Machine Vision and Applications (2025) 36:1
https://doi.org/10.1007/s00138-024-01616-1

RESEARCH

Assessing 3D volumetric asymmetry in facial palsy patients via advanced multi-view landmarks and radial curves

Tim Büchner[1] · Sven Sickert[1] · Gerd F. Volk[2] · Orlando Guntinas-Lichius[2] · Joachim Denzler[1]

Received: 14 March 2024 / Revised: 25 September 2024 / Accepted: 30 September 2024 / Published online: 4 November 2024

Abstract

The research on facial palsy, a unilateral palsy of the facial nerve, is a complex field with many different causes and symptoms. Even modern approaches to evaluate the facial palsy state rely mainly on stills and 2D videos of the face and rarely on dynamic 3D information. Many of these analysis and visualization methods require manual intervention, which is time-consuming and error-prone. Moreover, they often depend on alignment algorithms or Euclidean measurements and consider only static facial expressions. Volumetric changes by muscle movement are essential for facial palsy analysis but require manual extraction. We propose to extract an estimated unilateral volumetric description for dynamic expressions from 3D scans. Accurate landmark positioning is required for processing the unstructured facial scans. In our case, it is attained via a multi-view method compatible with any existing 2D predictors. We analyze prediction stability and robustness against head rotation during video sequences. Further, we investigate volume changes in static and dynamic facial expressions for 34 patients with unilateral facial palsy and visualize volumetric disparities on the face surface. In a case study, we observe a decrease in the volumetric difference between the face sides during happy expressions at the beginning (13.8 ± 10.0 mm^3) and end (12.8 ± 10.3 mm^3) of a ten-day biofeedback therapy. The neutral face kept a consistent volume range of $11.8-12.1$ mm^3. The reduced volumetric difference after therapy indicates less facial asymmetry during movement, which can be used to monitor and guide treatment decisions. Our approach minimizes human intervention, simplifying the clinical routine and interaction with 3D scans to provide a more comprehensive analysis of facial palsy.

Keywords Facial volumes · Facial palsy · Radial curves · Emotions

1 Introduction

Advancements in imaging techniques provide novel insights across various disciplines, particularly medicine. Especially in the area of facial palsy, a unilateral palsy of the facial nerve [1–7], 3D scans offer new capabilities to analyze the treatment progress [4, 8, 9]. The limited scope of 2D images cannot capture the palsy's full extent, such as the depth of the nasolabial fold or the volume differences between the two sides of the face. We think, that 3D models could bridge this gap [1, 9]. Obtaining volumetric information can be an indicator for facial changes [8, 10]. In reality, facial muscle contractions lead to 3D volume shifts of facial soft tissues, i.e., the muscle, fat tissue, and overlying skin. Movement is the primary driver of volumetric changes. Hence, static analysis alone is insufficient to capture the full complexity of facial palsy. In our work, we analyze the static state of the face in combination with the dynamic properties of *neutral* and *happy* expressions during volume changes. Our method is also applicable to other dynamic expressions, such as the six base emotions or the functional movements for evaluating facial palsy [1, 11]. Including dynamic facial surface information is crucial to bridging the gap between mimics and

✉ Tim Büchner
tim.buechner@uni-jena.de

Sven Sickert
sven.sickert@uni-jena.de

Gerd F. Volk
fabian.volk@med.uni-jena.de

Orlando Guntinas-Lichius
orlando.guntinas@med.uni-jena.de

Joachim Denzler
joachim.denzler@uni-jena.de

[1] Computer Vision Group, Friedrich Schiller University Jena, Ernst-Abbe-Platz 2, 07743 Jena, Thuringia, Germany

[2] Department of Otorhinolaryngology, Jena University Hospital, Am Klinikum 1, 07747 Jena, Thuringia, Germany

muscles. All of this together is highly relevant to describe the two halves of the face objectively and holistically in psychophysical experiments, especially in diagnosing facial muscle diseases. Unilateral facial palsy, i.e., a facial nerve disease with disturbed facial muscle function, is characterized by facial asymmetry. The method presented should make it possible to accurately describe changes during therapy or take targeted therapeutic measures based on volumetric changes.

Many methods for processing 3D facial data involve human interaction [9, 12]. It makes them time-consuming and prone to errors. Automatic approaches, such as 3D-morphable models [13–18], are not suitable for facial palsy patients, because their underlying templates have been trained predominantly on healthy faces. To avoid template-based limitations and errors in the fitting process, we leverage radial curves along the scan surface [8, 19–22]. These curves are computed on the original scan surface and provide accurate facial descriptions. Our work utilizes the radial properties to estimate lateral facial volumes and to visualize disparities between the facial sides. Consequently, the highlighted regions can support medical professionals in providing individualized treatment plans.

First, we provide a method to reliably place 3D facial landmarks for patients with facial palsy and reduce the runtime of exiting multi-view methods by 96.6% [22, 23]. Furthermore, we generalize the processing pipeline, enabling integration with any existing 2D landmark extraction technique to avoid dependency on specific predictors. We analyze the stability of landmark placement and assess the robustness against head rotation estimation during video sequences.

Starting at the nose tip, we extract radial curves from the 3D scan to obtain a dense facial description [8, 19–21]. These curves offer a detailed and structured description of the scan surface for lateral comparisons. We emphasize disparities in facial volume during static moments by projecting the curves onto the face and visualizing the asymmetries between the facial sides for *neutral* and *happy* expressions. Finally, we track and visualize them during dynamic expressions to combine volumetric changes associated with possible muscle contractions. From a single 3D facial scan acquired during routine clinical practice, we generate volumetric disparity maps, enabling medical experts to assess patient changes across multiple visits to monitor disease progression or assess treatment efficacy. This non-invasive approach allows for safe and repeated monitoring of patients with a standardized capture setup. Our approach minimizes human intervention during analysis, is fully automatic after parameter selection and is available as open-source software.[1]

[1] https://github.com/cvjena/corc and https://github.com/cvjena/mvlm.

2 Related work

Existing automated methods for assessing facial palsy [4, 6, 7, 24–28] rely on 2D images, either by using landmarks [2, 29, 30] or neural networks [2, 24, 31, 32] to estimate state of the nerve damage. Most often, they attempt to replicate existing grading methods such as Sunnybrook [33] or House-Brackmann [3]. However, many critical facial features cannot be captured by 2D images alone, such as the depth of the face or the volume differences between the two sides. Muscle activity is the cause of any facial change. Thus, the use of 2D images alone is insufficient to interpret these changes. It requires a combination with electromyography [11, 34–37]. We aim first to facilitate detailed analysis and visualization of the facial palsy state based on 3D scans. After that, we enhance it with electromyography data.

Automated facial palsy grading

Recent studies have explored the use of machine learning and deep learning methods for 2D facial palsy grading [4, 6, 7, 24–28, 38], which have shown promising results in terms of accuracy and efficiency [2, 24, 29–32, 39–42]. Most often, they attempt to replicate existing grading [31, 33, 41, 43–45] or attempt to utilize other facial features to estimate the patient health state [29, 46–48]. One main advantage of these approaches is their ability to analyze 2D facial images, which are widely available and easily accessible, making them a cost-effective solution for facial palsy diagnosis.

However, such machine learning methods also have some limitations. For example, if not fully represented in the training data, they may not capture the full range of facial expressions and movements, which can be necessary for accurate diagnosis. Furthermore, 2D methods may be more susceptible to errors due to variations in image quality and resolution or lighting and pose changes. Furthermore, black-box models often resort back to shortcut learning by leveraging biases in the dataset resulting in worse performance applied to new data [49–56]. For example, a model trained on data from a single hospital may not generalize well to data from other hospitals or may be biased toward certain patient demographics [55, 57].

As 2D images cannot capture facial depth information, we leverage 3D facial scans to understand facial asymmetry due to the underlying volumetric differences.

3D facial landmarking

The most common approach to integrate 3D information into facial palsy analysis is to extract 3D landmarks from the scan. Hence, existing 2D grading systems could be adopted [4]. However, existing 2D landmark positioning methods frequently fail for facial palsy patients due to

asymmetry in facial features. This issue gets even worse during facial expressions. Consequently, manual intervention is often required [9, 12]. Nevertheless, landmarks remain the fundamental component for subsequent analyses as they inject semantic information into the unstructured 3D point cloud or scan [8, 13, 16, 22]. A universal and stable method for 3D landmark extraction, adaptable to patients with facial palsy, is required.

Reliable placement of 3D landmarks based on 2D images necessitates camera parameters [4, 8]. Alternatively, given only a 3D scan, simulated camera views enable multi-view-based landmark prediction [22, 23]. This latter method is more adaptable and applicable to any facial 3D scan. Our work further refines the technique to enhance its generality and compatibility with any existing 2D landmark extraction method.

3D morphable models

Since 3D landmarks provide a sparse representation that fails to capture the full complexity of facial anatomy, a more detailed description is imperative to obtain an accurate facial characterization. 3D Morphable Models (3DMMs) represent a prevalent method for describing 3D facial scans [13–18, 58]. They utilize a learned template that encompasses a shape and expression space to characterize the face. During the template fitting process, 3D landmarks are critical for defining the rigid transformation regarding alignment and scaling [13, 14]. Therefore, reliable landmark placement is essential for fitting, particularly for facial palsy patients. Adapting 3DMMs for facial palsy patients is challenging, because the underlying templates are trained predominantly on healthy faces [13, 17, 59]. Depending on the expressiveness of the shape and expression space of the 3DMM, the fitting process might eliminate detailed information on the face surface to achieve a more overall fit [13]. Especially in [13], the expression space is learned via PCA [60] and changes of both face sides are included in a single expression space vector. Other works such as [17] separate the expression space into two distinct spaces for each face side but still inherently rely on using symmetric information during the fitting process. The current research comprehensively describes the face surface and creates highly detailed facial models. However, adapting 3DMMs directly to facial palsy patients is challenging and requires further research.

Consequently, we refrain from using 3DMMs for lateral comparisons, trading semantic knowledge for a more comprehensive description of the face surface. Current methods for tracking volume changes in the mid and lower face rely on sparse features [10]. Their findings indicate that subtle facial expressions lead to measurable volume changes, which we aim to capture using a dense representation.

Radical curves

Some 3D palsy assessment methods require manual interaction [9, 12, 61, 62], while others attempt automation [8, 20, 22] but focus only on face surface analysis. Utilizing radial curves [8, 21] shows potential for detailed anatomical descriptions. Initially aligned using landmarks, the curves remain on the original face scan surface, retaining their features and overcoming template-based limitations. Radial properties enable structured face surface descriptions to be divided into distinct regions. Existing facial palsy approaches demonstrated this [8, 22]. Nonetheless, the one-to-one mapping between curve points eliminates the influence of each side on the other. We overcome limitations by generating lateral face meshes using radial curves and revealing volume disparities between both facial sides. Our fully automatic approach minimizes human intervention, simplifying the clinical routine and interaction with 3D scans while providing a more comprehensive analysis of facial palsy.

3 Data acquisition

We investigate volumetric changes in patients with facial palsy. Facial muscle contractions lead to 3D volume shifts of facial soft tissues, i.e., the muscle, fat tissue, and overlying skin [10]. Given that the intricate network of facial muscle activation is the principal driver of volumetric changes [11], our focus is beyond static captures but includes dynamic facial expressions. We used the 3dMD face system (3dMD LCC, Georgia, USA) to capture these movement exercises, which generates a 3D facial mesh using multiple 2D images and infrared structured light. This setup ensures that the mesh faces towards the z-axis, and the face remains close to the coordinate system's center. The patients follow an instruction video. They are asked to maintain a neutral expression for three seconds, followed by 3.5 seconds of mimicking a shown *happy* expression. The recorded video is six seconds long, and the 3D scans are captured at 30 frames per second. We obtained 210 individual 3D face meshes per recording in the wavefront (.obj) data format. Each patient repeats the movement four times based on the images shown to capture varieties of muscle activation and ensure reproducibility.

We recorded 34 patients (age: 25–72; 27 female, seven male) suffering from postparalytic facial syndrome at their ten-day bio-feedback treatment's beginning and end with the above-described recording setup [63]. Therefore, our dataset consists of 272 video recordings of patients doing a dynamic expression going from natural to happy. Our data contains only patients with unilateral chronic synkinetic facial palsy with symptoms persisting for over six months. A common symptom of flaccid facial palsy is muscle atrophy. The affected side has a lower volume than the healthy side

[1]. Volume changes are expected due to the imbalance on the affected and the contralateral side in the synkinetic patients. For our visualizations, we will focus on six patients[2], with different palsy severities, to demonstrate the capabilities of our method. All remaining results for this case study are given for our complete patient cohort.

