A Review on Automated Facial Nerve Function Assessment from Visual Face Capture

Jianwen Lou, Hui Yu, *Senior Member, IEEE,* Fei-Yue Wang, *Fellow, IEEE,*

Abstract—**Assessing facial nerve function from visible facial signs such as resting asymmetry and symmetry of voluntary movement is an important means in clinical practice. By using image processing, computer vision and machine learning techniques, replacing the clinician with a machine to do assessment from ubiquitous visual face capture is progressing more closely to reality. This approach can do assessment in a purely automated manner, hence opens a promising direction for future development in this field. Many studies gathered around this interesting topic with a variety of solutions proposed in recent years. However, to date, none of these solutions have gained a widespread clinical use. This study provides a comprehensive review of the most relevant and representative studies in automated facial nerve function assessment from visual face capture, aiming at identifying the principal challenges in this field and thus indicating directions for future work.**

Index Terms—**automated assessment, facial nerve function, facial palsy, visual face capture**

I. INTRODUCTION

Facial palsy is associated with a myriad of functional [1] and psychosocial problems [2]-[5] that erode foundations of the patient's health and daily life. It generally refers to the weakness of facial musculature innervated by the facial nerve. The main obstacle in facial palsy management is the lack of an effective tool to objectively assess and document facial nerve function, which is crucial to clearly understand the progression or resolution of disease, evaluate the outcomes of therapeutic interventions, and make an accurate prognosis and appropriate treatment plan.

A major part of facial nerve function refers to the motor function manifested by various facial muscle movements, and is visually observable with clear static or dynamic facial signs, e.g. resting symmetry, symmetry of voluntary movement and synkinesis. Facial nerve function assessment from these facial signs is hence an important means in clinical practice. With ubiquitous visual face capture – images/videos, it is more widely accessible than those using obtrusive physical interventions such as electroneurography (ENoG) and electromyography (EMG). This motivated a branch of study [6]-[14] in this field to employ computational measures on

biomedical visual face capture to objectively and quantitatively evaluate the facial nerve function. Such a solution is capable of automatically quantifying facial nerve function in millimetric precision [7], [13] or with semantic grades [15]-[19] based on a machine learning model trained on clinician labelled data [6], [11], [12]. This provides a highly efficient and cost effective means whereby facial nerve function can be appraised in an objective manner. With the development of techniques in image processing, computer vision and machine learning, especially those of 2D/3D face tracking [8]-[10] and feature learning [14], the field of automated facial nerve function assessment has witnessed promising progress and developed various instruments in recent years. However, to date, none of these instruments has gained a wide clinical use. Their clinical effectiveness remains a big doubt mainly because of the limited data used for method development and validation. Meanwhile, important advancements in other areas, monocular 3D face tracking and face image synthesis for example, have not yet been fully utilized.

This study recaps the prerequisite knowledge of the facial nerve function and systematically reviews principal studies in automated facial nerve function assessment from visual face capture which contains rich physiological information [20], to provide readers with an overview of this critical research field and stimulate new ideas. It identifies existing challenges or problems and how they have been tackled so far, and indicates future directions. To the best of our knowledge, this is the first attempt of its kind to be reported so far, which is believed to benefit multiple groups of people, including clinical practitioners, neurologists, researchers in bioengineering, and even non-experienced readers.

II. REVIEW METHODS

A systematic review of the English language literature published from 1977 to 2019 was performed from the resources of PubMed and Google Scholar according to agreed inclusion and exclusion criteria: **Inclusion** – 1) facial nerve function assessment from face images/videos using computational measures; 2) 2D/3D face analysis from visual face capture, including tracking, reconstruction, synthesis and feature learning. **Exclusion** – 1) assessment from non-visual face capture, e.g. electroneurography and electromyography; 2) manual or subjective assessment methods; 3) non-English language.

Preliminary search was performed using key terms such as "facial", "nerve", "function", "assessment", "grading", "palsy", "paralysis", "automated", "automatic", "computer",
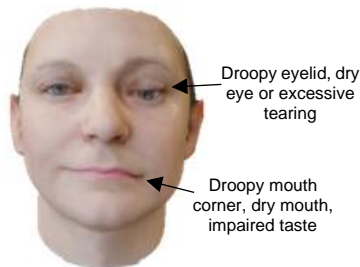
Fig. 1. Typical symptoms of facial palsy.

