face. This is because the BT injection is done locally in some specific muscles. The main motivation behind computing the shape deviation between an arbitrary 3D facial shape and its mirrored shape is to avoid the difficult task of mirror plane detection. In fact, our approach allows to quantify accurately the bilateral face asymmetry by (1) a fine alignment of the two shapes using the ICP algorithm, then (2) compute the shape deviation using elastic radial curves (i.e. DSFs). Thus, our approach is independent of any plane of symmetry. In particular, this is suitable when working with paralyzed faces when conveying expressions for whom the detection of a plane of symmetry is not an easy task.

4. EXPERIMENTAL RESULTS

In this section, we illustrate experimentally the use of DSFs to quantify the bilateral symmetry of the face. To this end, we consider a subset of samples taken from the new data set described in section 2. To conduct these experiments, given a 3D face sequence, each frame is first mirrored and then aligned to the original face. As a final step, the DSFs are calculated between the original and the mirrored (aligned) shapes.

4.1. Experiments and discussions

In order to observe and to analyze the difficulty level of the patients to move the facial muscle, to examine the hyperactivity of the healthy parts and to analyze which part of the face that the BT should be injected in order to reduce the bilateral asymmetry of the face, the patients have been asked to perform the 8 facial expressions in pre-and post-operative of the patient's face: (1) Neural face; (2) Smiling faces (natural); (3) Forced smiling ; (4) Close eyes softly; (5) Close eyes very tight; (6) Make a kiss; (7) Eyebrows raised; (8) repeat a specific sentence ("une jolie avenue bordée d'arbre et illuminer"). In the following, we also report quantitative and qualitative results on two control (healthy) samples faces. In these experiments, the facial surfaces are approximated with a total of 100 radial curves emanating from the nose tip, each curve consists of 50 points, thus, in total each DSF descriptor consists of 5000 scalars (i.e for each 3D frame).

The proposed 4D imaging method was evaluated on the whole data set with the helps of the clinicians. As result, the DSF curves for each patient before BT injection were always above of the controls samples. Some patients still had a higher asymmetry level overall the controls samples after BT injection and others had a lower asymmetry next to the one of the control samples. Our tool is able to distinguish a symmetrical face (control sample) from an asymmetrical face (patients before TB).

Then, we compare the results before and after TB in patients. The 4D tool measured a change in overall facial asymmetry after BT injection for all patients. We compare the asymmetry level for each patient before and after BT. Thus, for 11 patients the facial asymmetry was decreased, and for 5 patients it was increased. In the following, we will explain the results on each of the eight facial expressions:

1. **Neutral State** – in this experiment, we simply restrict our study to a static neutral face. Thus, a single 3D surface is considered for the patient's face before and after the BT injection. Figure 6 illustrates the DSFs computed to reveal densely the level of bilateral symmetry from the 3D face. One can

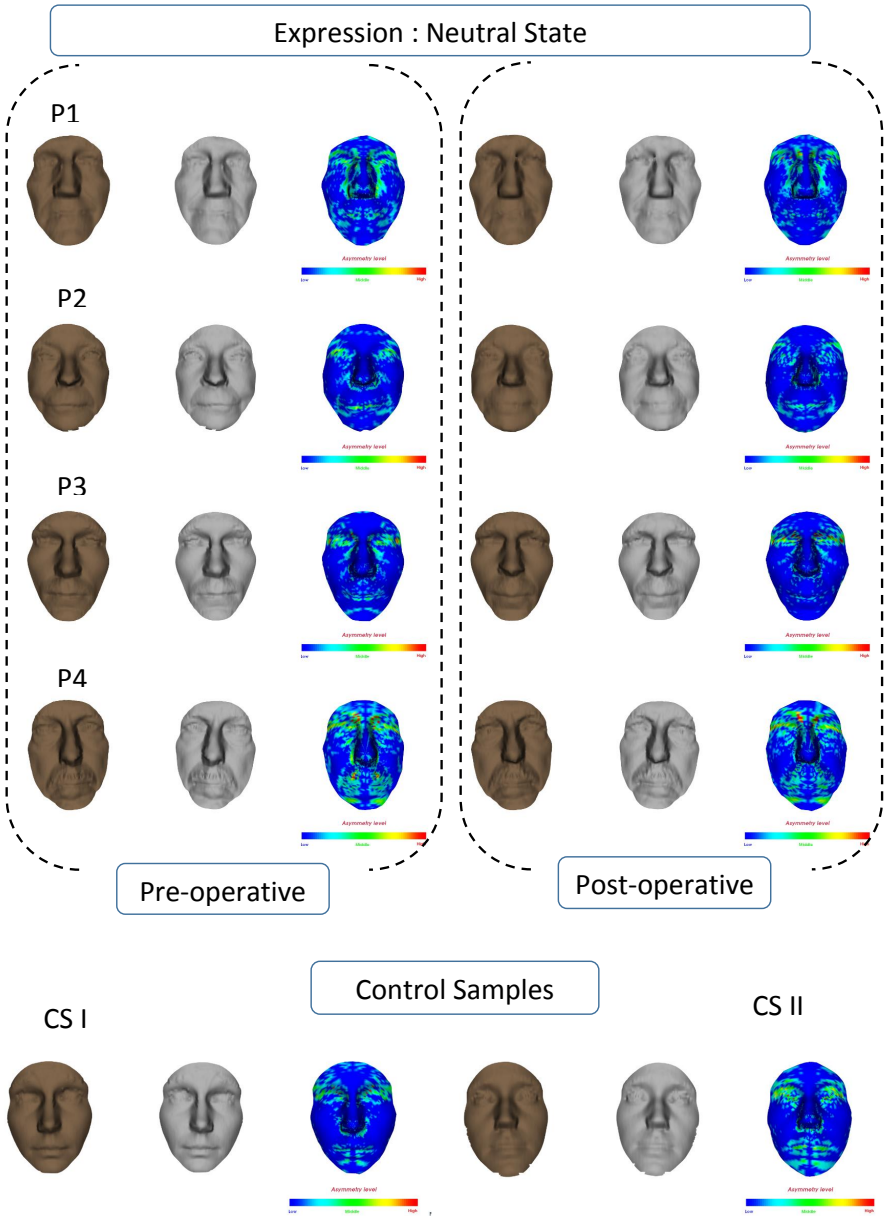

Figure 6: Amount of facial asymmetry of four patients (P1 to P4) before and after the BT injection. For each case, left : original 3D face, middle : mirrored face, and right : DSFs mapped on the face. The last row provides two examples of healthy faces for comparison.

note from this figure an improvement in terms of face symmetry, after the BT injection even when the face is in neutral state. When comparing with the two healthy faces (last row of the same figure), the latter are not perfectly symmetry. Hence, the comparison between paralyzed faces (before and after BT injection) to healthy faces is quite difficult, when consider static neutral faces. This was one of our motivations to extend our study to dynamic sequences. Let see the temporal evolution in the face with simple smiling.