4 Methods

Our goal is to achieve a fully automatic analysis of volumetric changes caused by facial tissue shifts during muscle contractions. Furthermore, we want to visualize the local volumetric differences between the face sides. First, we extract 3D landmarks from the 3D scan by adapting existing methods to patients with facial palsy [22, 23]. We generalized the multi-view landmarking approach to enhance compatibility with any pre-existing 2D landmark extraction technique for corresponding 3D landmark acquisition. Our method works with static and dynamic expressions without fine-tuning the feature extractors on our dataset.

To overcome the sparse nature of 3D landmarks, we extract radial curves from the 3D scan using these landmarks to obtain a dense facial representation [8, 22]. We construct a volume-based characterization of the facial sides utilizing the radial curves, see Fig. 3.

Lastly, we employ the lateral properties (side-by-side division) of radial curves to produce a disparity heatmap and compare it to the volumetric results. We aim to comprehensively analyze a cohort of patients with minimal human intervention, including static moments and dynamic movements. Our method can also be applied to point clouds because it exclusively utilizes the vertices of the 3D mesh.

4.1 Scan preprocessing: multi-view landmark extraction

A structured semantic representation of 3D facial features via landmarks is crucial for many processing steps to interact with an unstructured 3D point cloud or scan. These steps include the fitting of 3D Morphable Models [13, 14, 16–18] or head rotation estimation for radial curves [8, 19, 20, 22]. A standard approach involves using 3D landmarks estimated from 2D images using known intrinsic and extrinsic camera parameters or directly extracted from the 3D scan. The latter approach, requiring only the 3D scans, offers greater flexibility and broader applicability, especially when camera parameters are unknown. Head rotations, occlusions, movements, and facial palsy symptoms also contribute to landmark placement. Hence, a reliable and adaptable method for 3D landmark extraction is essential.

Fig. 1 Camera locations to approximate 3D landmarks: The camera positions are on a sphere around the original scan to generate lateral views [22, 23]. Patient with unilateral facial palsy on the right side of the face (left side of the image)

We opt for a multi-view 3D landmark extraction method based on [22, 23], which needs no knowledge about the 3D scan parameters. In previous work [23], Paulsen et al. utilize numerous simulated virtual camera views ($n = 96$) encircling the face to capture all possible nuances. The camera locations are randomly sampled on a half-sphere in front of the face with the same distance to the face center. Furthermore, the authors employ RGB and depth images as inputs for a custom-trained convolutional neural network to predict potential landmark locations through generated heatmaps. They designate the heatmaps' maximal value as the predicted landmark location. With the known virtual camera parameters and predicted 2D landmarks, the authors calculate the potential intersections of the camera rays and predicted landmarks to estimate the 3D landmark positions. Owing to the high number of cameras and the potential for erroneous predictions due to occlusion, RANSAC was used to identify the most reliable landmark from the line intersections. The authors accommodate two distinct landmark sets: the BU-3DFE [59] (84 landmarks) and the DTU3D [64] (73 landmarks) schemes.

Implementation details

The original implementation takes 30 s to extract the landmarks for one single scan on an Nvidia GeForce RTX 3070 [22, 23]. Hence, processing a six-second video (at 30 frames per second) would take around 5400 s (90 min). To promote use in clinical practice, we adapt their methodological approach, focusing on improvements in speed and generality. As shown in [22], we achieve suitable landmarks with only eight fixed camera locations. Our fixed camera locations are: *yaw* angles of −40, −20, 20, and 40 degrees and *pitch* angles of −30 and 30 degrees. Employing lateral views of the face, as illustrated in Fig. 1, reduces potential symmetry biases in the prediction models [5]. Since patients frequently

[2] All shown individuals agreed to have their images published in terms with the GDPR.

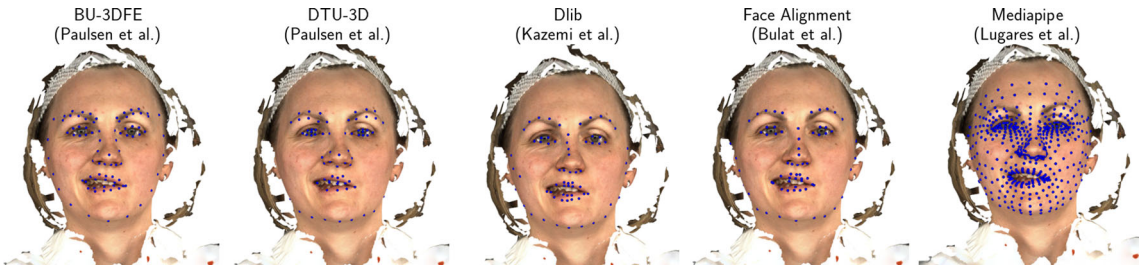


Fig. 2 The figure shows the predicted facial landmarks of each landmark extraction technique (blue dots). The facial palsy patient shows asymmetry during the *happy* expression in the eye contour, nose tip, and mouth region. The visualization is normalized such that the palsy side is on the left side of the face (right side of the image) (colour figure online)

exhibit unilateral palsy where only one side of the face is affected, we analyze each side independently. By limiting rendering calls to two per camera view, one for RGB and one for depth, we achieve considerable improvements in processing speed. Additionally, we eliminate rendering pipeline overhead, minimize superfluous memory access, and reduce duplicate I/O operations and cache misses. We have substantially decreased the computation time for a single surface scan to one second. The processing time is reduced to 180 s for a six-second video.

Generalization and integration

We have also expanded the method's versatility, allowing it to be compatible with any pre-existing 2D landmark extraction technique. The original method relies on the maximal heatmap value location as the predicted landmark aligns with a generic approach to 2D landmark extraction. Consequently, we can divide the workflow into generalized discrete steps: multi-view rendering, 2D prediction, and 3D landmark estimation. Only the 2D prediction step is specific to the chosen predictor, and the other steps are independent of the predictor. Adopting this strategy, we have natively integrated five different extractors for landmark estimation [15, 22, 23, 65, 66], though our pipeline readily accommodates the integration of additional predictors. Figure 2 presents the predicted facial landmarks for each of the five 2D landmark extraction techniques. We evaluate the stability of landmark placement, especially for the nose tip, eye contour, and mouth region, by assessing their robustness regarding head rotations over the video sequence.

Our generalized approach offers enhanced flexibility and is designed so that its integration has no adverse effects on runtime performance, dependent only on the predictor's speed. Consequently, the method can be adapted to any facial 3D scan and integrated with existing 2D landmark extraction techniques with minimal effort, requiring no additional fine-tuning.

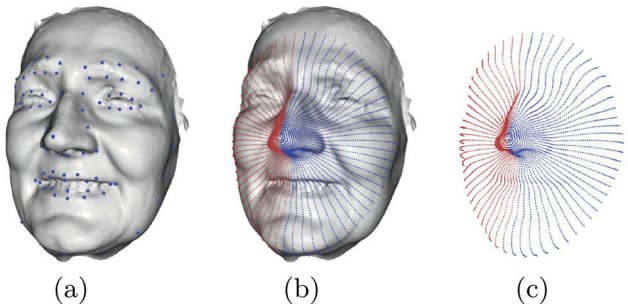


Fig. 3 We highlight the extracted landmarks (a) and the resulting radial curves Γ (b) on the face scan of a patient with facial palsy. In (c), the curves Γ_{left} describe the left (blue) and right Γ_{right} (red) sides of the face [22] (colour figure online)

4.2 Volume feature extraction: facial description via radial curves

In a matrix $\mathcal{X} \in \mathbb{R}^{k\times 3}$ with k points, either a point cloud or the vertices of a mesh, radial curves are a structured description of the scanned surface [8, 19–22]. Radial curves, emitting from a common start point and limited by a maximal distance r to the start point, are surface descriptors defined by a tensor $\Gamma \in \mathbb{R}^{n\times m\times 3}$ with n curves consisting of m curve points. In the case of faces, the nose tip is the start point, and the chin limits the distance [8, 19, 20], as depicted in Fig. 3.

To extract the curve group Γ, alignment is crucial to compare facial regions later. Thus, we normalize $\mathcal{X}$ with $\tilde{\mathcal{X}} = (\mathcal{X} - T) \cdot R_H^{-1}$. Hence, we use the estimated 3D landmarks on the scan to obtain the orientation and position of the face. The translation vector $T \in \mathbb{R}^3$ is the nose tip, and $R_H \in \mathbb{R}^{3\times 3}$ is the head pose obtained from the eye and mouth landmarks [67]. If the 3D landmarks are misaligned, the estimated head rotation matrix and the radial curves may be skewed. Nonetheless, if misalignment remains consistent throughout a video sequence, volume differences can still be estimated, as the misalignment constantly affects both sides of the face. In our experimental evaluation, we will assess the stability of the head rotation estimation and its influence on the radial curves.

Given the normalized scan $\bar{\mathcal{X}}$, each radial curve γ_α resides on a plane p_α, defined by normal d_α^n and vector along it d_α, rotated through the nose tip with an angle $\alpha = \frac{360°}{n}$. For lateral descriptions, n has to be an even number and greater equals than four [8, 20]. All points in $\bar{\mathcal{X}}$ that lie around the plane p_α with maximal distance δ are selected [8, 20]:

$$S_\alpha = \left\{ x_i \in \bar{\mathcal{X}} \middle| |x_i^T d_\alpha^n| < \delta \wedge x_i^T \cdot d_\alpha > 0 \right\}. \quad (1)$$

With the 3dMD camera system (3dMD LCC, Georgia, USA), we empirically set $\delta = 0.8$ mm to ensure a dense point set S_α. Other sensors might require different values. A temporary projection of S_α onto the plane p_α simplifies ordering and fitting, as seen in Fig. 4a. We sort the points by distance to the nose tip [20] and afterward for a correct surface description with a graph traversal algorithm [8]. A 2D spline approximates the curve γ_α in the ordered points $\bar{S}_\alpha$. We equidistantly sample m point between the nose tip and the most distant point, shown in Fig. 4d. Lastly, we project the spline points back into the original 3D coordinate system. After computing all n curves, we obtain Γ. Aligning the curve tensor to the scan is only necessary for visualization and is done with $\Gamma := \Gamma \cdot R_H + T$, as seen in Fig. 3. The curve tensor Γ forms the foundation for generating lateral face meshes and estimating volume. In previous studies, we have utilized curvature measurements along individual curves, denoted as γ_α, to predict the severity of facial palsy [8, 20].

Lateral face mesh generation

The tensor Γ represents a structured face surface. The pointwise difference between the lateral curve pairs (γ_α and $\gamma_{360-\alpha}$) computes asymmetry [8]. However, this approach neglects the overall state of each face, especially volume information, which is measurable even for small movements [10]. We overcome this limitation by estimating unilateral facial volume using a watertight mesh of Γ. A watertight mesh is a 2-manifold mesh [68], where every edge is part of precisely two faces [69]. Despite receiving a mesh surface, we create a new volumetric mesh to ensure watertight properties and support point cloud scanners. As we already have a face surface descriptor Γ, we do need general surface reconstruction algorithms [65, 70]. The left tensor Γ_{left} includes all curves γ_α with $\alpha \in [0, 180]$ and the right tensor Γ_{right} all curves γ_α with $\alpha \in [180, 360]$. Please note that curve γ_0 and γ_{360} are identical, and the views are from the patient's point of view.

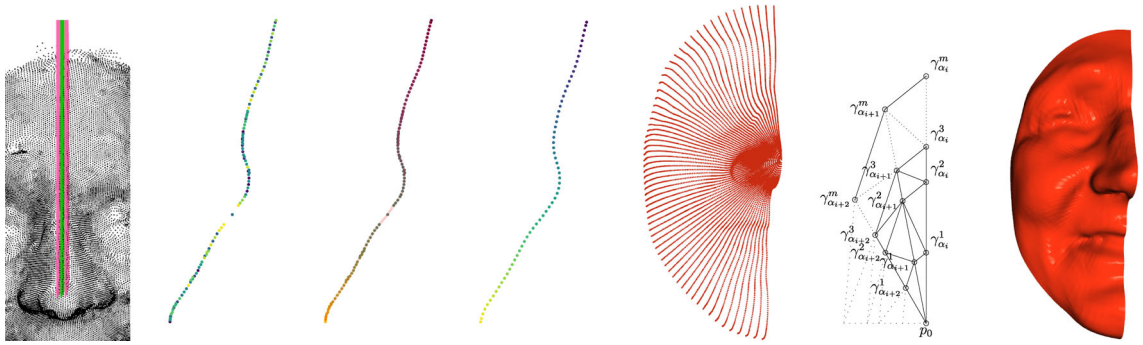
Γ_{left} and Γ_{right} construct the closed lateral face surfaces $\mathcal{M}_{\text{left}}$ and $\mathcal{M}_{\text{right}}$, as seen in Fig. 4g. The borders of $\mathcal{M}_{\text{left}}$ and $\mathcal{M}_{\text{right}}$ are connected with an underlying sphere mesh, as seen in Fig. 5b, and a side mesh along the vertical face center, as seen in Fig. 5c. All steps are watertight, and we apply the combined mesh for the volume approximations. A tetrahedron volume can be computed using each triangle, with the origin as the fourth point. We obtain a correct volume by ensuring the triangle winding order is counter-clockwise [69].