"vision", "machine", "learning", "image", "video", "processing". These terms were manually grouped as key words for the search of titles which were then screened for potential relevance. 86 titles were searched and 47 articles were retrieved after carefully examining their abstracts and main contents with regard to relevance. Applying the same selection criteria, the searching scope was then extended to the bibliographies of all the selected publications for relevant reports that were not covered by database searching. This process yielded 15 more articles. The articles were grouped in terms of three aspects: 1) what kind of visual face capture was used such as still face image or dynamic facial expression image sequence, RGB or RGBD image; 2) whether provided quantification of static, dynamic and synkinetic facial features; 3) whether predicted semantic grades using machine learning techniques. These three aspects also grounded the subsequent categorization of different automated assessment methods.

In addition, 15 more articles that introduce facial nerve function, facial palsy, their clinical assessment, and clinical facial nerve grading scales were reviewed and summarized to briefly introduce the medical background of this review paper.

III. FACIAL NERVE FUNCTION

Facial nerve function represents a group of fundamental functions performed by the facial nerve - the seventh paired cranial nerve or simply CN VII [21]. It mainly consists of: 1) Motor functions, supplying the muscles of facial expression, the posterior belly of the digastric, the stylohyoid and the stapedius muscles with motor fibres. 2) Sensory functions, providing special taste sensation from the anterior 2/3 of the tongue and general sensation from a small area around the concha of the auricle. 3) Parasympathetic functions, innervating a portion of head and neck glands, including submandibular and sublingual salivary glands. Since the facial nerve is principally composed of motor fibres, facial nerve function generally refers to the motor function manifested by various facial muscle movements, thus can be effectively evaluated from outer facial features without any obtrusive physical intervention.

A. Relationship with Facial Palsy

Once the facial nerve is damaged, the aforementioned functions will be partially or completely lost, hence causing paralysis to the affected side of the face, which is also known as facial palsy. Typical symptoms of facial palsy (see Fig. 1) are inability to frown, reduced elevation of the eyebrow and closure of the eye, loss of blinking and squinting control,

droopy lower eyelid, decreased tearing, dropping of the mouth to the affected side, inability to whistle or blow, altered taste, etc. Facial palsy patients may subsequently suffer from various sequelae [22], including hyperkinesis, synkinesis and atrophy. All of these conditions could result in marked facial disfigurement, interrupt basic human function such as eating, drinking and speaking. The functional disability or impairment may further lead to a wide range of psychosocial problems [2]-[5]. Investigations carried out in Japan, UK and USA show that only the annual incidence of Bell's palsy[1] is 20 to 30 per 100,000 population [22], [23]. It thus calls for immediate and effective action to understand and alleviate the suffering of such a large group of affected people, in which the primary step is to perform an accurate and efficient facial nerve function assessment which is a prerequisite for facial palsy diagnosis and therapy [24], [25].

B. Assessment with Facial Nerve Grading Scales

To date, clinical facial nerve function assessment still relies on clinician to subjectively evaluate features such as resting symmetry, symmetry of voluntary movement and synkinesis. Targeting at providing a more uniform and accurate method for assessing facial nerve function, a variety of facial nerve grading scales [15]-[19] such as House-Brackmann [16], Sunnybrook [17], Yanagihara [15], FNGS 2.0 [18] and eFACE [19] have been developed over the years. These scales divide the degree of facial nerve damage into a series of discrete levels based on some rigorously-validated measures, including facial symmetry at rest, differential voluntary facial muscle movement, and secondary features such as synkinesis. Clinicians summarized the ideal characteristics of a facial nerve grading scale with current technologies: 1) perform regional scoring of facial nerve function; 2) conduct static and dynamic measures; 3) assess secondary sequelae such as synkinesis; 4) generate reproducible results with low interobserver and intraobserver variability; 5) sensitive enough to track changes over time and following interventions; 6) convenient for clinical use. A 2015 systematic review [26] found only Sunnybrook (see Table I) fulfilled all criteria among previous grading scales.