2. Natural Smile – in this experiment, the patients are asked to perform a natural smile. As in the previous experiment, we report in Figure 7 the bilateral symmetry for the apex (maximum expression) of the smile. Compared to the previous experiment (Figure 6) the DSFs reveal more differences, in terms of asymmetry, between the pre-operative and the post-operative result. That is, a clear improvement after the BT injection could be seen, for all the patients faces. In particular, the amount of asymmetry is reduced along the zygomatic muscle, area of the BT injection. We report in Figure 8, at each frame of the video a scalar which represents the distance between the 3D face and its mirror. This distance represents the length of the geodesic (shortest) path which connects the facial shape and its mirror represented in the same shape space. We have selected four patients (Figure 8, P1 to P4) to illustrate the dynamics of the bilateral symmetry over the time, before and after the BT injection. We have added to these graphs the distances for two control (healthy) patients to allow comparison. We should note first the differences in sequences lengths and the temporal non-alignment of the four sequences (for the same graph).

It is clear from these graphs that for the smiling expression marked an improvement in terms of facial behavioral symmetry, is quantified (red and blue curves). In general, the curves related to the healthy faces are always under those of the paralyzed faces. It can be seen in sub-figure (P4) a huge amount of asymmetry before the BT injection which disappears after the injection.

The results presented in this figure confirms the conclusions made before (for a static shape) and extends them to the temporal domain. They showed clearly the importance of considering the temporal evolution of the facial 3D shape for

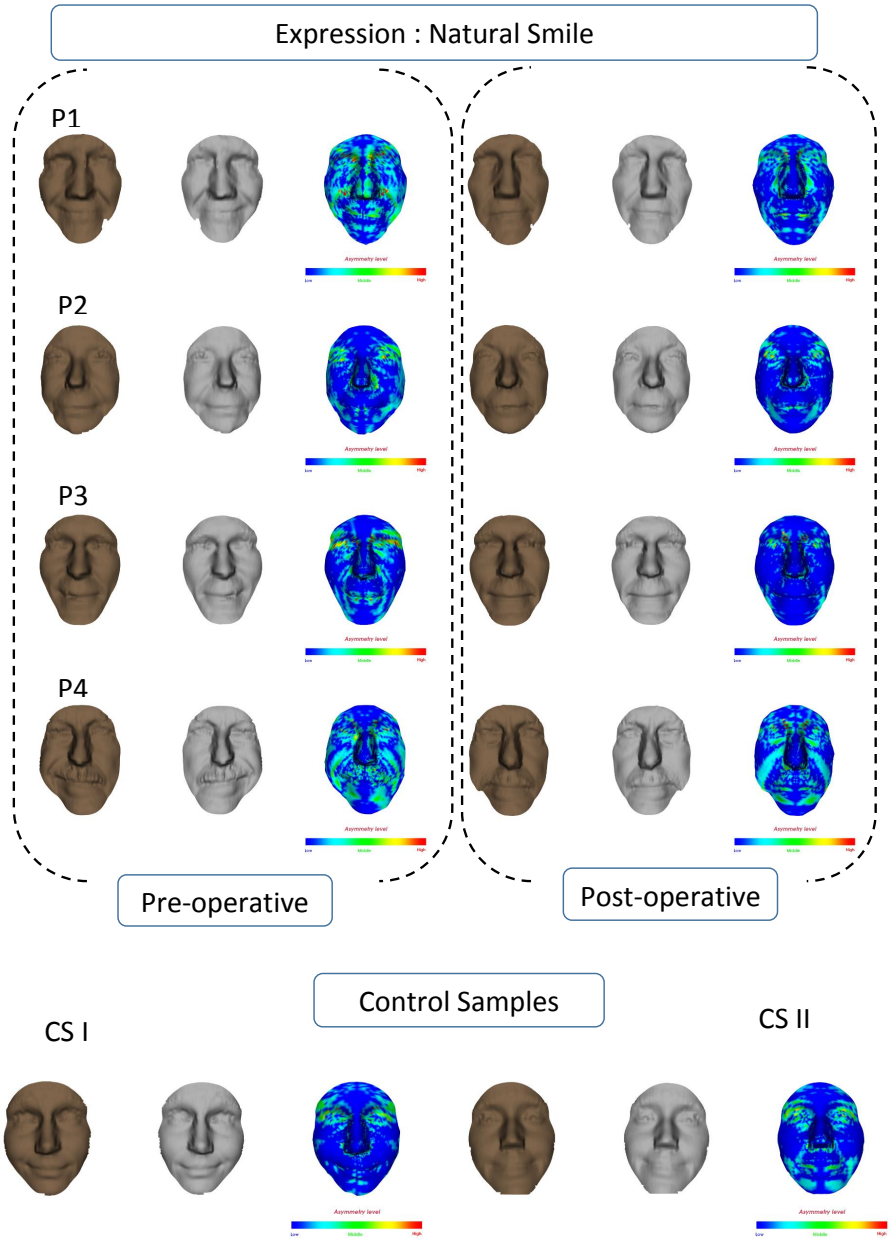

Figure 7: Natural smiling experiment from patients (P1 to P4) in pre-operative, and post-operative.

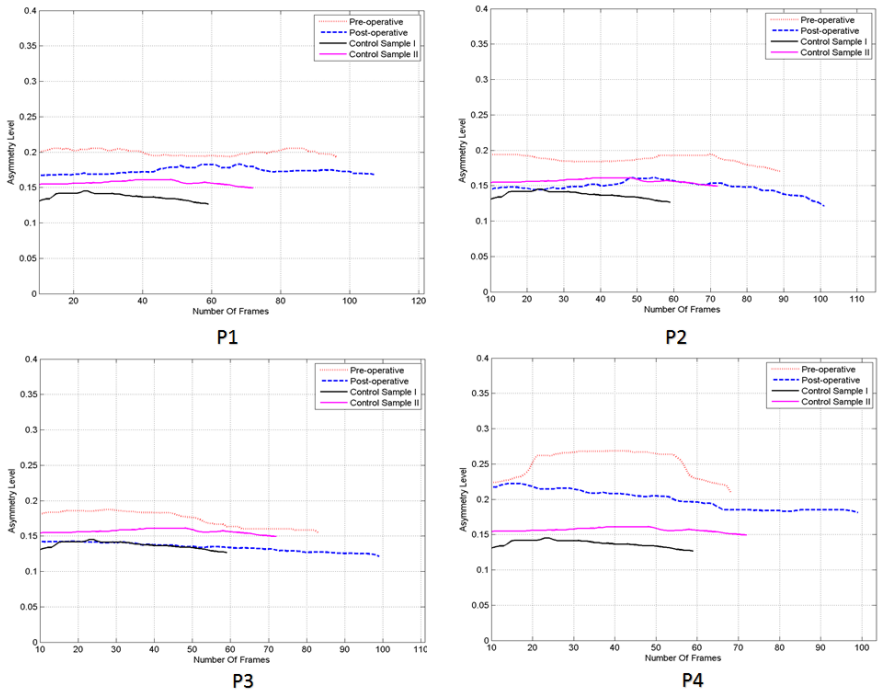

Figure 8: Degree of asymmetry across the video for the natural smiling experiment.

the clinician. When comparing the post-operative results with the healthy faces (curves in hyphen blue, the black and magenta) in Figure 8, one can note the efficiency of the BT injection to symmetrizing the facial shape and its temporal evolution. However, what happen if the expression is more prominent than a natural smile?