Radial curve mesh generation

The generation algorithm is based on the radial curves tensor Γ_{left} and Γ_{right}, see Fig. 4e. As each curve γ lies on the face's scan surface, the generated mesh also represents the face surface, see Fig. 4g. The degree of detail is adjustable by the number of points m and the number of curves n. Each curve starts at the same point p_0, which is the nose tip of the face. The equidistantly sampled points ensure that points between adjacent curves describe similar perimeter locations. Between two adjacent curves γ_{α_i} and $\gamma_{\alpha_{i+1}}$, we create a triangulation pattern shown in Fig. 4f. The triangulation $\mathcal{M}$ includes all points, and the mesh is without holes and watertight. Our approach relies only on the extracted radial curve tensors Γ_{left} and Γ_{right} without any hyperparameters. The outer edges of $\mathcal{M}_{\text{left}}$ and $\mathcal{M}_{\text{right}}$ connect to the underlying sphere and side meshes, linking every edge precisely to two faces.

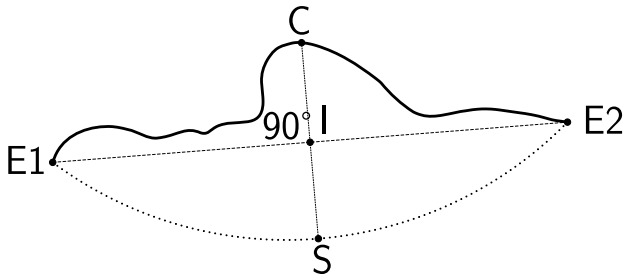
Underlying sphere mesh

Several approaches are possible to close the volume below the face mesh to enable volume calculations. First, we require a south pole point S as a reference point. We use the endpoint $E1$ of the curve γ_{180} and the endpoint $E2$ of the curve γ_0 as orientation. Additionally, we use the nose tip C as a third reference point, shown in Fig. 5a. S is the perpendicular intersection of the centerline and the line connecting $E1$ and $E2$ with the distance r, the same value as for the radial curve extraction, to C. The case of $E1$, $E2$, and C being co-linear cannot occur since these points reference anatomical facial structures and would otherwise indicate a severe facial deformity or a severe error during previous extraction steps. We could use the south pole point S and the boundary points of the face to form simple triangles. However, different facial shapes (e.g., a deep eye socket or mouth being open) will lead to self-intersections [69]. A spherical approach prevents trivial self-intersections. They are theoretically possible but unlikely in practice, as the sphere is convex, and the head will fit inside the sphere. For each radial curve γ_α, we calculate the arc on the corresponding sphere using the corresponding endpoint and S, ensuring alignment with the outer edges of the surface mesh. We triangulate sphere segments like the facial mesh, as depicted in Fig. 5b. Consequently, the underlying sphere is closed and directly connected to the scan surface mesh.

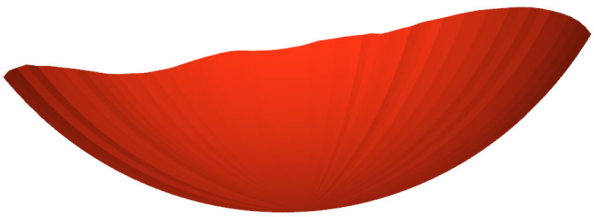


(a) δ area (b) S_α (c) $\bar{S}_\alpha$ (d) γ_α (e) Γ_{right} (f) Meshing (g) $\mathcal{M}_{\text{right}}$

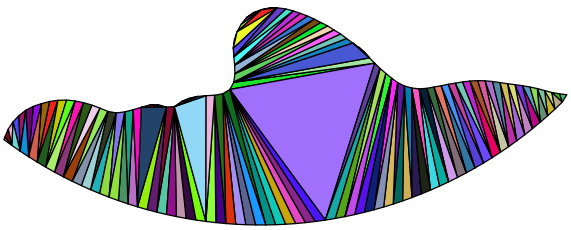
Fig. 4 We visualize each curve extraction process step and the resulting triangulation $\mathcal{M}_{\text{right}}$ of the right facial surface scan. The color coding of S_α, $\bar{S}_\alpha$, and the γ_α indicate the order of the points. Details are found in the text [22]



(a) Construction of the arc segments



(b) Triangulated underlying sphere mesh



(c) Mesh using constrained Delaunay triangulation

Fig. 5 The side mesh (b) fills the gap between the facial and underlying sphere mesh (a). The inner edges of the radial curves (a) serve as orientation [22]

Face side mesh

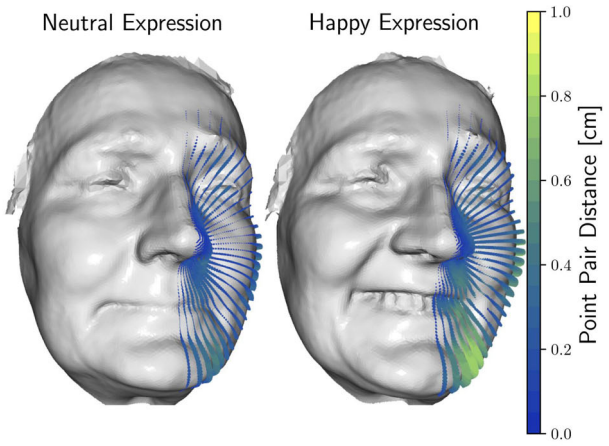
The final component is the face side, closing the facial and sphere mesh gap. A triangulation scheme as in Fig. 4f is not feasible due to the different point densities. However, we can use a constrained Delaunay triangulation to fill the gap [71, 72]. We use the curves γ_0 and γ_{180} of the curve tensor Γ and sphere meshes boundary edges to constrain the triangulation, as depicted in Fig. 5a. The face side mesh is useable for both the left and right sides of the face. This operation guarantees a watertight mesh, as the boundary edges constrain the triangulation. Therefore, we obtain three watertight meshes: the facial mesh $\mathcal{M}_{\text{left}}$ and $\mathcal{M}_{\text{right}}$, the underlying sphere- and the face side mesh.

Volume estimation for lateral face sides

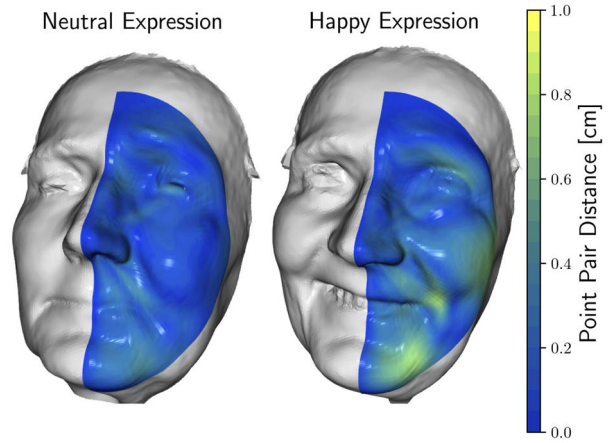
We create a 2-manifold mesh from the radial curve tensor Γ_{left} and Γ_{right}, as described in Sect. 4.2. Please note that self-intersections are possible but unlikely in practice. The meshes are watertight and closed, allowing us to estimate the patient's lateral face volume; examples are shown in Fig. 10. Summing up the signed tetrahedrons' volumes yields the total volume of the mesh [69]:

$$V = \sum_{i=1}^{n} \frac{1}{6} \cdot (p_i^{(1)} \times p_i^{(2)} \cdot p_i^{(3)}), \tag{2}$$

where $p_i^{(1)}$, $p_i^{(2)}$ and $p_i^{(3)}$ are the vertices of the triangle i. The $\times$ operator denotes the cross product. We assume all tetrahedrons share the same origin $[0, 0, 0]^T$ as a fourth point. Overlapping tetrahedron bodies cancel each other out due to the triangle's counter-clockwise winding order, thus providing a correct volume estimation. As before, the degree of detail is adjustable by the number of points m and the number of curves n. We observed that the relative difference among the face sides is not affected significantly. We chose $n = 64$ curves and $m = 64$ spline points in our analysis to balance accuracy and computation time. The calculations are conducted in the original coordinate system, resulting in real-world values measured in millimeters.



(a) Vector field visualization $\mathcal{V}$ between tensors Γ_{left} and $\bar{\Gamma}_{\text{left}}$.



(b) Projection of the volumetric difference onto the facial surface

Fig. 6 Both visualizations show the disparities (maximum 1 cm) between the left and right sides of the face between the *neutral* and *happy* expressions [22]. The visualization is normalized such that the palsy side is on the left side of the face (right side of the image)

4.3 Interpretable heatmaps: volume difference visualization

Volume estimation provides crucial information about the disparities between the left and right sides of the face. Additionally, identifying the specific locations of volumetric differences can aid doctors in optimizing the treatment of facial palsy patients. We assume the left side is the palsy side, and the right is the healthy side. Mirroring the curve tensor Γ_{right} along the vertical centerline, defined by γ_0 and γ_{180}, yields $\bar{\Gamma}_{\text{left}}$, which describes the left side without palsy. As the curve tensors and face scan align, we can directly compare the points of Γ_{left} and $\bar{\Gamma}_{\text{left}}$. Due to the equidistant spline sampling, this pointwise correspondence enables us to indicate where the palsy and contralateral sides differ.

For a 3D visualization, we create a vector field $\mathcal{V}$ between pairwise points from Γ_{left} to $\bar{\Gamma}_{\text{left}}$. The vector length measures the side disparities. As the radial curves γ_α lie on the face's surface, the vectors also originate on the surface. For the visualization, we use a sequential colormap (Imola) without dark shades to reduce interference with shadows introduced by the render engine [22, 73], as depicted in Fig. 6a. We set the upper limit of the color range to 1 cm to ensure comparable visualizations between the different time steps. This visualization aids doctors in identifying the affected areas of the palsy side.

Additionally, we project the volumetric differences onto the facial surface, see Fig. 6b. This visualization is more intuitive for doctors, who can more easily identify the affected areas without interpreting a vector field or interacting with the 3D renderer. We use identical color mapping for the vector field visualization. The projection requires computing the intersection between $\mathcal{V}$ and the facial surface mesh $\mathcal{M}_{\text{left}}$. The vertex colors are then assigned to the corresponding vector length color range value. This approach is similar to [1], indicating the differences between time steps but does not require registration of the facial surface meshes. For example, Fig. 6b shows that the chin area is affected more than the cheek area.

5 Volume analysis for facial expressions

3D facial analysis relies on 3D landmarks to interact with the unstructured 3D point cloud or scan. Therefore, we assess the reliability of 3D landmark extraction across a video sequence by analyzing the consistency of head rotation angles. We observe that head rotation angle consistency is a robust indicator of landmark stability, forming a reliable foundation for subsequent volume estimation. Next, we compare the static *neutral* and *happy* expressions of 34 patients, highlighting the volumetric differences between the left and right sides of their faces. The volumetric differences are compared at the beginning and end of ten-day biofeedback therapy. This case study allows us to check whether our proposed algorithm could be used to measure therapy success objectively. Lastly, the method's applicability is extended by analyzing dynamic changes occurring during single expressions.

5.1 3D landmark reliability

The primary objective is to analyze volumetric changes during dynamic movements. Ensuring the reliability of 3D landmarks throughout the video sequence is essential. Hence, we assess the stability of the 3D landmarks by analyzing the smoothness of the head rotation angles. As a first step, the visualized landmark positions qualitatively reveal each pre-

Fig. 7 3D landmarks for the *happy* expression are obtained from five predictors across six patients. Due to their significance in computing head rotation, we emphasize the nose tip, eye contour, and mouth landmarks. These landmarks are accurately positioned on asymmetrical faces without requiring manual intervention or predictor fine-tuning. The visualization is normalized such that the palsy side is on the left side of the face (right side of the image). (Best viewed digitally.)

dictor's performance, as seen in Fig. 7. The nose tip, eye contour, and mouth landmarks are emphasized, given their significance for computing head rotation. Should these landmarks deviate strongly from their expected positions, head rotation estimates become unreliable, potentially introducing bias into volume estimations.