Although sophisticated grading scales [19] are being developed for clinical applications and the discussion [27] over the clinical effectiveness of the scales continues, all these grading scales are limited by the subjective nature of clinician-based assessment and have inherent problems such as labor-intensive, time-consuming and might yield low reproducible results with interobserver and intraobserver variability [26], [27]. As an alternative, automated instruments enabling cost-effective, efficient, objective and quantitative facial nerve function assessment from ubiquitous visual face capture are invaluable and highly expected.

IV. AUTOMATED ASSESSMENT FROM VISUAL FACE CAPTURE

As mentioned above, most facial nerve dysfunction is visually observable with clear static or dynamic facial signs,

---

[1] The most common acute facial nerve paralysis without known causes. It is thought to account for 60-75% facial palsy cases [22], [23].

TABLE I
SUNNYBROOK GRADING SCALE

Measure Description			Score
Resting Symmetry (compared to normal side)	0 – normal, 1 – narrow, wide or eyelid surgery	Eye	
	0 – normal, 2 – absent, 1 – less pronounced or more pronounced	Cheek (naso-labial fold)	
	0 – normal, 1 – corner dropped or corner pulled up/out	Mouth	
Symmetry of Voluntary Movement (degree of muscle excursion compared to normal side)	1 – unable to initiate movement/no movement 2 – initiated slight movement 3 – initiated movement with mild excursion 4 – movement almost complete 5 – movement complete	Forehead Wrinkle	
		Gentle Eye Closure	
		Open Mouth Smile	
		Snarl	
		Lip Pucker	
Synkinesis (degree of involuntary muscle contraction)	0 – none: no synkinesis or mass movement 1 – mild: slight synkinesis 2 – moderate: obvious but not disfiguring synkinesis 3 – severe: disfiguring synkinesis/gross mass movement of several muscles	Forehead Wrinkle	
		Gentle Eye Closure	
		Open Mouth Smile	
		Snarl	
		Lip Pucker	
Resting Symmetry Score (RSS) = score(eye, cheek, mouth) x 5 Voluntary Movement Score (VMS) = score(facial expressions) x 4 Synkinesis Score (SS) = score(facial expressions)		Composite Score = VMS - RSS - SS	

which motivated a lot of studies on automated facial nerve function assessment from biomedical visual capture of the face. A typical paradigm of such an instrument is illustrated in Fig. 2. It first uses an ordinary camera to take pictures of the patient's face when it is at rest or performing specified facial expressions. Then, computational techniques [11]-[14] in various areas such as computer vision, image processing and machine learning are employed to objectively and quantitatively assess the facial nerve function within a certain feature space. The resulting solution can significantly reduce the subjective bias in assessment and would be easily ported to ubiquitous mobile devices such as smartphones and tablets, hence has promising applicability in facial palsy diagnosis and therapy. In the following, we will systematically review the principal studies in this important area along two main dimensions - computational measures and assessment outcomes. According to the modality of the input data, computational measures can be further divided into 2D measures and 3D measures.

A. Computational Measures in 2D

Numerous computational measures on facial palsy images have been developed. They are all based on clinical measurement of facial nerve function, mainly including evaluation of facial symmetry at rest, facial movements and secondary deficits such as synkinesis [28]. Two fundamental categories of computational measures are static measures [7], [29] and dynamic measures [13], [30], whereby facial resting symmetry and muscle movements are principally evaluated.

The Role of Facial Landmarks. A large portion of computational measures are built on top of a group of facial fiducial points called landmarks to quantify facial symmetry and movements. The pioneer work of Burres [31] calculated 13 distances among ten landmarks on faces at rest and during expressions to evaluate facial motor function. The points (see Fig. 3) were manually marked on the face with a grease pencil, and the distance was gauged with a hand-held caliper. This inefficient process was then significantly improved by applying
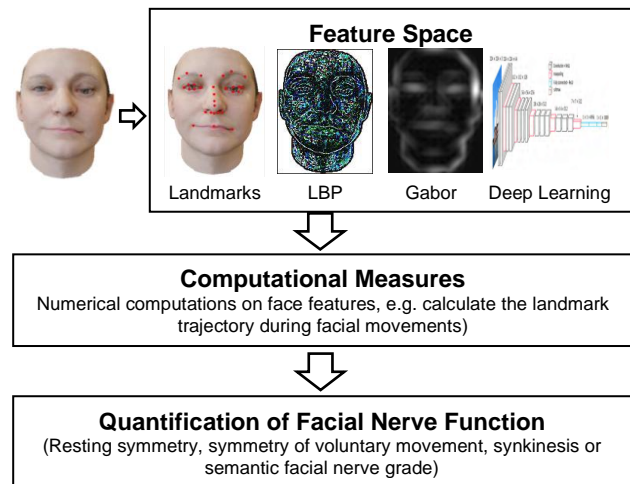