3. Forced Smiling – in response of the previous question, the patients are asked to convey a forced (maximum) smile. Then the same pipeline is applied as previously. From Figure 9, one can establish the same conclusions than in the previous experiment. That is, the degree of asymmetry of static and dynamic faces decreases after the BT injection. It can be seen from this graph in Figure 10 that during the time-interval to the end, when the expression happened the amount of bilateral asymmetry before the injection is higher (red dot curves) than after. The blue curve (after the BT injection) is always higher than the black and the magenta (healthy subjects), but the patient P3 has a very low

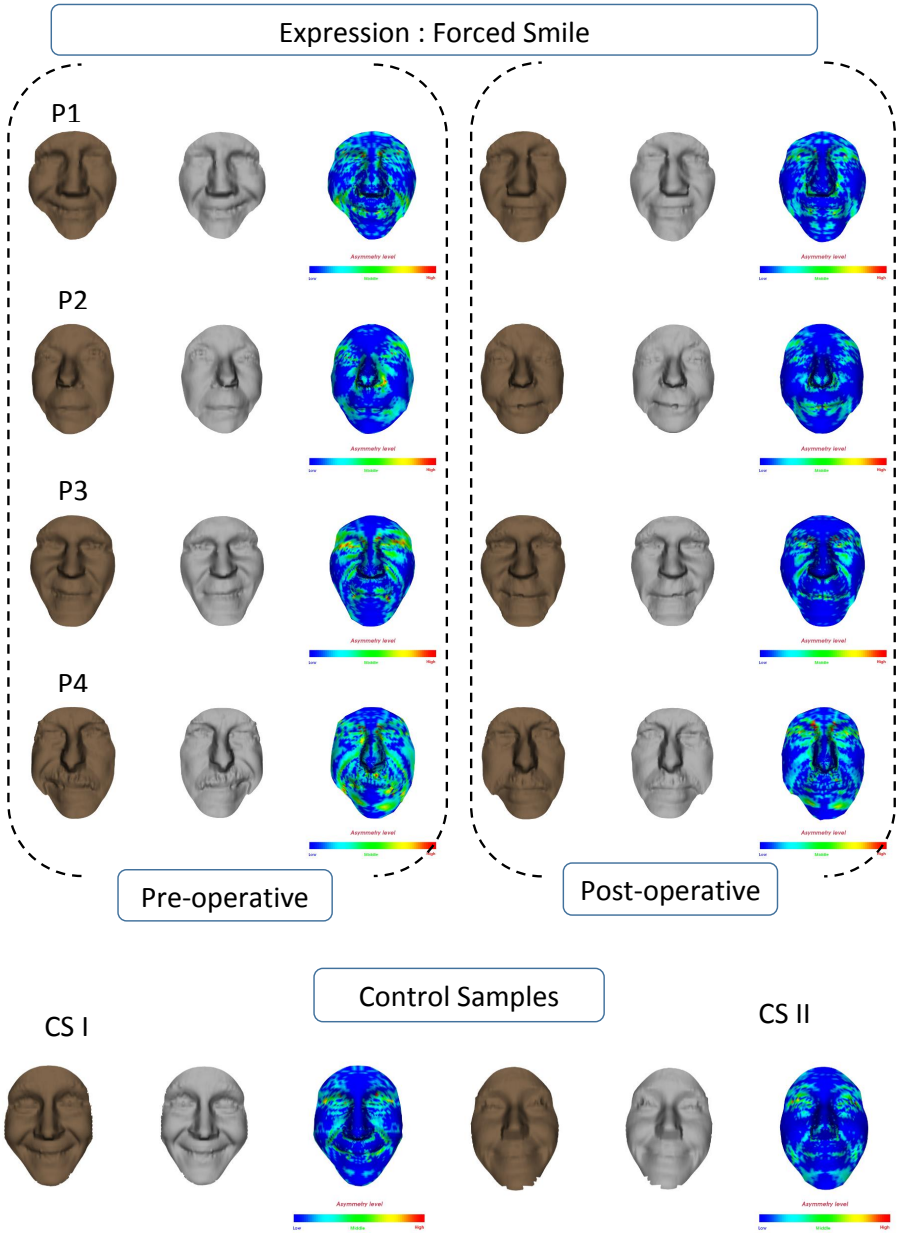

Figure 9: Forced (maximum) smiling experiment.

asymmetry level that is better than the control sample (SC II) in magenta. This expression is done twice with the patient in (P4) due to the difficulty to make

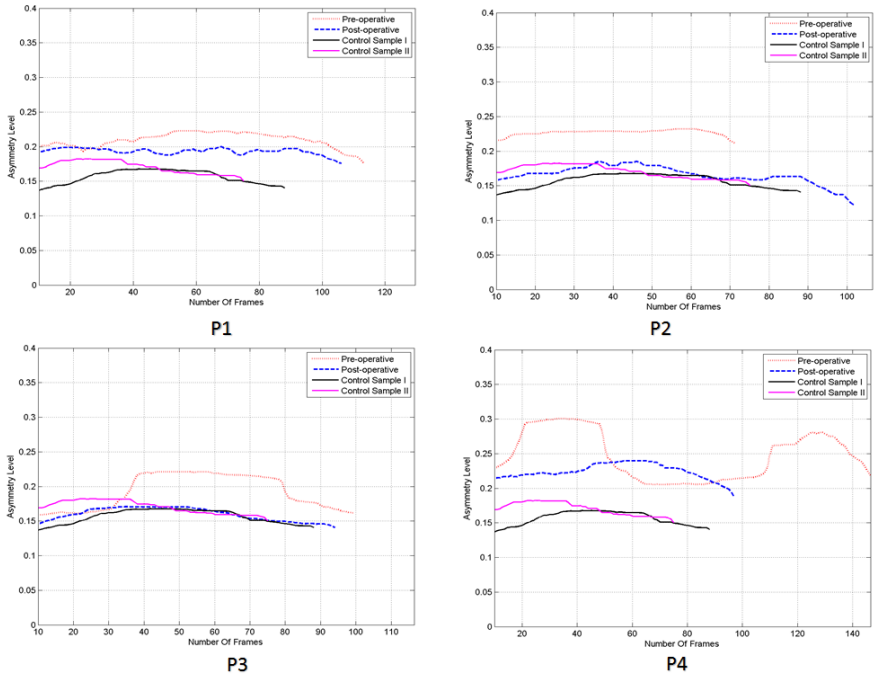

Figure 10: Forced smiling in the experiment.

a force smiling.