As demonstrated in Fig. 7, the original method detailed by [22, 23] reliably positions the landmarks at their expected locations. Slight differences in the placement of the landmarks are visible between the different predictors. For the eye contour landmarks, the *BU-3DFE* scheme [59] and the 68 scheme of *face-alignment* [66] place the landmarks on the outer eye corner. In contrast, the other extractors place landmarks in the inner eyelid. Of all the predictors assessed, only *Dlib* [74] encounters difficulty accurately fitting landmarks around the mouth area. *Mediapipe* [15] demonstrates consistency in landmark placement around the mouth area. This approach is possibly attributable to the numerous landmarks that enforce local stability.

The reliability of 3D landmarks is quantified by calculating the smoothness of head rotation across different predictors. We define smoothness as the absolute difference between the original angle signal and a smoothed signal obtained by applying a Gaussian kernel with $\sigma = 5$. We follow [67] to compute the head rotation angles based on the 3D landmarks. The resulting time series for yaw, pitch, and roll angles are depicted in Fig. 8. This metric allows for the quantification of landmark stability throughout the video. Stable landmarks will yield a low deviation from the smoothed signal, approaching zero. If the angle time series deviates from the smooth signal, the landmarks shift their position between consecutive frames. Table 1 presents the mean and standard deviation of rotation smoothness for each predictor across all three rotational angles for all recorded video sequences. The quantitative results corroborate the visual assessment, indi-

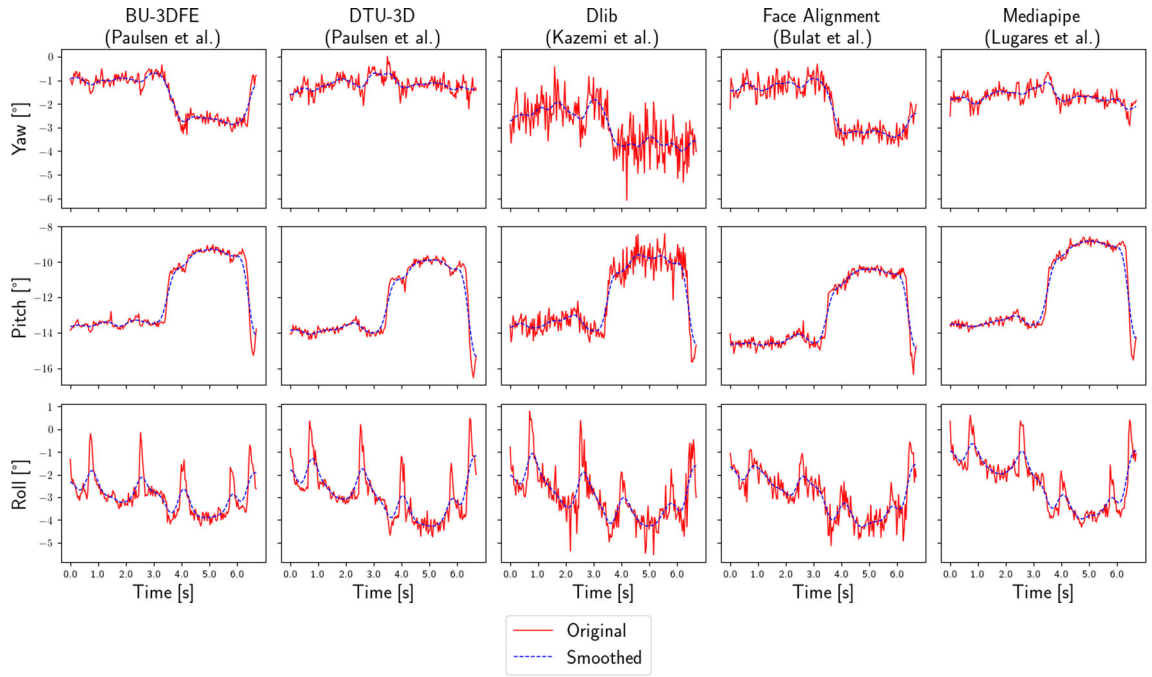


Fig. 8 The original and smoothed rotation angles (yaw, pitch, and roll) for the different predictors for a single patient. The original method by [22, 23] and *Mediapipe* [15] are the most stable among the investigated predictors. However, *Dlib* [74] has the most visible noise in the rotation angles and indicates that the landmarks are not stable throughout the video (best viewed online.)

Table 1 The mean and standard deviation of rotation smoothness are calculated for the various predictors

	BU-3DFE [23, 59]	DTU3D [23, 64]	Dlib [74]	Face-alignment [66]	Mediapipe [15]
Pitch	0.17 ± 0.04	0.26 ± 0.12	0.68 ± 0.53	0.22 ± 0.02	0.19 ± 0.03
Roll	0.22 ± 0.09	0.32 ± 0.09	0.58 ± 0.19	0.34 ± 0.03	0.25 ± 0.07
Yaw	0.14 ± 0.05	0.25 ± 0.14	0.70 ± 0.44	0.30 ± 0.05	0.17 ± 0.03

Rotation smoothness is the absolute difference between the original and smoothed angles obtained using a Gaussian kernel with $\sigma = 5$. A low score on this measure implies that the landmark predictor provides stable landmark placement throughout the video

cating that the original method by [22, 23] and *Mediapipe* [15] exhibit the smoothest performance among the investigated predictors. High mean and standard deviation values for rotation smoothness for *Dlib* [74] confirm the prior visual observation of incorrect placement.

Computation of volumetric differences between the lateral faces builds on head rotation measurements. A typical assumption is constant head rotation during an exercise, yet the angle graphs in Fig. 8 exhibit a notable rotation drift around the three-second mark. This drift corresponds with the moment when the patient begins to mimic the *happy* expression. The rotation drift is observed across all predictors, suggesting it likely stems from the underlying estimation algorithm [67]. The mouth landmarks contribute to head rotation estimation; the computed pitch angle increases as the mouth corners move upwards during expressions like the *happy* expression. Nevertheless, the effect on the pitch angle does not significantly impact the estimation of lateral face volume since the face halves are divided based on the roll angle. The roll angle rotation shows minor drift during the video sequence; see Table 1. Figure 8 reveals five spikes in the roll angle occurring within all predictors, likely attributable to blinking events. The impact of angle deviation on volume estimation can be disregarded since the deviation is minor (approximately 2 degrees). For most facial regions, a 2-degree cone around the face's center along the nose back affects volume estimation merely as a constant factor.

This analysis affirms the reliability of the 3D landmark extraction for the method developed by [22, 23]. Moreover, it is shown that 2D predictors utilizing a multi-view approach are also reliable, thereby endorsing the use of 3D landmarks in volume estimation. Although a method for estimating head rotation without relying on landmarks would be advantageous, the current approach is reliable, and the volume estimation remains unaffected by minor drifts in head

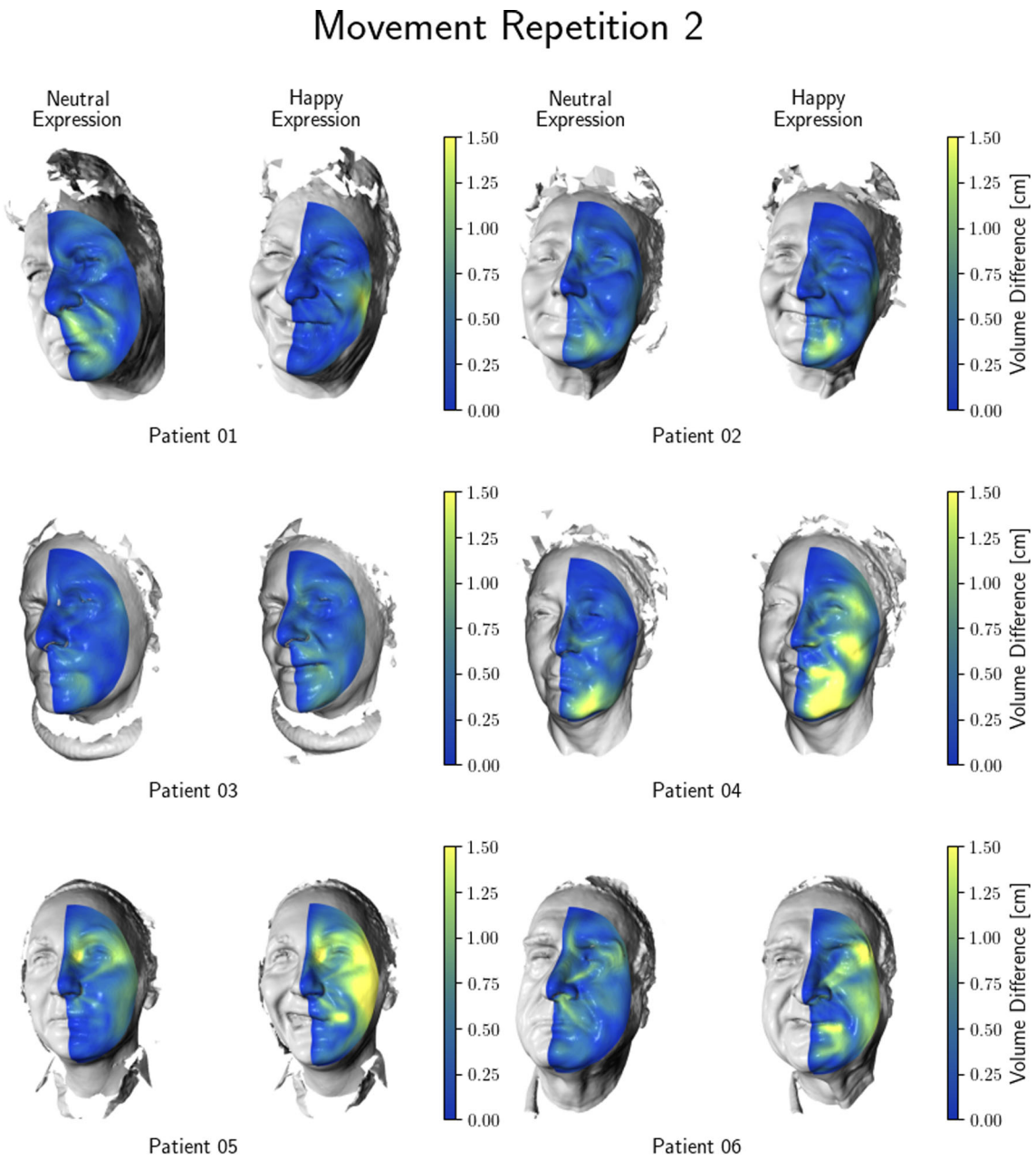


Fig. 9 The volumetric disparities between the left and right sides of the face are displayed for the six facial palsy patients for the **second** measurement. Our visualizations are normalized such that the palsy-affected side is positioned on the left (right side of the image), and the maximal difference is 1.5 cm

rotation. In subsequent experiments, the *BU-3DFE* landmark scheme is utilized for the head rotation estimation [8, 23, 59].

5.2 Static facial volume analysis

Having confirmed the reliability of the 3D landmarks, we examine the volumetric differences between the *neutral* and *happy* expressions. Therefore, we analyze our proposed method as an objective evaluation tool in a case study of 34 patients undergoing a ten-day biofeedback therapy [63]. Static frames are automatically selected at 1.5 s post-instruction and 1.5 s before the conclusion of the *happy* expression to achieve comparable results across patients. The volumetric disparities between the lateral facial sides are depicted in Fig. 9 for six patients, utilizing radial curves to project the differences in volume [8, 22]. The second measurement of the four replicates per patient is presented; the remaining measurements are included in the appendix (B). All visualizations are normalized to present the palsy-

affected side on the left and set the maximum difference to 1.5 cm. We used $n = 64$ curves and $m = 64$ spline points.