Fig. 2. Pipeline of the automated facial nerve function assessment system.

the reflective marker [32]-[34], image-editing software [7], [35], image processing [36], [37] and computer vision techniques [11]-[13], [38] to automate landmark placement and distance calculation on a digitized face photograph. Evaluation of the angle and area among landmarks is also incorporated to augment the facial function quantification (see Fig. 4) [35], [39]. This initiates a basic measurement which is sensitive to facial abnormalities of spatial (topological) nature and has been widely applied in automated facial nerve function assessment. Whereas various facial landmarks have been proposed in subsequent studies [11]-[13], [35]-[39], there is a simple rule: the landmarks of interest are located close to the facial area responsible for facial movements, or on anatomical points. For example, as shown in Fig. 3, SO (eyebrow) and IO (lower lid) for eye closure, M and Mid for smile, Lc is on lateral canthus and Mc is on medical canthus.

Static Measures. The resting asymmetry is a result of muscle weakness on one side of the face. Typical symptoms (see Fig. 5) are droopy lower eyebrow and lower eyelid, the mouth corner droops, and the depth or orientation of the nasolabial fold alters. Most of these features can be effectively quantified
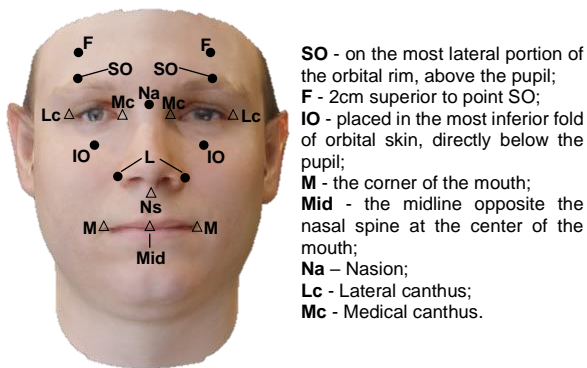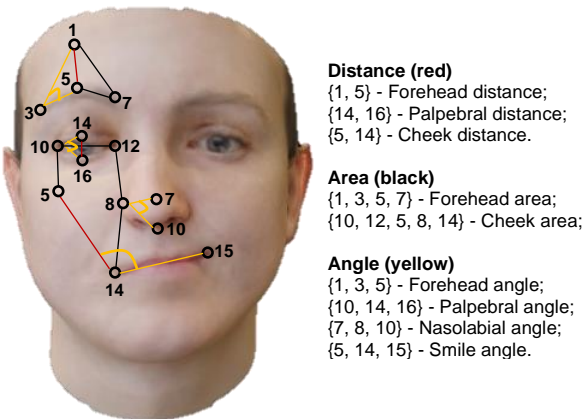
SO - on the most lateral portion of the orbital rim, above the pupil;
F - 2cm superior to point SO;
IO - placed in the most inferior fold of orbital skin, directly below the pupil;
M - the corner of the mouth;
Mid - the midline opposite the nasal spine at the center of the mouth;
Na – Nasion;
Lc - Lateral canthus;
Mc - Medical canthus.

Fig. 3. Facial landmarks applied in [31].

Distance (red)
{1, 5} - Forehead distance;
{14, 16} - Palpebral distance;
{5, 14} - Cheek distance.

Area (black)
{1, 3, 5, 7} - Forehead area;
{10, 12, 5, 8, 14} - Cheek area;

Angle (yellow)
{1, 3, 5} - Forehead angle;
{10, 14, 16} - Palpebral angle;
{7, 8, 10} - Nasolabial angle;
{5, 14, 15} - Smile angle.