4. Close the Eyes Softly– To continue with the experiments, the clinician invited the patient to close the eyes Softly. Thus, this expression allowed the clinician to observe firstly the difficulty level of the muscles of the eyes to close, and secondly to examine if other muscle is participated in the expression (see Figure 11, P1 to P4). However, for a healthy face, only the muscle of eyes is reacted (see Figure 11, SC I and SC II) . It is clear in this experiment that the patients do not receive any BT injection in the muscles of the eyes. In fact, the bilateral asymmetry in pre-operative and post-operative case should be the same, but in Figures 11 and 12, specially for the patients P3, we observe a very low bilateral asymmetry in the post-operative case. This can be explained by the fact, that the healthy part of the face of this patient is well participated in this expression in the preoperative. As the hyperactivity of the healthy part is

reduced in post-operative, this allows less deformation of the face. For the others patients in Figure 12, the asymmetry level in the post-operative (blue hyphen curve) is not lower than the pre-operative (red dot curve) and the asymmetry level of the patient P3 is almost equal to the control sample II. One can note that, the amount of the bilateral asymmetry of the two control samples is still very low during the whole expression. Thus, what happen if this expression is more tight ?

5. Close the Eyes very Tight– To continue to analyze the deformation of the paralyze face, the clinician asked the patient to closed the eyes very tight. In Figure 13, we can observe that the healthy part of the face is also participated in the expressions while the sickness part does not move. By looking at the Figure 13, one can also observe the difficulty level of some patients to achieve this expression. We can establish the same conclusions than in the previous experiment (close the Eyes Softly), but the amount of the asymmetry level is reduced in the post-operative case.

However, in Figure 14, the asymmetry level in the pre-operative case (red dot curves) is higher than the post operative case (blue hyphen curves) for patients P1, P3, P4. Nevertheless, for the patients P2, and during the frame interval 40-47, the asymmetry level of the post-operative case is little higher than the pre-operative because the difficulty level of this patient to do this expression is very lower out of BT injection. In Figure 14 the amount of bilateral asymmetry level is very low for the control sample I and II, and the control sample I has the lowest asymmetry level (black curve).

6. Make a Kiss – In this part, the patients are asked to convey a kiss expression. This expression help the clinician to observe the auricular muscle of the mouth and the lips position. In Figure 15 it is clear that the amount of the bilateral asymmetry in pre-operative is better than the post-operative case and the BT injection reduced considerably the amount of the bilateral asymmetry level. Moreover, all the graphs in Figure 16 confirm that the face of the all patients become more symmetry after the BT injection. The patient P3 has a low asymmetry level which can be compare with the asymmetry level of the

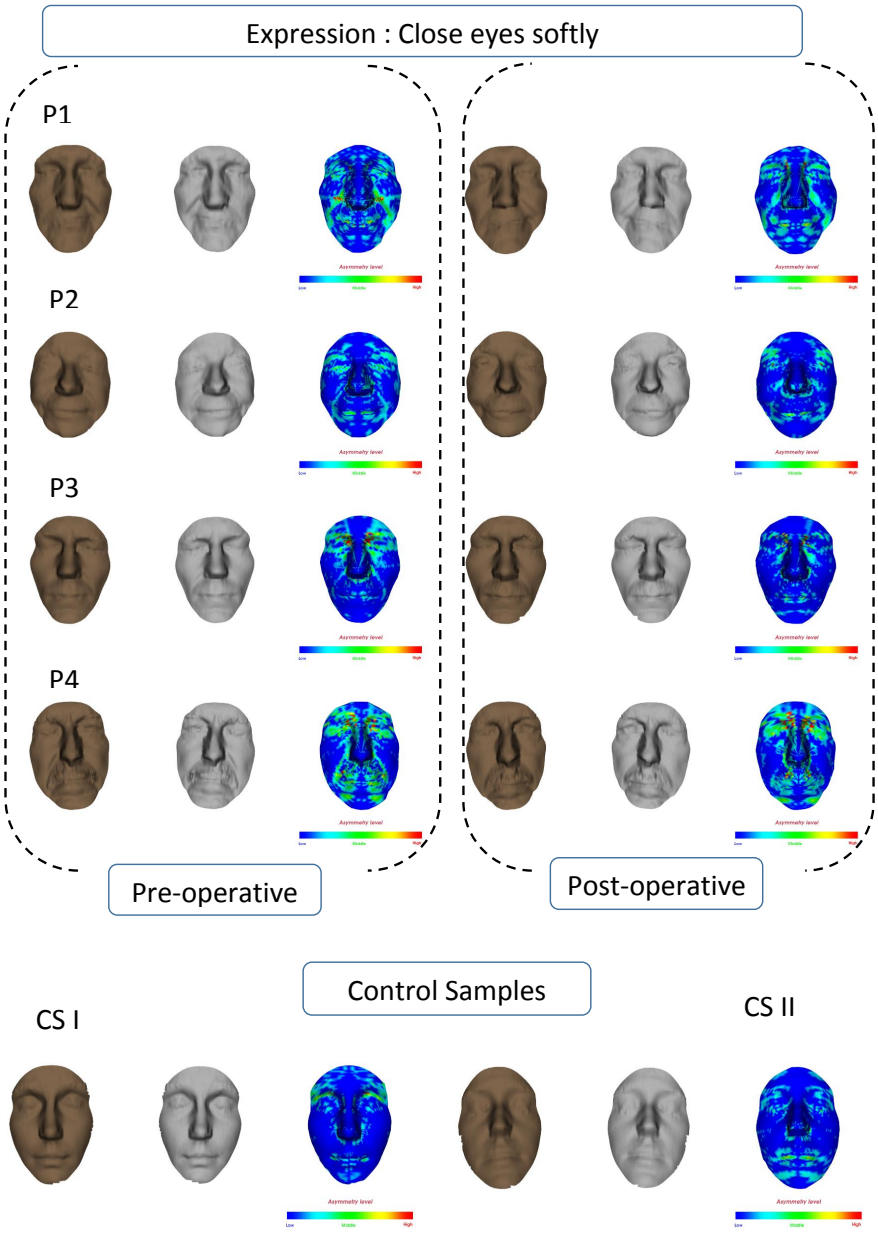

Figure 11: Closed Eyes experiment before and after the BT injection in the zygomatic muscle.

control sample I, the amount of the bilateral asymmetry of the control II is the lowest. In Figure 16 specially the patient P4, the difficulty to do this expression

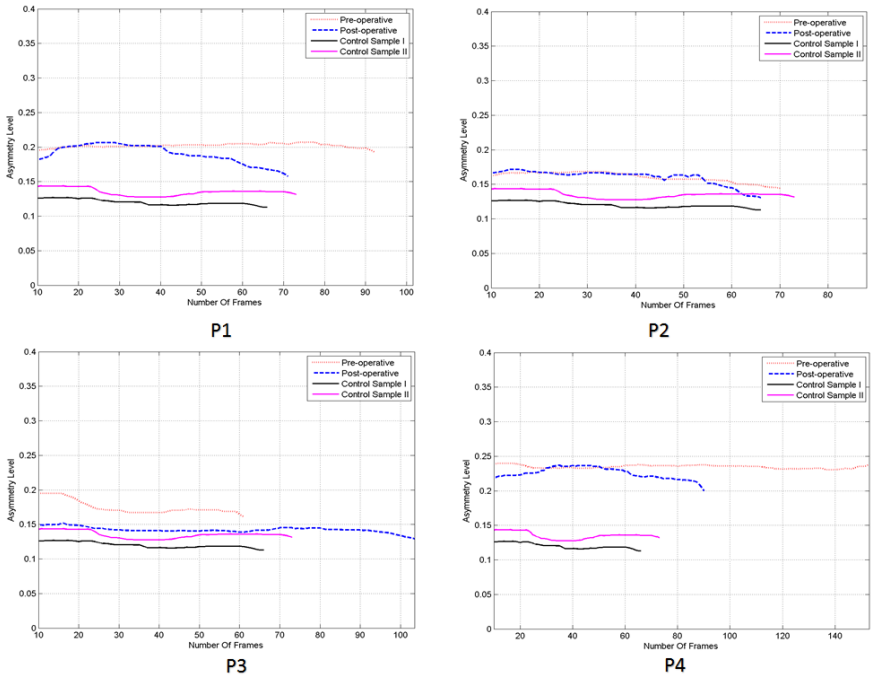

Figure 12: Temporal evolution of the expression close eyes softly in the experiment.

is more complicated in the preoperative case.