The mean volumetric differences between the *neutral* and *happy* expressions, aggregated over all four repetitions at the beginning and end of therapy, are listed in Table 2. Although these variations are patient-dependent, we note considerable and observable volumetric differences between the face sides during happy expressions at the beginning ($13.8 \pm 10.0\,\text{mm}^3$) and end ($12.8 \pm 10.3\ \text{mm}^3$) of the ten-day biofeedback therapy. Furthermore, we performed a Wilcoxon signed-rank test to assess the effect of the patient training [75–77]. The results of this test indicate that the volume changes are statistically significant ($p = 0.0305$, $z = -2.1627$, $r = 0.3709$). The neutral face kept a consistent volume range of $11.8-12.1\ \text{mm}^3$. The post-therapy reduction in volumetric differences suggests decreased facial asymmetry during movement, providing valuable insights for monitoring and informing treatment decisions.

However, these values should be interpreted cautiously, especially when analyzing only a single frame during a dynamic facial expression. Factors such as frame selection, patient non-compliance with movement, failure to relax into a neutral state, or variance in the happy expression could influence the results. Especially, patients might also learn to activate their unaffected face side less to achieve a more symmetrical face. Although the provided value alone does not disclose the underlying cause for reduced volumetric difference, it is still an essential marker of facial symmetry. Additionally, since our current analysis focuses on the absolute difference between sides of the face, areas may negate each other's changes. This can be qualitatively observed especially for patient four, see Fig. 9. As the mouth is open during the *happy* expression, the total approximated facial volume decreases, reducing the volumetric difference. This occurrence could explain the reduction in volumetric difference for patient four but the high visual disparity in the volumetric difference, see Fig. 9.

Some of the limitations of static analysis may be mitigated by dynamic analysis, which can offer an enhanced understanding of the effects of facial palsy on facial volume. Thus, we avoid the dependence on a single frame and the potential impact of frame selection by using the entire video sequence. Yet the static analysis provides a foundation for the dynamic analysis, as it allows us to identify the areas of the face most affected by facial palsy. The combination of radial curves and semantic analysis via 3D morphable models could enhance the localization of the volumetric differences. Additionally, we abstain from medically interpreting the results, given that the study is not intended to offer medical recommendations.

5.3 Dynamic facial volume analysis

Lastly, we are interested in the dynamic analysis of the facial volume during an instructed movement, as static analysis may not fully capture the impact of facial palsy on facial volume. Thus, we analyzed a patient's volumetric changes and locations mimicking facial expressions as a case study. During dynamic movements, temporal noise might occur from the 3dMD camera system (3dMD LCC, Georgia, USA), which propagates to the 3D landmarks, impacting head rotation and lateral volume estimation. Thus, we apply a sliding window of five frames to the 3D landmarks to reduce the impact on subsequent processing steps to remove the noise as displayed in Fig. 7. The radial curve range is set to $r = 85$ mm, the distance between the nose tip and the chin. Patients were instructed to mimic a *happy* facial expression after a *neutral* phase that lasted about three seconds. We visualize the time progression in Fig. 10. The top row displays four face scans, the lateral meshes, and the projected volumetric differences of a patient during selected time steps. The second row illustrates the volume of the left (blue) and right (red) sides of the face, as well as the absolute volume difference between the two sides (black). The third row shows the averaged one-to-one pointwise distance of $\mathcal{V}$ between the affected and contralateral sides [8].

The initial difference might be due to inherent facial asymmetries, muscle atrophy or compensatory hypertrophy, a mixture of both, or inaccuracy in the method. However, as it remains constant during the neutral phase, we assume our method is stable throughout dynamic movements. During the movement phase, the volume difference increases, indicating increasing facial asymmetry for both measurement methods. Notably, volumetric analysis enables insight into the impact of both sides. The palsy side mesh (left/blue) decreases in volume, whereas the healthy side (right/red) experiences an increase in volume. A change is expected as the facial muscles contract [10]. The volume reduction on the palsy side could be due to tissue pulled towards the healthy side. The facial muscles of both sides form a more extensive interwoven network. Suppose one side (the healthy contralateral side) is contracting stronger. In that case, this automatically pulls over the facial soft tissue (that we observe as volume change) from the palsy side to the healthy side, as the contraction on the healthy side is not counterbalanced by a symmetric activity on the palsy side.

For patient two, the observed volume difference increases during movement, signaling a rise in facial asymmetry. This observation aligns with the static analysis in Table 2. The dynamic analysis reveals that selecting a single frame for

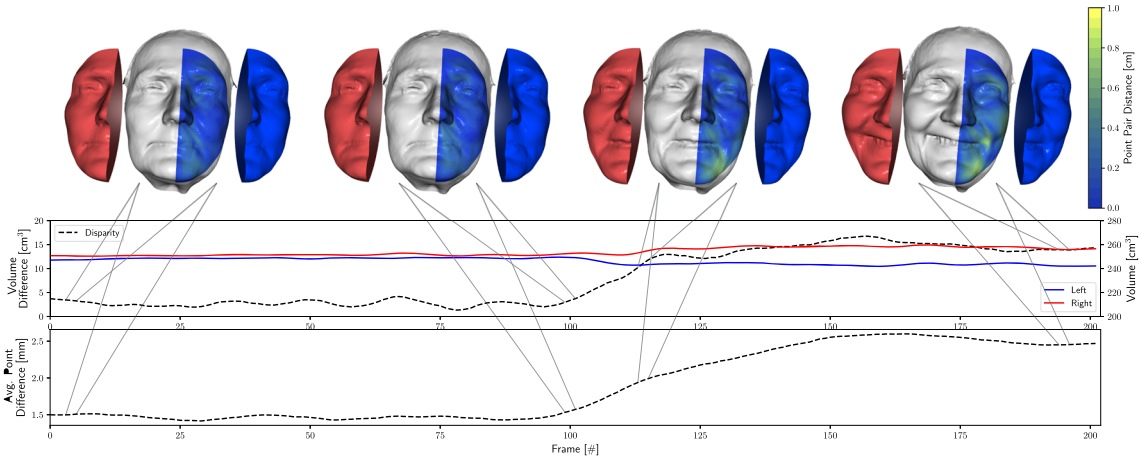


Fig. 10 We measure the volumetric changes from a `neutral` to a `happy` facial expression. The palsy side (blue) decreases while the healthy side (red) increases in volume, indicating a shift from the palsy side to the healthy side [22] (colour figure online)

static analysis may not fully capture the extent of facial palsy's impact on facial volume. Thus, dynamic analysis is essential for a comprehensive understanding of the impact of facial palsy on facial volume and should be considered in future studies. Our method exhibits similar behavior to the pointwise distance [8], indicating that it measures the same asymmetry. However, pointwise distance cannot reveal the palsy side's impact on the healthy side, making volumetric analysis more suitable for assessing this effect. Based on pointwise disparities, our proposed visualization effectively illustrates the volumetric differences between the healthy and palsy sides. The dynamic analysis provides more insight into facial volume changes. However, the combination of radial curves and semantic analysis via 3D morphable models could enhance the localization of the volumetric differences.

6 Conclusions and future work

We introduced a method for calculating volumetric facial disparities, improving on existing approaches and providing insights into the palsy side's impact. In addition, our approach offers visualization of volumetric differences, enhancing understanding of facial asymmetry. Moreover, we can analyze changes during a single movement, extending the method's usefulness and bridging the gap from 2D static to 3D dynamic analysis.

As our approach is automatic and requires minimal parameter tuning (number of radial curves n, spline points m, and δ based on the sensor), we do not rely on any assumptions and estimations about facial symmetry. The joint visualization, see Fig. 10, of the volume differences, the facial expression, and the difference heatmap helps to understand the behavior of the facial muscles during a single movement. The projected volumetric differences between the meshes, see Fig. 6b and 9, can be used to identify the affected areas of the face in a static scan. Visualizations guide treatment decisions during the clinical routine, and our approach can help doctors better understand facial asymmetry. Frame selection is crucial; our dynamic analysis provides a more comprehensive understanding of facial palsy's impact on facial volume. This insight is significant for treating facial palsy, as the muscle contractions create facial expressions. Our work offers a combined tool to analyze facial volume changes during dynamic movements, especially in facial palsy. We open up future medical research to define instance exercises that address counter-actions to volume shifts towards the healthy side [1, 36].

The effectiveness of our approach depends on 3D scan quality and intermediate facial landmark estimation, which are only dependent on pre-existing 2D landmark estimation. The growing accessibility of 3D sensors allows for the application of our approach to a broader patient population with facial palsy. We plan to integrate radial curves with 3D morphable models to discern patient-specific facial muscle behavior, enabling automated detection of affected facial regions. This integration would also negate the need to estimate head rotation, which currently has a minimal effect on volume estimation.

Appendix A Volumetric changes

See Table 2.

Table 2 We list the absolute lateral volume difference for the *neutral* and *happy* expressions for all 34 patients

ID	Begin therapy		End therapy	
	Neutral	Emotion	Neutral	Emotion
01	6.2 ± 4.0	8.0 ± 3.2	12.5 ± 4.2	5.3 ± 2.0
02	10.3 ± 0.8	2.2 ± 1.3	5.5 ± 1.3	5.3 ± 1.7
03	8.0 ± 0.9	17.3 ± 0.8	7.8 ± 3.4	17.1 ± 1.4
04	26.8 ± 1.1	9.7 ± 2.9	20.0 ± 0.7	6.5 ± 0.8
05	10.7 ± 2.5	8.6 ± 1.9	9.3 ± 5.0	2.1 ± 1.2
06	3.5 ± 1.9	12.8 ± 1.5	3.4 ± 1.4	8.0 ± 2.1
08	3.3 ± 2.1	3.4 ± 1.9	1.9 ± 0.9	1.7 ± 0.8
09	33.0 ± 1.3	17.2 ± 2.0	25.4 ± 1.7	16.0 ± 2.4
10	35.5 ± 2.3	28.2 ± 1.5	38.3 ± 1.2	24.0 ± 2.9
11	16.2 ± 2.1	5.3 ± 2.4	18.5 ± 1.2	2.2 ± 0.8
12	22.0 ± 2.4	22.6 ± 1.2	25.4 ± 2.3	23.7 ± 4.0
13	23.4 ± 1.8	19.1 ± 2.3	18.1 ± 6.6	20.3 ± 7.9
14	3.2 ± 0.5	8.3 ± 3.3	3.7 ± 0.8	5.9 ± 2.3
15	2.5 ± 1.5	11.2 ± 3.6	2.3 ± 1.4	14.2 ± 1.4
16	8.8 ± 2.7	19.9 ± 2.4	7.2 ± 1.9	17.7 ± 4.4
17	3.1 ± 1.0	15.0 ± 1.5	13.6 ± 3.2	7.1 ± 1.7
18	13.0 ± 1.7	29.3 ± 1.1	6.1 ± 3.0	26.0 ± 1.7
19	15.4 ± 5.0	16.3 ± 3.0	18.2 ± 2.2	9.9 ± 1.9
20	3.6 ± 1.0	11.6 ± 1.0	2.0 ± 0.9	11.4 ± 3.9
21	4.9 ± 3.3	10.0 ± 5.2	8.1 ± 1.7	13.5 ± 3.1
22	10.2 ± 1.1	3.2 ± 1.8	9.7 ± 1.7	2.0 ± 0.7
23	2.0 ± 1.3	2.3 ± 0.8	2.2 ± 0.5	5.1 ± 1.2
24	19.5 ± 3.5	28.7 ± 2.4	19.5 ± 1.3	24.4 ± 0.3
25	13.2 ± 6.5	32.8 ± 3.9	11.3 ± 0.6	23.0 ± 0.6
26	10.7 ± 1.9	11.3 ± 2.8	16.0 ± 2.5	20.1 ± 3.2
27	1.9 ± 0.7	2.7 ± 1.5	3.3 ± 0.6	2.3 ± 0.6
28	28.0 ± 4.2	44.6 ± 1.5	41.4 ± 1.8	48.3 ± 2.3
29	4.1 ± 1.5	7.3 ± 2.2	3.4 ± 2.0	4.0 ± 1.4
30	12.9 ± 1.6	5.6 ± 2.7	11.0 ± 1.7	11.5 ± 1.8
31	5.6 ± 3.1	18.7 ± 2.7	2.8 ± 0.7	14.2 ± 0.9
32	9.7 ± 1.5	2.7 ± 1.5	12.1 ± 3.1	4.0 ± 3.0
33	4.1 ± 1.8	14.2 ± 1.1	6.1 ± 2.3	8.3 ± 2.6
34	15.9 ± 3.9	3.0 ± 1.8	17.7 ± 1.1	2.1 ± 1.3
35	9.2 ± 3.2	13.9 ± 3.2	3.9 ± 4.2	10.0 ± 0.7
ϕ	11.8 ± 9.2	13.7 ± 10.0	12.1 ± 9.7	12.8 ± 10.3

All values are presented in cubic millimeters, aggregated from all four instances of the movement repetition (given as standard deviations per patient). We observe a reduction in the volumetric difference at begin and end of the ten-day biofeedback therapy, indicating possible treatment success. Note that we omit patient 07 as no valid scan was obtained at the treatment

Appendix B Static volume

See Figs. 11, 12 and 13.