Fig. 4. Typical distance, angle and area among landmarks [55].
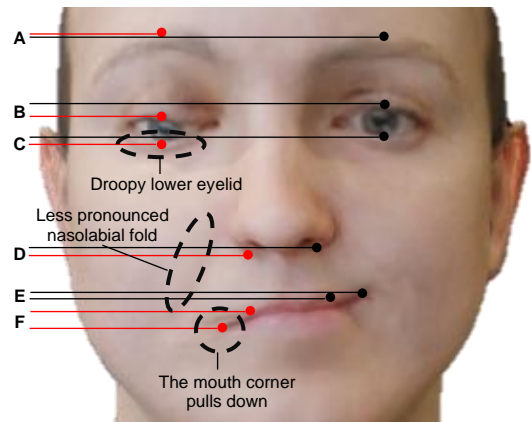

Fig. 5. Landmark position deviations for measuring the resting asymmetry [7]. Red (paralyzed side) and black (normal side) dots represent landmarks on top edge of eyebrow in mid-pupillary line (MPL), margin of upper eyelid in MPL, margin of lower eyelid in MPL, alar base, mid-upper lip position, and oral commissure position. Horizontal black and red lines indicate height of these landmarks. The vertical lines represent facial midline (based on bisection of the inter-pupillary line) (black) and the actual center of the philtrum (red). A - Resting brow ptosis, B - Superior eyelid malposition, C - Inferior eyelid malposition, D - Nasal base ptosis, E – Mid-upper lip ptosis, F - Oral commissure malposition, G - Philtrum deviation.

Forehead Wrinkle   Eye Closure   Nose Wrinkle   Smile   Lip Pucker

Fig. 6. Typical facial expressions involved in evaluation of voluntary movement.

with vertical deviations of landmark positions compared against the normal side of the face, e.g. brow ptosis, superior eyelid malposition, inferior eyelid malposition, nasal base ptosis, mid-upper lip malposition, oral commissure malposition, and philtrum deviation toward the healthy side [7]. Fig. 5 demonstrates such deviations with paired red and black lines on a paralyzed face in repose. Difference between landmark-based triangle areas [39] and angle degrees [35] from the two sides of the face is also frequently used in quantifying the asymmetry. All these measures are initially represented in image pixels, which could be further scaled by the inter-pupillary distance (the average human iris diameter is 11.77mm [40]) to allow "real-life" millimetric measurements on the image.

However, landmarks can hardly depict the contrast between the nasolabial folds in two sides of the face, which exhibits non-pronounced variations in topology. To address this problem, measures upon image pixel intensities could be adopted, e.g. distances between pixel intensities [41]-[43] or mediate visual texture descriptors such as Local Binary Pattern (LBP) histogram features [30] and circular Gabor features [44] from two sides of the face.

Dynamic Measures. Evaluation of facial movements evoked by voluntary muscle contraction lays the basis of almost every facial nerve function assessment instrument. Despite the nearly limitless ways in which humans may move the muscles of facial expression, typical attempted movements critical in facial function and communication are frequently evaluated: forehead wrinkle, eye closure, nose wrinkle, smile and lip pucker [7], [39], [45] (see Fig. 6).

Photographs of the face in repose and with facial expressions, or videos of the face performing facial expressions are normally required and analyzed. Similar with the resting symmetry, the symmetry of voluntary muscle movement can also be efficiently measured using landmarks. Generally, changes lying in difference between landmark-based line distances [7], [45], [46] (or triangle areas [12], [47], angle degrees [13], [35]) introduced in static measures between rest and maximum movements are first gauged to quantify the muscle excursion. Then, the symmetry of voluntary movement is denoted as deviations between quantified muscle excursions on two sides of the face. Secondary defects such as synkinesis resulting from abnormal activation of muscles during expression could be measured in the same way as those mentioned above [28]. Instead of just factoring in two states (neutral and peak) during facial movement, a few measures [10], [32] are based on the trajectories (position over time) of facial landmarks. They can not only appraise abnormalities of spatial nature, but also assess temporal characteristics such as the velocity and moving direction.