7. Raising-up the Eyebrows – in this experiment, the clinician asked to the patients to raised up the eyebrows. From figure 17, one can note again that the amounts of the asymmetry before and after the BT injection are comparable. This is mainly due to the fact that the BT injection is concentrated only on the zygomaticus muscles and the muscles related to the eyebrows movements (i.e. the corrugator muscle) do not received any treatments, the conclusion can be the same as the expressions, closed the eyes Softly and tight. This result is also confirmed with the control sample I and II, the amount of asymmetry is huge in the patient's face compared to the normal face. Due to a short smiling at the expression by the control sample I, the asymmetry level (the black curve) is increased from the frame 40-60. The graph reported in Figure 18 confirm this observation. We note here a shift in the time interval which corresponds

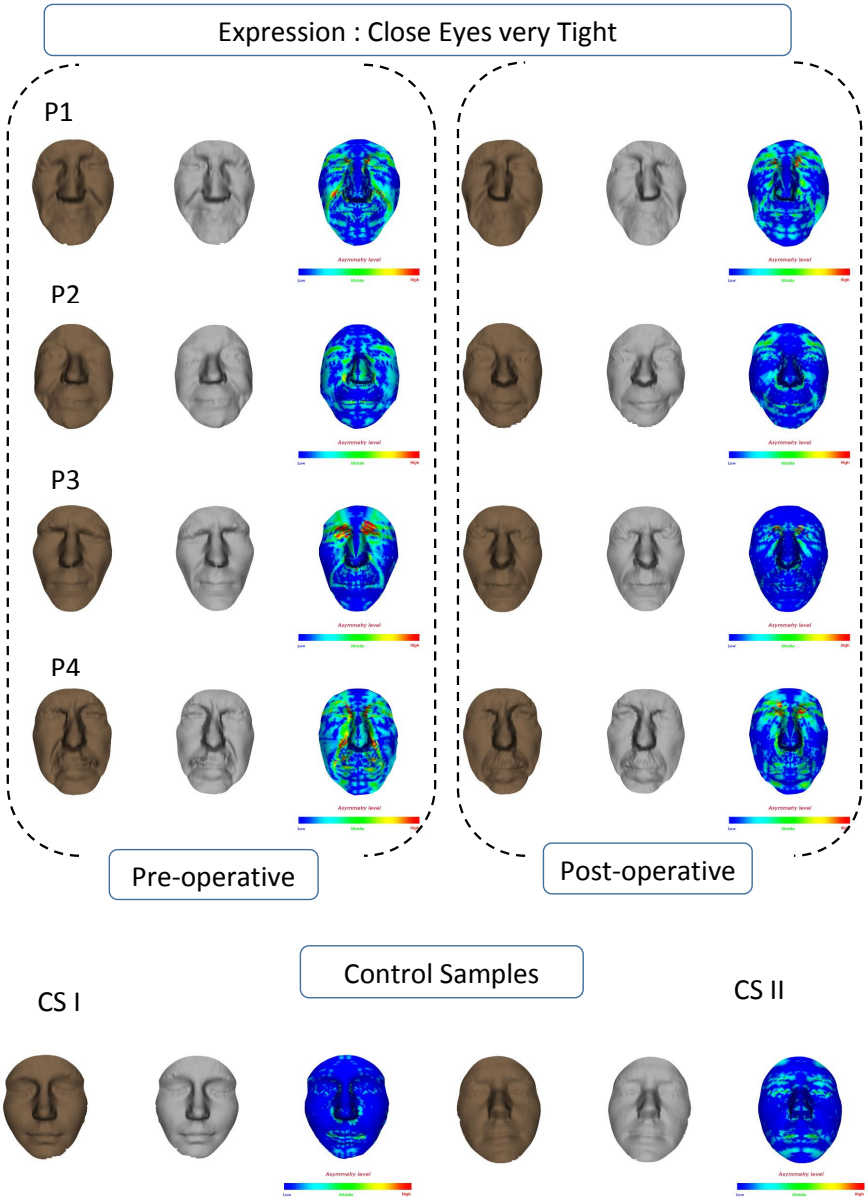

Figure 13: Closed eyes very tight in this experiment.

to different starting and ending time of the expression performed before and after receiving the treatments. What is the behavior all face muscle while say a

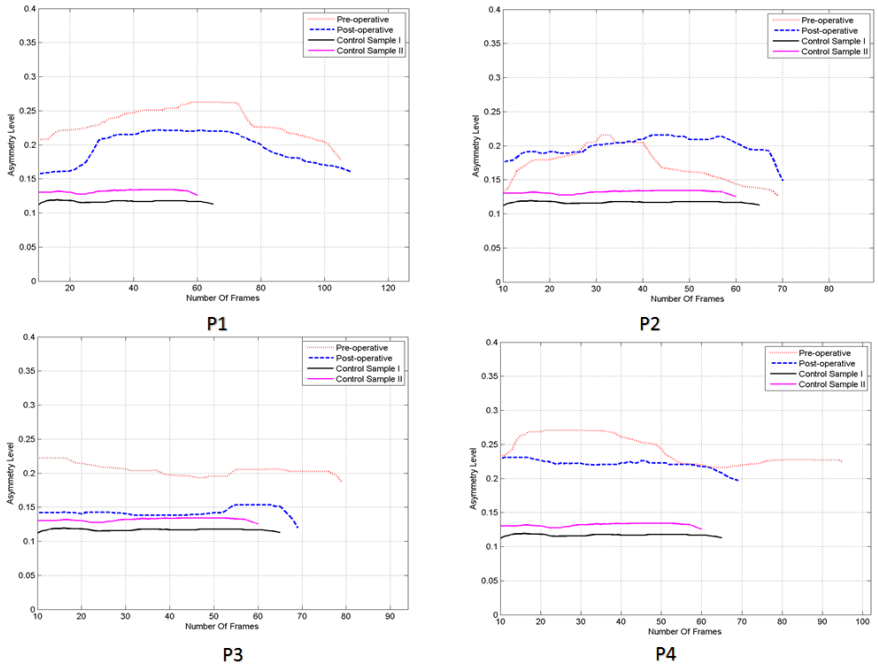

Figure 14: Temporal evolution of the expression of close eyes very tight in this experiment from patients P1 to P4.

sentence ?