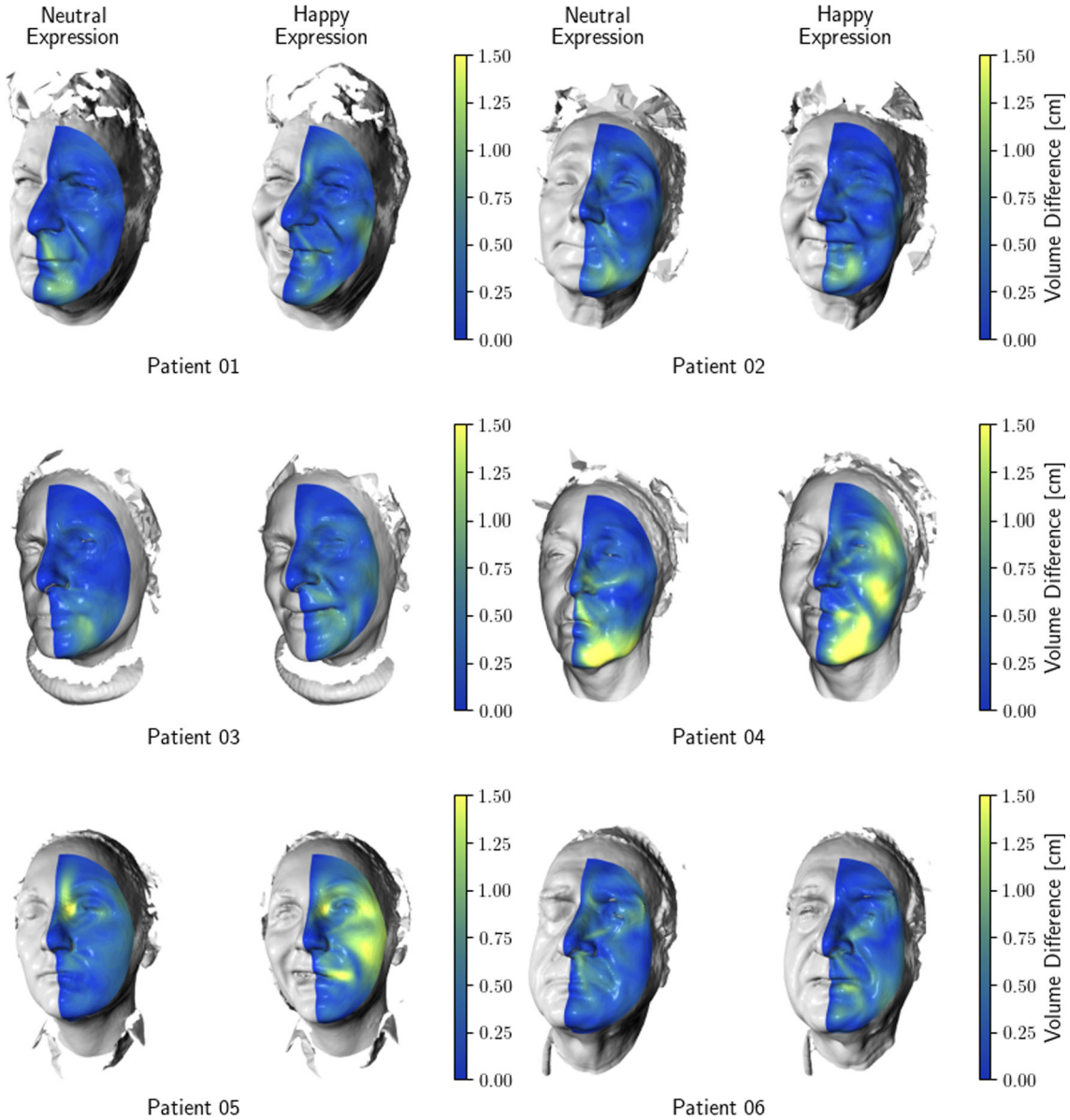


Fig. 11 The volumetric disparities between the left and right sides of the face are displayed for the six facial palsy patients for the **first** measurement. Our visualizations are normalized such that the palsy-affected side is positioned on the left (right side of the image), and the maximal difference is 1.5 cm

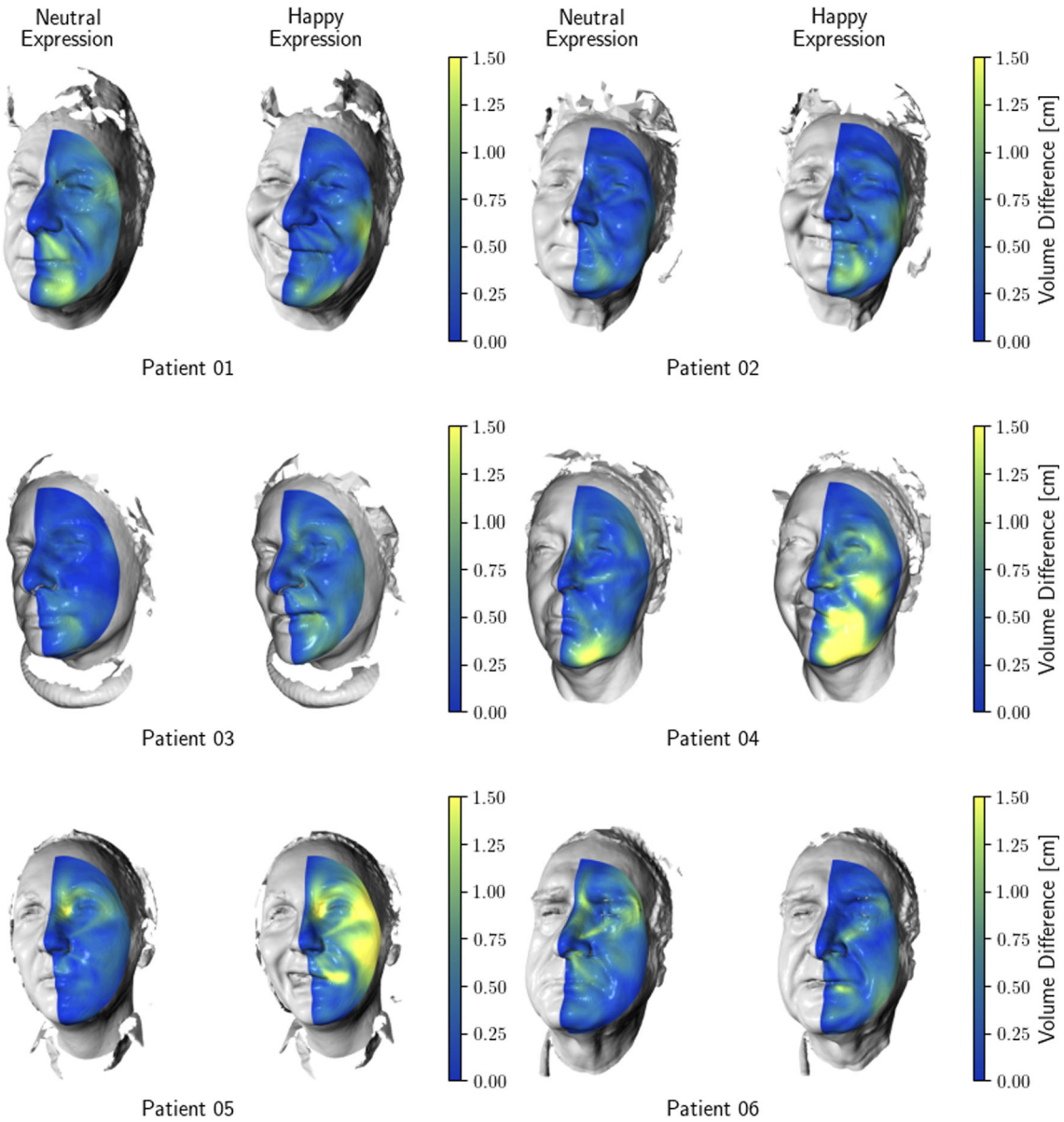


Fig. 12 The volumetric disparities between the left and right sides of the face are displayed for the six facial palsy patients for the **third** measurement. Our visualizations are normalized such that the palsy-affected side is positioned on the left (right side of the image), and the maximal difference is 1.5 cm

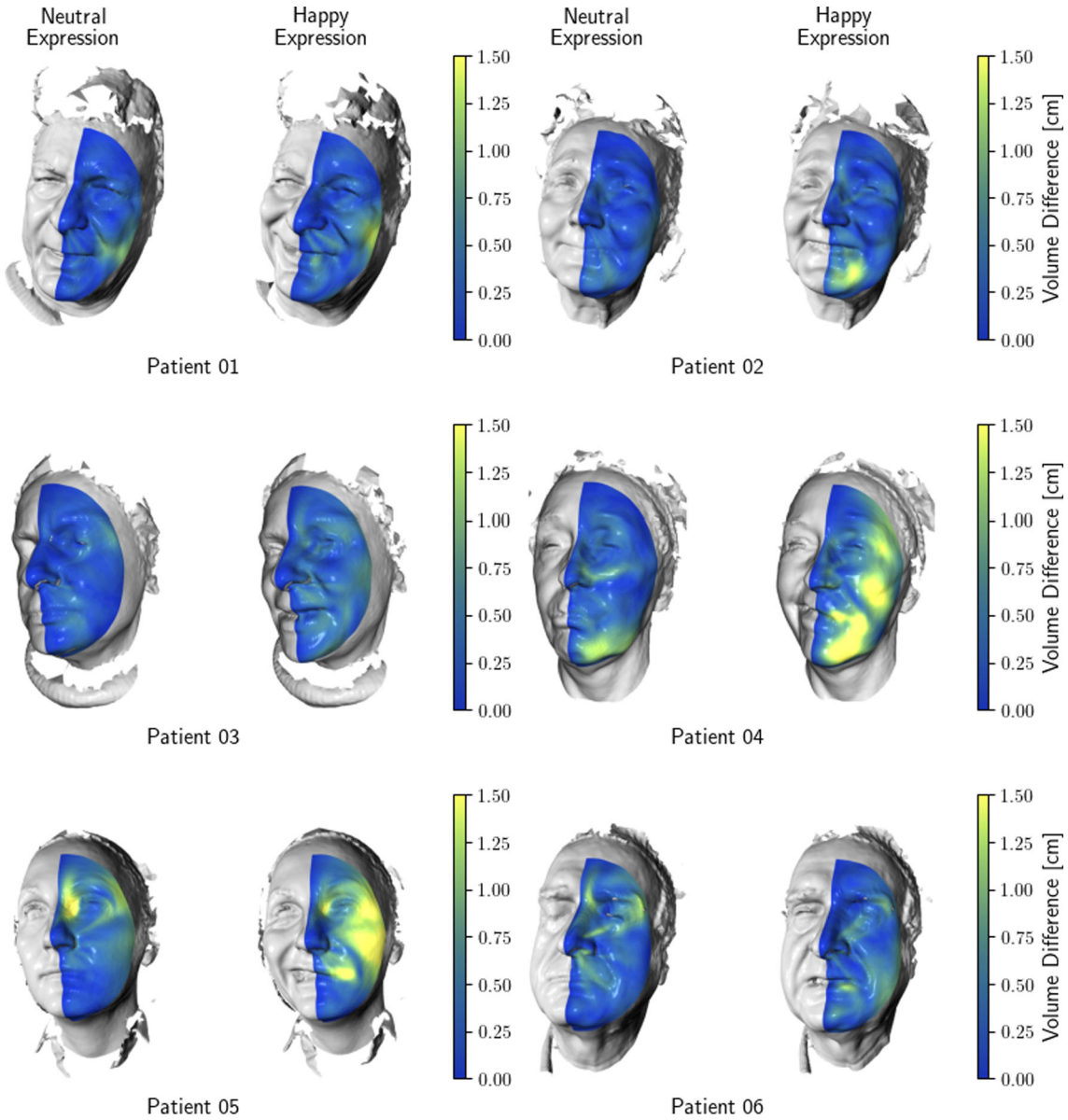


Fig. 13 The volumetric disparities between the left and right sides of the face are displayed for the six facial palsy patients for the **fourth** measurement. Our visualizations are normalized such that the palsy-affected side is positioned on the left (right side of the image), and the maximal difference is 1.5 cm

Acknowledgements Supported by Deutsche Forschungsgemeinschaft (DFG - German Research Foundation) project 427899908 BRIDGING THE GAP: MIMICS AND MUSCLES (DE 735/15-1 and GU 463/12-1).