As discussed in static measures, an inherent deficiency of landmark-driven measures is that they are insensitive to abnormality with obscure topological features such as changes in the nasolabial fold. The deficiency will be magnified in

voluntary movement symmetry evaluation. As the zygomaticus muscle contracts during smiling, the nasolabial fold commonly deepens a lot on the normal side, while it keeps almost unchanged on the affected side. This contrast manifests clear evidence of asymmetry, thus should not be excluded. To this end, analogous measures [30], [38], [44], [48]-[50] driven by pixel intensities as those used in static measures could be applied. A simple solution is to perform a subtraction between images obtained at rest and during facial movement, then compare the luminance changes of a specific paretic area with that of the healthy side [48], [49]. Such kind of methods however is sensitive to illumination changes, which is restricted to environment with controlled lighting. To cross this constraint, some studies resorted to robust visual texture descriptors [30], [38], [44], [50]. He et al. [30] employed the multi-resolution LBP (MLBP) on temporal-spatial domain to extract the motion features from each region of the face. Then they assessed the symmetry of facial motion by the Resistor-Average Distance (RAD, a distance measure between two probability distributions that is closely related to the Kullback-Leibler distance) between MLBP features. NGO et al. further extended the facial texture analysis from spatial domain to frequency domain by using Gabor filters [50], circular Gabor filters [44]. More recent studies [14], [51] turned to deep learning methods such as convolutional neural networks (CNNs) which have revolutionized the visual imagery analysis to extract high-level features from the face image. The extracted features are supposed to embed the most prominent image patterns probably including the facial abnormality into a compact numerical vector. A major concern about this method is that deep learning always requires a huge amount of data (typically more than 10K images) for training. Creating such a large scale dataset is extremely expensive and time-consuming, let alone it might involve intractable ethics problems as the data exposes the privacy of patients.

As facial movements are driven by muscles located in specific facial areas (this does not apply to synkinesis which is a kind of abnormal muscle activation), regional analysis is important in measuring facial motions. For example, smiling only accounts for facial movement around the mouth region. Therefore, it is beneficial to restrict computational measures within facial regions that are responsible for the corresponding facial movements [37], [38], [43], [52]. The face can be divided into regions according to facial landmarks [38] or using other image segmentation techniques [37], [43].

B. Computational Measures in 3D

An inherent shortcoming of 2D measures is that they cannot deal with out-of-plane facial movement due to the anatomical nature of skull. Gross et al. [53] found that 2D analysis underestimates 3D facial motion amplitudes by up to 43%. Mendes et al. [54] measured the cornea surface on a 3D eyeball model created with a CAD (computer-assisted design) software, which was identified to be far more accurate than calculating only the 2D distance between the two eyelids for corneal exposure measurement. 3D analysis is hence crucial for more accurate assessment of complex facial function.

Fig. 7. Multi-camera setup [55], RGB-D cameras [57], [58] and 3D hand-held scanner [60] used in 3D facial motion capture systems.

Landmark-based Measures. Many existing 3D measures [8]-[10], [55]-[59] are built upon the analysis of 3D facial landmark's trajectory during standardized facial movements. Distances, angles and surface between 3D landmarks on the normal side of the face are typically calculated and compared with that on the paralyzed side [55]. During this procedure, a 3D motion capture system is employed to reconstruct and track the 3D facial geometry. Such systems were usually established with a multi-view camera setup [55] (see Fig. 7) or a mirror structure [8]. These systems required a tedious calibration process and invasive reflective markers attached on the subject's face to track 3D facial landmarks. Mehta et al. [56] applied a different system called 3D VAS which was calibration free and was able to track a dense 3D shape in real-time. However, the 3D VAS required color fringe patterns to be projected on the face during motion capture. It either didn't provide an efficient means to track facial landmarks which had to be manually annotated frame by frame. A few recent studies [10], [57]-[59] developed more compact and cost-effective 3D motion capture systems which only comprised a portable RGB-D camera (see Fig. 7). Meanwhile, advanced computer vision facial tracking algorithms were incorporated to further automate the 3D capture system [8], [10], [58].