8. Repeat a specific sentence – In response to previous question, the clinician asked the patient to repeat this French sentence " Une jolie avenue bordée d'arbre et illuminée ". This french sentence makes work all the zigomaticus muscles in the face. It is possible to observe the deformation of the face for each word that the patient pronounces, and we report in Figure 19 the asymmetry level of the face. We can observe that for the patient P1 and P3, the amount of bilateral asymmetry is lower in post-operative than pre-operative case. This conclusion is not hold for the patients P2 and P4, this is due to the difficulty to repeat and pronounce the word correctly.

One can note that the patients covey the expressions more longer than the control samples, this is due to the difficulty of the patient to convey the expressions.

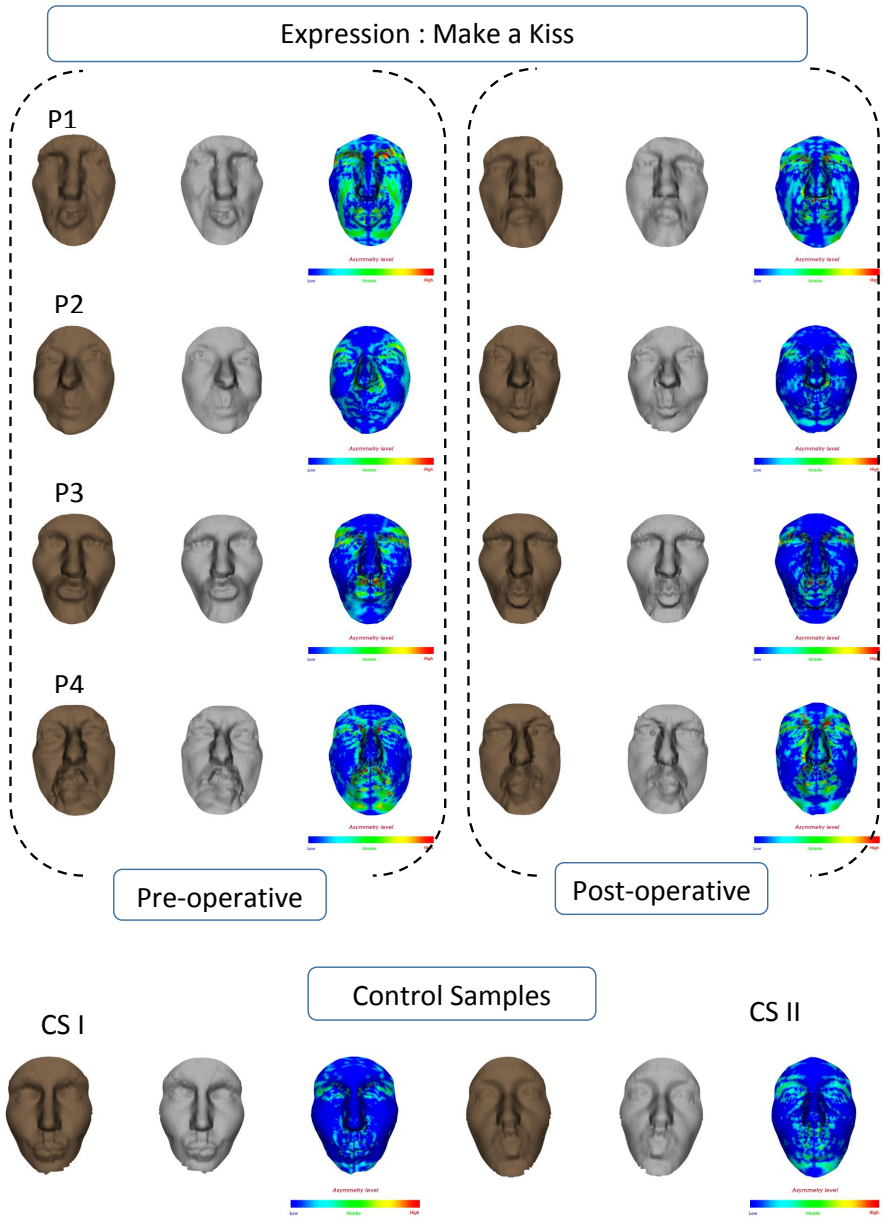

Figure 15: Amount of bilateral asymmetry in a kiss experiment.

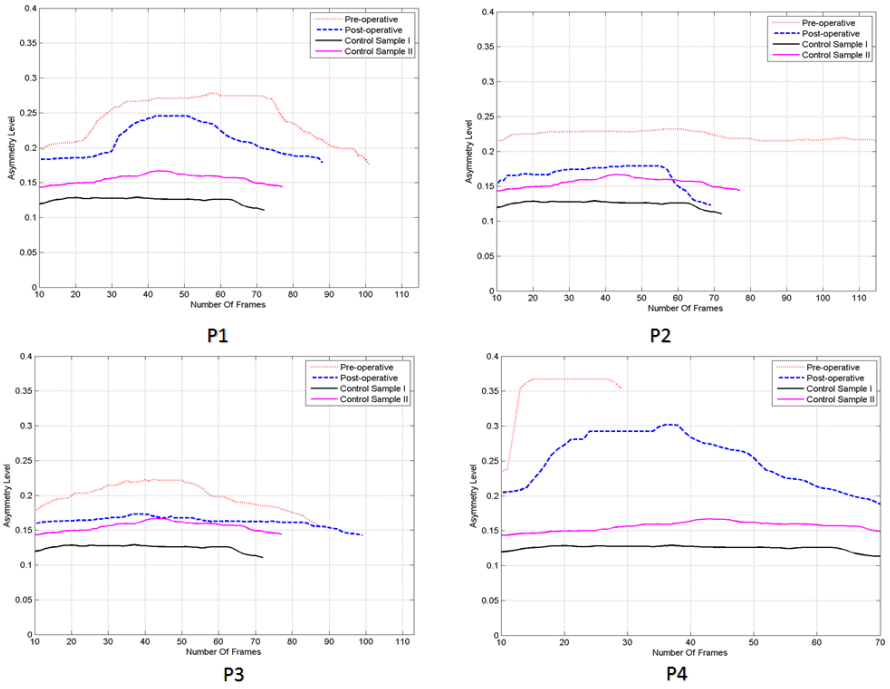

Figure 16: Asymmetry along time for Kissing experiment.

4.2. Computational time

Our pipeline to process an arbitrary 3D mesh, taken individually, is around 2 seconds. It includes (1) the preprocessing (denoising, nose tip detection and cropping), (2) mirror transformation, alignment of the face and its mirror and (3) radial curves extraction and DSF computations. The time needed to process a video depends on the number of frames. For example, a sequence of 100 frames will be processed around 2 minutes 40 seconds. All our programs are written in C++ and run under a 2.3 GHz CPU laptop with a memory size of 16 GB[4].

[4]The processing time exclude the acquisition time of the 3D video.