Author Contributions Conceptualization: TB, JD, OG-L; Methodology: TB; Formal analysis and investigation: TB; Writing—original draft preparation: TB; Writing—review and editing: TB, SS, GFV, JD, OG-L; Funding acquisition: JD, OG-L; Resources: GFV, OG-L; Supervision: SS, JD, OG-L.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availibility The data generated during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Consent to participate The ethics committee of the Jena University Hospital approved the study (No. 2019-1539)

Consent for publication Written informed consent was obtained from all participants.

Code availability The code used to generate the presented results is available at https://github.com/cvjena/corc and https://github.com/cvjena/mvlm.

References

1. Volk, G.F., Thümmel, M., Mothes, O., Arnold, D., Thielker, J., Denzler, J., Mastryukova, V., Mayr, W., Guntinas-Lichius, O.: Long-term home-based surface electrostimulation is useful to prevent atrophy in denervated facial muscles. In: Vienna Workshop on Functional Electrical Stimulation (FESWS) (2019)
2. Lou, J., Yu, H., Wang, F.-Y.: A review on automated facial nerve function assessment from visual face capture. IEEE Trans. Neural Syst. Rehabil. Eng. **28**(2), 488–497 (2020). https://doi.org/10.1109/TNSRE.2019.2961244
3. House, J.W., Brackmann, D.E.: Facial nerve grading system. Otolaryngol. Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg. **93**(2), 146–147 (1985). https://doi.org/10.1177/019459988509300202
4. Gaber, A., Taher, M.F., Wahed, M.A.: Quantifying facial paralysis using the Kinect v2. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2015, 2497–2501 (2015) https://doi.org/10.1109/EMBC.2015.7318899
5. Büchner, T., Penzel, N., Guntinas-Lichius, O., Denzler, J.: The power of properties: Uncovering the influential factors in emotion classification. In: Pattern Recognition and Artificial Intelligence (2024). https://doi.org/10.48550/arXiv.2404.07867 . arXiv:2404.07867
6. Steinhäuser, J., Volk, G.F., Thielker, J., Geitner, M., Kuttenreich, A.-M., Klingner, C.M., Dobel, C., Guntinas-Lichius, O.: Multidisciplinary care of patients with facial palsy: treatment of 1220 patients in a German facial nerve center. J. Clin. Med. **11**(2), 427 (2022). https://doi.org/10.3390/jcm11020427
7. Ali, W., Imran, M., Yaseen, M.U., Aurangzeb, K., Ashraf, N., Aslam, S.: A transfer learning approach for facial paralysis severity detection. IEEE Access **11**, 127492–127508 (2023). https://doi.org/10.1109/ACCESS.2023.3330242
8. Büchner, T., Sickert, S., Volk, G.F., Guntinas-Lichius, O., Denzler, J.: Automatic objective severity grading of peripheral facial palsy using 3D radial curves extracted from point clouds. In: Challenges of Trustable AI and Added-Value on Health, 179–183 (2022) https://doi.org/10.3233/SHTI220433
9. Özsoy, U., Sekerci, R., Hizay, A., Yildirim, Y., Uysal, H.: Assessment of reproducibility and reliability of facial expressions using 3D handheld scanner. J. Cranio-Maxillo-Fac. Surg. Off. Publ. Eur. Assoc. Cranio-Maxillo-Fac. Surg. **47**(6), 895–901 (2019). https://doi.org/10.1016/j.jcms.2019.03.022
10. Rawlani, R., Qureshi, H., Rawlani, V., Turin, S.Y., Mustoe, T.A.: Volumetric changes of the mid and lower face with animation and the standardization of three-dimensional facial imaging. Plast. Reconstr. Surg. **143**(1), 76–85 (2019). https://doi.org/10.1097/PRS.0000000000005082
11. Guntinas-Lichius, O., Trentzsch, V., Mueller, N., Heinrich, M., Kuttenreich, A.-M., Dobel, C., Volk, G.F., Graßme, R., Anders, C.: High-resolution surface electromyographic activities of facial muscles during the six basic emotional expressions in healthy adults: A prospective observational study. Sci. Rep. **13**(1), 19214 (2023). https://doi.org/10.1038/s41598-023-45779-9
12. ten Harkel, T.C., Vinayahalingam, S., Ingels, K.J.A.O., Bergé, S.J., Maal, T.J.J., Speksnijder, C.M.: Reliability and agreement of 3D anthropometric measurements in facial palsy patients using a low-cost 4D imaging system. IEEE Trans. Neural Syst. Rehabil. Eng. **28**(8), 1817–1824 (2020). https://doi.org/10.1109/TNSRE.2020.3007532
13. Li, T., Bolkart, T., Black, M.J., Li, H., Romero, J.: Learning a model of facial shape and expression from 4D scans. ACM Trans. Gr. **36**(6), 1–17 (2017). https://doi.org/10.1145/3130800.3130813
14. Paysan, P., Knothe, R., Amberg, B., Romdhani, S., Vetter, T.: A 3D face model for pose and illumination invariant face recognition. In: 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 296–301. IEEE, Genova (2009). https://doi.org/10.1109/AVSS.2009.58
15. Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M.G., Lee, J., Chang, W.-T., Hua, W., Georg, M., Grundmann, M.: MediaPipe: A Framework for Building Perception Pipelines. arXiv (2019). https://doi.org/10.48550/arXiv.1906.08172
16. Egger, B., Smith, W.A.P., Tewari, A., Wuhrer, S., Zollhoefer, M., Beeler, T., Bernard, F., Bolkart, T., Kortylewski, A., Romdhani, S., Theobalt, C., Blanz, V., Vetter, T.: 3D morphable face models-past, present, and future. ACM Trans. Gr. **39**(5), 157–115738 (2020). https://doi.org/10.1145/3395208
17. Li, R., Bladin, K., Zhao, Y., Chinara, C., Ingraham, O., Xiang, P., Ren, X., Prasad, P., Kishore, B., Xing, J., : Learning formation of physically-based face attributes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3410–3419 (2020)
18. Cao, Chen, Weng, Yanlin, Zhou, Shun, Tong, Yiying, Zhou, Kun: Facewarehouse: a 3D facial expression database for visual computing. IEEE Trans. Vis. Comput. Gr. **20**(3), 413–425 (2014). https://doi.org/10.1109/TVCG.2013.249
19. Berretti, S., Del Bimbo, A., Pala, P., Mata, F.J.S.: Face recognition by svms classification of 2d and 3d radial geodesics. In: 2008 IEEE International Conference on Multimedia and Expo, pp. 93–96. IEEE, Hannover (2008). https://doi.org/10.1109/ICME.2008.4607379
20. Thümmel, M., Sickert, S., Denzler, J.: Facial behavior analysis using 4d curvature statistics for presentation attack detection. In: 2021 IEEE International Workshop on Biometrics and Forensics (IWBF), pp. 1–6 (2021). https://doi.org/10.1109/IWBF50991.2021.9465080
21. Bowman, A.W., Katina, S., Smith, J., Brown, D.: Anatomical curve identification. Comput. Stat. Data Anal. **86**, 52–64 (2015). https://doi.org/10.1016/j.csda.2014.12.007
22. Büchner, T., Sickert, S., Volk, G.F., Guntinas-Lichius, O., Denzler, J.: From faces to volumes - measuring volumetric asymmetry in 3D facial palsy scans. In: Bebis, G., Ghiasi, G., Fang, Y., Sharf, A., Dong, Y., Weaver, C., Leo, Z., LaViola Jr., J.J., Kohli, L. (eds.) Advances in Visual Computing. Lecture Notes in Computer Science, pp. 121–132. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-47969-4_10
23. Paulsen, R.R., Juhl, K.A., Haspang, T.M., Hansen, T., Ganz, M., Einarsson, G.: Multi-view consensus CNN for 3D facial landmark placement. arXiv:1910.06007 [cs] 11361, 706–719 (2019) https://doi.org/10.1007/978-3-030-20887-5_44
24. Raj, A., Mothes, O., Sickert, S., Volk, G.F., Guntinas-Lichius, O., Denzler, J.: Automatic and objective facial palsy grading index prediction using deep feature regression. In: Papież, B.W., Namburete,

A.I.L., Yaqub, M., Noble, J.A. (eds.) Medical Image Understanding and Analysis, vol. 1248, pp. 253–266. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-52791-4_20
25. Haase, D., Minnigerode, L., Volk, G.F., Denzler, J., Guntinas-Lichius, O.: Automated and objective action coding of facial expressions in patients with acute facial palsy. Eur. Arch. Otorhinolaryngol. **272**(5), 1259–1267 (2015). https://doi.org/10.1007/s00405-014-3385-8
26. Zhuang, Y., McDonald, M.M., Aldridge, C.M., Hassan, M.A., Uribe, O., Arteaga, D., Southerland, A.M., Rohde, G.K.: Video-based facial weakness analysis. IEEE Trans. Biomed. Eng. **68**(9), 2698–2705 (2021). https://doi.org/10.1109/TBME.2021.3049739
27. Sakai, T., Seo, M., Matsushiro, N., Chen, Y.-W.: Simulation of facial palsy using conditional generative adversarial networks and face shape normalization. In: 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), pp. 793–797 (2021). https://doi.org/10.1109/GCCE53005.2021.9621817
28. Gaber, A., Taher, M.F., Abdel Wahed, M., Shalaby, N.M., Gaber, S.: Comprehensive assessment of facial paralysis based on facial animation units. PLoS One **17**(12), 0277297 (2022). https://doi.org/10.1371/journal.pone.0277297
29. Miller, M.Q., Hadlock, T.A., Fortier, E., Guarin, D.L.: The auto-eFACE: machine learning-enhanced program yields automated facial palsy assessment tool. Plast. Reconstr. Surg. **147**(2), 467–474 (2021). https://doi.org/10.1097/PRS.0000000000007572
30. Sajid, M., Shafique, T., Baig, M., Riaz, I., Amin, S., Manzoor, S.: Automatic grading of palsy using asymmetrical facial features: a study complemented by new solutions. Symmetry **10**(7), 242 (2018). https://doi.org/10.3390/sym10070242
31. Rodríguez Martínez, E.A., Polezhaeva, O., Marcellin, F., Colin, É., Boyaval, L., Sarhan, F.-R., Dakpé, S.: Deepsmile: anomaly detection software for facial movement assessment. Diagnostics **13**(2), 254 (2023). https://doi.org/10.3390/diagnostics13020254
32. Gaber, A., Taher, M.F., Wahed, M.A., Shalaby, N.M., Gaber, S.: Classification of facial paralysis based on machine learning techniques. Biomed. Eng. Online **21**(1), 65 (2022). https://doi.org/10.1186/s12938-022-01036-0
33. Ross, B.G., Fradet, G., Nedzelski, J.M.: Development of a sensitive clinical facial grading system. Otolaryngol. Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg. **114**(3), 380–386 (1996). https://doi.org/10.1016/S0194-59989670206-1
34. Büchner, T., Sicker, S., Volk, G.F., Anders, C., Guntinas-Lichius, O., Denzler, J.: Let's get the facs straight: Reconstructing obstructed facial features. In: Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2023) - Volume 4: VISAPP, pp. 727–736 (2023). https://doi.org/10.5220/0011619900003417 . INSTICC
35. Büchner, T., Guntinas-Lichius, O., Denzler, J.: Improved obstructed facial feature reconstruction for emotion recognition with minimal change cyclegans. In: Blanc-Talon, J., Delmas, P., Philips, W., Scheunders, P. (eds.) Advanced Concepts for Intelligent Vision Systems. Lecture Notes in Computer Science, pp. 262–274. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-45382-3_22
36. Arnold, D., Thielker, J., Klingner, C.M., Puls, W.C., Misikire, W., Guntinas-Lichius, O., Volk, G.F.: Selective surface electrostimulation of the denervated zygomaticus muscle. Diagnostics **11**(2), 188 (2021). https://doi.org/10.3390/diagnostics11020188
37. Diederiks, V.L.: Multichannel Surface EMG and Machine Learning for Classification of Facial Expressions. University of Twente (2021)
38. Boonipat, T., Asaad, M., Lin, J., Glass, G.E., Mardini, S., Stotland, M.A.: Using artificial intelligence to measure facial expression following facial reanimation surgery. Plast. Reconstr. Surg. **146**, 1147–1150 (2020)
39. Guarin, D.L., Yunusova, Y., Taati, B., Dusseldorp, J.R., Mohan, S., Tavares, J., van Veen, M.M., Fortier, E., Hadlock, T.A., Jowett, N.: Toward an automatic system for computer-aided assessment in facial palsy. arXiv:1910.11497 (2019)
40. Miller, M.Q., Hadlock, T.A., Fortier, E., Guarin, D.L.: The auto-eFACE: machine learning-enhanced program yields automated facial palsy assessment tool. Plastic & Reconstructive Surgery (2020)
41. Mothes, O., Modersohn, L., Volk, G.F., Klingner, C.M., Witte, O.W., Schlattmann, P., Denzler, J., Guntinas-Lichius, O.: Automated objective and marker-free facial grading using photographs of patients with facial palsy. Eur. Arch. Otorhinolaryngol. **276**, 3335–3343 (2019)
42. Guarin, D.L., Dusseldorp, J.R., Hadlock, T.A., Jowett, N.: A machine learning approach for automated facial measurements in facial palsy. JAMA Fac. Plast. Surg. **20**(4), 335–337 (2018)
43. House, J.W., Brackmann, D.E.: Facial nerve grading system. Otolaryngol. Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg. **93**(2), 146–147 (1985). https://doi.org/10.1177/019459988509300202
44. Zhang, Y., Gao, W., Yu, H., Dong, J., Xia, Y.: Artificial intelligence-based facial palsy evaluation: a survey. IEEE Trans. Neural Syst. Rehabil. Eng. **32**, 3116–3134 (2024). https://doi.org/10.1109/TNSRE.2024.3447881
45. Knoedler, L., Miragall, M., Kauke-Navarro, M., Obed, D., Bauer, M., Tißler, P., Prantl, L., Machens, H.-G., Broer, P.N., Baecher, H., Panayi, A.C., Knoedler, S., Kehrer, A.: A Ready-to-Use Grading Tool for Facial Palsy Examiners-Automated Grading System in Facial Palsy Patients Made Easy. Journal of Personalized Medicine **12**(10), 1739 (2022). https://doi.org/10.3390/jpm12101739
46. Samsudin, W.S.W., Zarini, E.F., Samsudin, W.N.A.W., Shaharum, S.M., Sundaraj, K., Ahmad, M.Z.: Eye blinking assessment of bell's palsy. In: Proceedings of the 7th International Conference on Electrical, Control and Computer Engineering—Volume 2: InECCE 2023, Kuala Lumpur, p. 457. Springer
47. Schuhmann, L., Büchner, T., Heinrich, M., Volk, G.F., Denzler, J., Guntinas-Lichius, O.: Automated analysis of spontaneous eye blinking in patients with acute facial palsy or facial synkinesis. Sci. Rep. **14**(1), 17726 (2024). https://doi.org/10.1038/s41598-024-68707-x
48. Xie, Y., Büchner, T., Schuhmann, L., Guntinas-Lichius, O., Denzler, J.: Unsupervised learning of eye state prototypes for semantically rich blinking detection. In: Digital Health and Informatics Innovations for Sustainable Health Care Systems, pp. 1607–1611. IOS Press (2024)
49. Büchner, T., Penzel, N., Guntinas-Lichius, O., Denzler, J.: The power of properties: uncovering the influential factors in emotion classification. arXiv preprint arXiv:2404.07867 (2024)
50. Reimers, C., Penzel, N., Bodesheim, P., Runge, J., Denzler, J.: Conditional dependence tests reveal the usage of ABCD rule features and bias variables in automatic skin lesion classification. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1810–1819 (2021)
51. Blunk, J., Penzel, N., Bodesheim, P., Denzler, J.: Beyond debiasing: Actively steering feature selection via loss regularization. In: DAGM German Conference on Pattern Recognition, pp. 394–408. Springer, (2023)
52. Piater, T., Penzel, N., Stein, G., Denzler, J.: When medical imaging met self-attention: A love story that didn't quite work out. (2024) arXiv preprint arXiv:2404.12295
53. Jones, C., de Castro, D.C., Ribeiro, F.D.S., Oktay, O., McCradden, M., Glocker, B.: No fair lunch: a causal perspective on dataset bias in machine learning for medical imaging. arXiv:2307.16526 (2023)