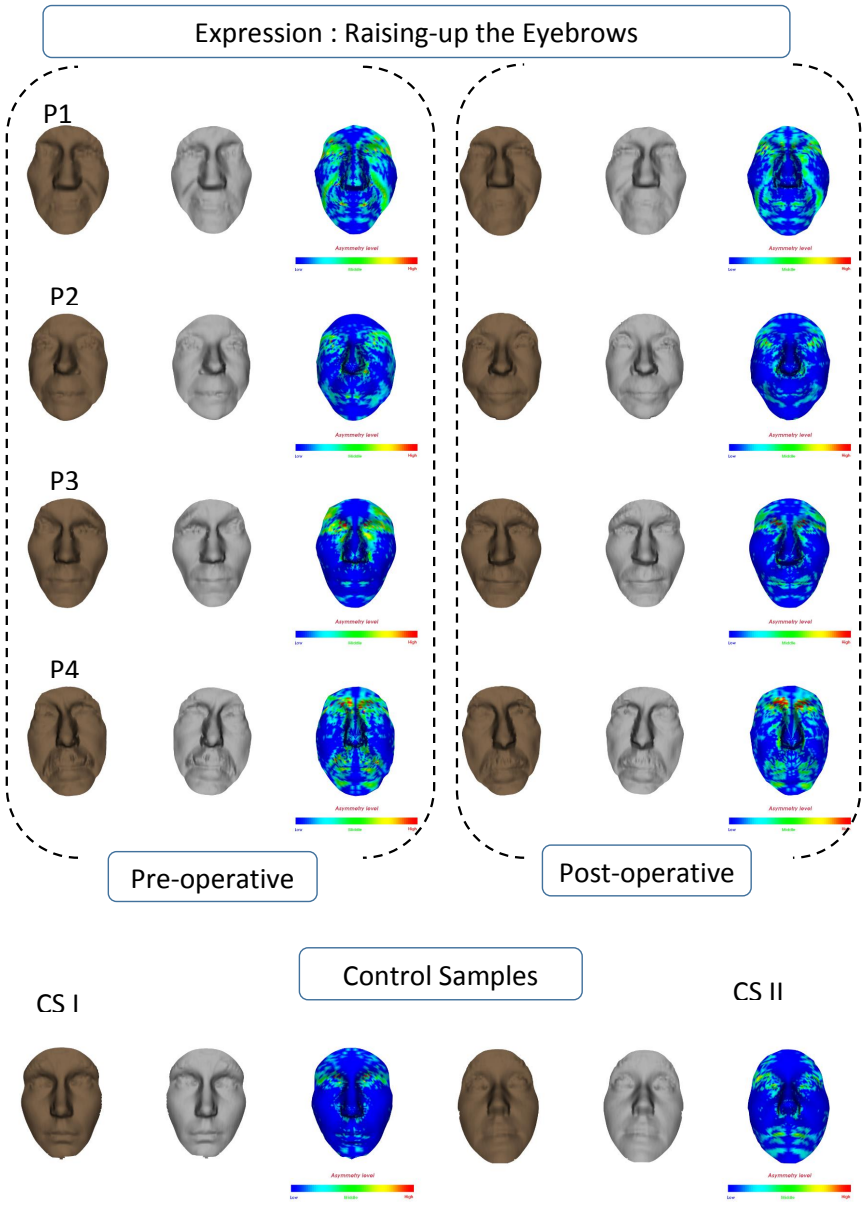

Figure 17: Raising-up the eyebrows experiment.

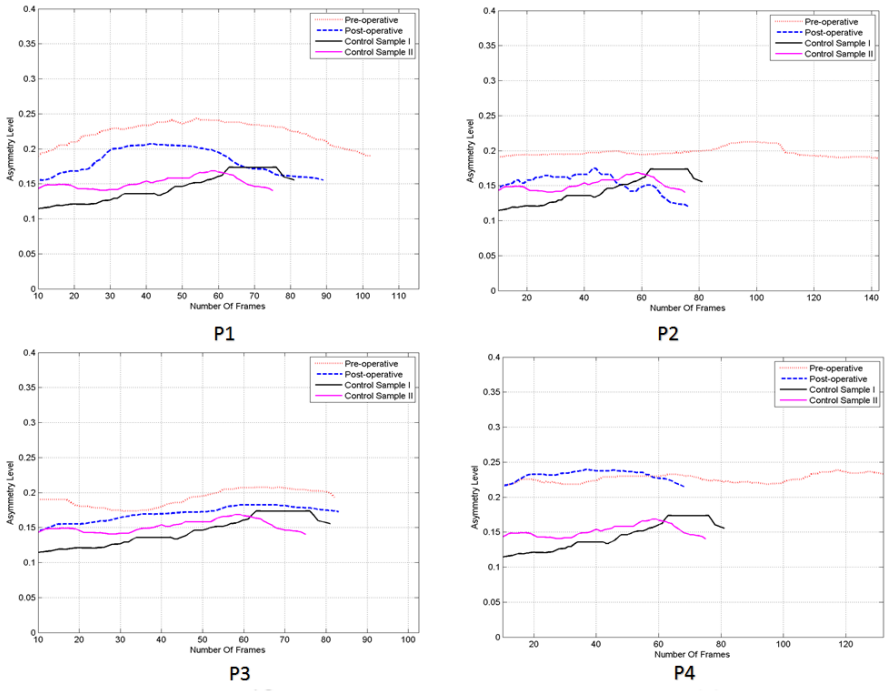

Figure 18: Asymmetry along time for the raising-up the eyebrows experiment.

5. CONCLUSIONS

In this paper we have proposed a new approach to quantify facial asymmetry from 3D sequences. A new data set of 16 patients has been collected in clinical conditions and a new experimental setup to quantify the facial deformations was proposed. We have demonstrated using the collected 3D sequences the effectiveness of the proposed approach to quantify the facial asymmetry. In particular, the comparison of the facial asymmetry by using Dense Scalar Fields (DSF) features before and after the Botulinum Toxin's injection reveals and shows that the proposed approach is a promising solution. This work provides a quantitative tools to the clinicians in order to evaluate the treatment.

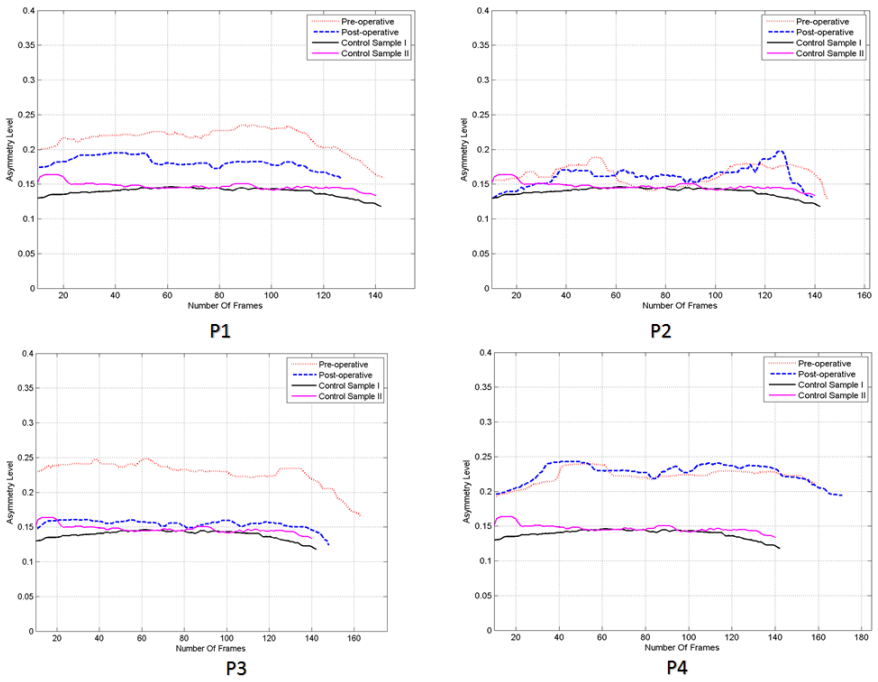

Figure 19: Asymmetry along time while the patients repeat the sentence in the experiment.