54. Vokinger, K.N., Feuerriegel, S., Kesselheim, A.S.: Mitigating bias in machine learning for medicine. Commun. Med. **1**, 25 (2021)
55. Saunders, D., Byrne, B.: Reducing gender bias in neural machine translation as a domain adaptation problem. In: Annual Meeting of the Association for Computational Linguistics (2020)
56. Büchner, T., Penzel, N., Guntinas-Lichius, O., Denzler, J.: Facing Asymmetry - Uncovering the Causal Link between Facial Symmetry and Expression Classifiers Using Synthetic Interventions. arXiv:2409.15927 (2024)
57. Li, X., Grandvalet, Y., Davoine, F.: Explicit inductive bias for transfer learning with convolutional networks. arXiv:1802.01483 (2018)
58. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '99, pp. 187–194. ACM Press (1999). https://doi.org/10.1145/311535.311556
59. Lijun Yin, Xiaozhou Wei, Yi Sun, Jun Wang, Rosato, M.J.: A 3D facial expression database for facial behavior research. In: 7th International Conference on Automatic Face and Gesture Recognition (FGR06), pp. 211–216 (2006). https://doi.org/10.1109/FGR.2006.6
60. Pearson, K.: Liii. On lines and planes of closest fit to systems of points in space. Lond. Edinb. Dublin Philos. Mag. J. Sci. **2**(11), 559–572 (1901). https://doi.org/10.1080/14786440109462720
61. Patel, A., Islam, S.M.S., Murray, K., Goonewardene, M.S.: Facial asymmetry assessment in adults using three-dimensional surface imaging. Prog. Orthod. **16**(1), 36 (2015). https://doi.org/10.1186/s40510-015-0106-9
62. Sarhan, F.-R., Olivetto, M., Ben Mansour, K., Neiva, C., Colin, E., Choteau, B., Marie, J.-P., Testelin, S., Marin, F., Dakpé, S.: Quantified analysis of facial movement: a reference for clinical applications. Clin. Anat. (2023). https://doi.org/10.1002/ca.23999
63. Volk, G.F., Finkensieper, M., Guntinas-Lichius, O.: EMG biofeedback training at home for patient with chronic facial palsy and defective healing. Laryngorhinootologie **93**(1), 15–24 (2014). https://doi.org/10.1055/s-0033-1345166
64. Fagertun, J., Harder, S., Rosengren, A., Moeller, C., Werge, T., Paulsen, R.R., Hansen, T.F.: 3d facial landmarks: inter-operator variability of manual annotation. BMC Med. Imaging **14**(1), 35 (2014). https://doi.org/10.1186/1471-2342-14-35
65. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: Symposium on Geometry Processing, vol. 7 (2006). https://doi.org/10.2312/SGP/SGP06/061-070
66. Bulat, A., Tzimiropoulos, G.: How far are we from solving the 2d & 3d face alignment problem? (and a dataset of 230,000 3d facial landmarks). In: International Conference on Computer Vision (2017)
67. Derkach, D., Ruiz, A., Sukno, F.M.: Head pose estimation based on 3-d facial landmarks localization and regression. In: 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), pp. 820–827 (2017). https://doi.org/10.1109/FG.2017.104
68. Edelsbrunner, H.: Surface reconstruction by wrapping finite sets in space. In: Aronov, B., Basu, S., Pach, J., Sharir, M. (eds.) Discrete and Computational Geometry: The Goodman-Pollack Festschrift. Algorithms and Combinatorics, pp. 379–404. Springer, Berlin, Heidelberg (2003). https://doi.org/10.1007/978-3-642-55566-4_17
69. Giblin, P.: Graphs, Surfaces and Homology, 3rd edn. Cambridge University Press, Cambridge (2010). https://doi.org/10.1017/CBO9780511779534
70. Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G.: The ball-pivoting algorithm for surface reconstruction. IEEE Trans. Vis. Comput. Gr. **5**(4), 349–359 (1999). https://doi.org/10.1109/2945.817351
71. Delaunay, B.N.: Sur la sphère vide. Bulletin de l'Académie des Sciences de l'URSS. VII. Série 1934(6), 793–800 (1934)
72. Paul Chew, L.: Constrained delaunay triangulations. Algorithmica **4**(1), 97–108 (1989). https://doi.org/10.1007/BF01553881
73. Büchner, T., Sickert, S., Graßme, R., Anders, C., Guntinas-Lichius, O., Denzler, J.: Using 2d and 3d face representations to generate comprehensive facial electromyography intensity maps. In: Bebis, G., Ghiasi, G., Fang, Y., Sharf, A., Dong, Y., Weaver, C., Leo, Z., LaViola Jr., J.J., Kohli, L. (eds.) Advances in Visual Computing. Lecture Notes in Computer Science, pp. 136–147. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-47966-3_11
74. Kazemi, V., Sullivan, J.: One millisecond face alignment with an ensemble of regression trees. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1867–1874. IEEE, Columbus (2014). https://doi.org/10.1109/CVPR.2014.241
75. Fritz, C.O., Morris, P.E., Richler, J.J.: Effect size estimates: current use, calculations, and interpretation. J. Exp. Psychol. Gen. **141**(1), 2–18 (2012). https://doi.org/10.1037/a0024338
76. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Psychology Press, New York (2009)
77. Wilcoxon, F.: Individual comparisons by ranking methods. Biometr. Bull. **1**(6), 80–83 (1945). https://doi.org/10.2307/3001968

Tim Büchner received a B.Sc. and M.Sc. in computer science at the Friedrich Schiller University Jena. He is currently pursuing a Ph.D. degree at the Computer Vision Group at the Friedrich Schiller University Jena, under supervision of Joachim Denzler. His research interests include 3D vision for facial analysis bridging the gap between mimics and muscles.

Sven Sickert earned the Diploma degree in Computer Science in 2012 from the Friedrich Schiller University Jena, Germany. He received his Ph.D. in 2018 for his work on 3D semantic segmentation using context features, which was done under supervision of Joachim Denzler. He is currently senior researcher and lecturer at the Computer Vision Group Jena. His research interests are in the field of machine learning, local features and analysis of three-dimensional data with a focus on learning using unstructured data.

Gerd F. Volk is active since 2006 as a physician at the Department of Otorhinolaryngology of Jena University Hospital. Since 2012 he is the Leader of the interdisciplinary Facial-Nerve-Center Jena. His clinical interests are: diagnostic and treatment of facial nerve palsy, telemedicine to improve diagnostic and treatment of rare diseases, application of botulinum toxin and electrostimulation in the head and neck region, functional diagnostics and therapy of peripheral nerve lesions electrophysiological and imaging techniques for the evaluation and rehabilitation of the muscles and nerves. Experimental topics: development of new methods of reconstruction of the facial and laryngeal nerve, electrostimulation as diagnostic and therapeutic tool, central changes after brain nerve failure, in particular of the facial and vestibular nerves, telemedicine and health care networks to improve treatment quality.

Orlando Guntinas-Lichius studied medicine in Cologne, Germany. He completed his specialist training in otorhinolaryngology at the Department of Otorhinolaryngology, University of Cologne in 1998 and worked in the department until 2006. Since 2006, he is full professor for Otorhinolaryngology and head of the Department of Otorhinolaryngology, Jena University Hospital, Jena, Germany. His main research interests are biophotonic diagnostics, head and neck cancer, salivary gland tumors, facial palsy, and facial nerve reconstruction. He is author and co-author of more than $>$ 700 journal articles. He is a member of the German National Academy of Sciences Leopoldina, Collegium Oto-Rhino-Laryngologicum Amicitiae Sacrum, German Society of Oto-Rhino-Laryngology, Head and Neck Surgery, German Cancer Society.

Joachim Denzler earned the diploma (DiplomInformatiker), Dr-Ing, and Habilitation degrees from the University of Erlangen, Germany, in 1992, 1997, and 2003, respectively. Currently, he holds a position as full professor for computer science and is head of the Computer Vision Group, Department of Mathematics and Computer Science, Friedrich Schiller University Jena, Germany. He is also director of the ELLIS Unit in Jena. His research interests comprise the automatic analysis, fusion, and understanding of sensor data, especially development of methods for visual recognition tasks, and dynamic scene analysis. He contributed in the area of active vision, 3D reconstruction, as well as object recognition and tracking. He is author and co-author of more than 300 journal and conference papers as well as technical articles. He is a member of the IEEE, IEEE Computer Society, DAGM, and GI.