Acknowledgment

This work is partially supported the Programme d'Investissements d'Avenir (PIA) and Agence Nationale pour la Recherche (grant ANR-11-EQPX-0023), and European Founds for the Regional Development (Grant FEDER- Presage 41779).

References

[1] P. A. Desrosiers, Y. Bennis, B. Ben Amor, M. Daoudi, P. Guerreschi, Facial asymmetry assessment from 3D shape sequences – the clinical case of facial paralysis, in: The International Conference on Computer Vision Theory and Applications , Rome, Italy, February 27th - 29th.

[2] M. Risoud, N. Aljudaibi, V. Duquennoy-Martinot, P. Guerreschi, Long-term sequelae treatment of peripheral facial paralysis with botulinum toxin type a:

Repartition and kinetics of doses used, May 2015. doi:10.1016/j.anplas.2015.04.002.

[3] L. Benichou, D. Labbe, P. Guerreschi, Facial palsy sequel and botulinum toxin, Oct 2015, 60(5):377-92. doi:10.1016/j.anplas.2015.07.011.

[4] R. Clark, C. Berris, Botulinum toxin: a treatment for facial asymmetry caused by facial nerve paralysis., Plast Reconstr Surg 84 (2) (1989) 353—5.

[5] R. Filipo, I. Spahiu, E. Covelli, M. Nicastri, G. Bertoli, Botulinum toxin in the treatment of facial synkinesis and hyperkinesis., The Laryngoscope 122 (2) (2012) 266–70.

[6] X. Yang, D. Huang, Y. Wang, L. Chen, Automatic 3d facial expression recognition using geometric scattering representation, in: 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2015, Ljubljana, Slovenia, May 4-8, 2015, 2015, pp. 1–6. doi:10.1109/FG.2015.7163090.
URL http://dx.doi.org/10.1109/FG.2015.7163090

[7] H. Li, L. Chen, D. Huang, Y. Wang, J. Morvan, 3d facial expression recognition via multiple kernel learning of multi-scale local normal patterns, in: Proceedings of the 21st International Conference on Pattern Recognition, ICPR 2012, Tsukuba, Japan, November 11-15, 2012, 2012, pp. 2577–2580.
URL http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=6460694

[8] G. Sandbach, S. Zafeiriou, M. Pantic, D. Rueckert, Recognition of 3D facial expression dynamics, Image and Vision Computing 30 (10) (2012) 762–773.

[9] B. Ben Amor, H. Drira, S. Berretti, M. Daoudi, A. Srivastava, 4-D facial expression recognition by learning geometric deformations, IEEE T. Cybernetics 44 (12) (2014) 2443–2457.

[10] Q. Zhen, D. Huang, Y. Wang, L. Chen, Muscular movement model-based automatic 3d/4d facial expression recognition, IEEE Trans. Multimedia 18 (7) (2016) 1438–1450. doi:10.1109/TMM.2016.2557063.
URL http://dx.doi.org/10.1109/TMM.2016.2557063

[11] X. Zhao, E. Dellandréa, J. Zou, L. Chen, A unified probabilistic framework for automatic 3d facial expression analysis based on a bayesian belief inference and

statistical feature models, Image Vision Comput. 31 (3) (2013) 231–245. doi:10.1016/j.imavis.2012.10.001.
URL http://dx.doi.org/10.1016/j.imavis.2012.10.001

[12] X. Zhao, E. Dellandréa, L. Chen, I. A. Kakadiaris, Accurate landmarking of three-dimensional facial data in the presence of facial expressions and occlusions using a three-dimensional statistical facial feature model, IEEE Trans. Systems, Man, and Cybernetics, Part B 41 (5) (2011) 1417–1428. doi:10.1109/TSMCB.2011.2148711.
URL http://dx.doi.org/10.1109/TSMCB.2011.2148711

[13] X. Zhao, J. Zou, H. Li, E. Dellandréa, I. A. Kakadiaris, L. Chen, Automatic 2.5-d facial landmarking and emotion annotation for social interaction assistance, IEEE Trans. Cybernetics 46 (9) (2016) 2042–2055. doi:10.1109/TCYB.2015.2461131.
URL http://dx.doi.org/10.1109/TCYB.2015.2461131

[14] T. Al-Anezi, B. Khambay, M. Peng, E. O'Leary, X. Ju, A. Ayoub, A new method for automatic tracking of facial landmarks in 3D motion captured images (4D), International Journal of Oral and Maxillofacial Surgery 42 (1) (2013) 9 – 18.

[15] S. Shujaat, B. Khambay, X. Ju, J. Devine, J. McMahon, C. Wales, A. Ayoub, The clinical application of three-dimensional motion capture (4d): a novel approach to quantify the dynamics of facial animations, International Journal of Oral and Maxillofacial Surgery 43 (7) (2014) 907 – 916.

[16] W. Quan, B. J. Matuszewski, L. Shark, Facial asymmetry analysis based on 3D dynamic scans, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, SMC 2012, Seoul, Korea (South), October 14-17, 2012, 2012, pp. 2676–2681.

[17] A. Gaber, M. F. Taher, M. A. Wahed, Quantifying facial paralysis using the kinect v2, in: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2015, Milan, Italy, August 25-29, 2015, 2015, pp. 2497–2501.

[18] S. Cheng, I. Marras, S. Zafeiriou, M. Pantic, Active nonrigid ICP algorithm, in: 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2015, Ljubljana, Slovenia, May 4-8, 2015, 2015, pp. 1–8.

[19] T. Fang, X. Zhao, O. Ocegueda, S. K. Shah, I. A. Kakadiaris, 3D/4D facial expression analysis: An advanced annotated face model approach", Image and Vision Computing 30 (10) (2012) 738–749.

[20] A. Srivastava, E. Klassen, S. H. Joshi, I. H. Jermyn, Shape analysis of elastic curves in euclidean spaces, IEEE Trans. Pattern Anal. Mach. Intell. 33 (7) (2011) 1415–1428.

[21] S. Joshi, E. Klassen, A. Srivastava, I. Jermyn, A novel representation for Riemannian analysis of elastic curves in $\mathbb{R}^n$, in: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Minneapolis, MN, 2007, pp. 1063–6919.