Surface-based Measures. The landmark trajectory only outlines the facial movement in a coarse manner, so it is unable to depict more in-depth morphological changes in facial soft tissue. To solve this problem, a few studies [60]-[64] introduced 3D surface-based measures. They first applied commercial 3D scanners such as 3dMDflex™ to repetitively capture the detailed 3D geometry of the face with repose and during facial expressions in a specified period of time. Then, measures such as point-to-point root mean square (RMS) between the registered 3D point clouds of the normal side and the paralysed side, the neutral face and the morphed face to quantify face symmetry and the intensity of the facial expression (see Fig. 8). Statistical analysis such as ANOVA and t-test showed high intra-observer and inter-observer reproducibility of such surface-based measures, which implied a potential more reliable and accurate method to assess facial nerve function. Recent studies [60] found that using mobile hand-held 3D scanners, e.g. Artec™ Eva (see Fig. 7), could achieve similar measure accuracy as that using stand-by immobile 3D scanners. This indicates that 3D surface-based measures could probably be widely applied in clinic without the
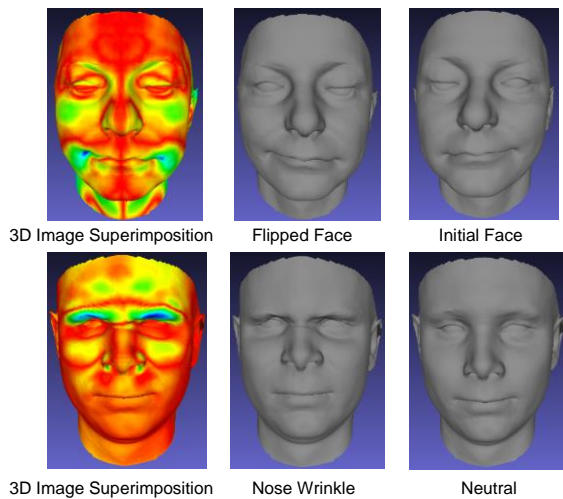
Fig. 8. Point-to-point distance between 3D images of face and flipped face, facial expression and neutral face, for constructing 3D surface-based measures. Please note that the distance value increases from red to blue.

need of a complicated laboratory setup.

C. Assessment Outcomes

Although a variety of automated facial nerve function assessment solutions have been proposed, their outcomes fall into two main categories: 1) non-semantic numerical values quantifying static, dynamic and synkinetic facial features; 2) semantic grade of facial nerve function designed by the clinician.

The majority of existing solutions belongs to the first category, which output at least one aforementioned computational measure in high precision. For instance, as reported in [55], results from a 3D measurement instrument called FACIAL CLIMA varied from the caliper results an average of 0.11% regarding the distance measured and 0.41% regarding the angles measured. The intra-rater (test-retest) reliability of these measurements is quite high, with an intra-rater correlation greater than 0.9 [35]. Most of these solutions however stays at the method discussion phase, only a few of them [7], [8], [10], [13], [35] have been implemented into prototypes. As presented in [10], a typical system of its kind embeds the facial function measuring algorithms into a user-friendly graphical interface to acquire and process facial motion data. The analysis outcomes are organized into a graph named facegram to present the measures with plots and tables. Tools such as pointers, zooming and line axis tracings are provided to facilitate the user interaction. Whilst these solutions provide detailed insights and quantifications about abnormal conditions, they still need clinicians to judge the severity of facial nerve dysfunction.

Solutions in the second category instead aim to quantify the facial nerve function according to a specific facial nerve grading scale designed by the clinicians. To achieve this target, machine learning techniques should be applied to build a predictive model which is trained on labelled data and capable of making predictions on new data. The model is called classifier when the prediction is of assigning an unseen data sample into one or more predefined classes, or regressor if the prediction output is continuous. When applying to our case, the data refers to images of facial movement from either a healthy subject or a facial palsy patient, and the prediction is the grade of facial nerve function. If the grade is discrete, a classifier is employed, otherwise a regressor is employed. The classifier or regressor is trained on a group of facial movement data (from both healthy subject and facial palsy patient) which has been graded by clinicians, using methods such as support vector machine (SVM) [11], [12], [46], [47], artificial neural network (ANN) [49], k-nearest-neighbor (KNN) [30] or hybrid classifier [37]. For a new subject, the solution first extracts computational features from his/her facial movement data, then calls a pre-trained classifier to map the features to the facial nerve function grade defined in the grading scale. The most frequently used grading scale is House-Brackmann scale (HBS) which divides the facial nerve function into six levels [11], [12], [30], [51], followed by Yanagihara scale (YGS) [44], [50] and Sunnybrook scale (SGS, see Table I) [47]. The grade could also simply be a binary value indicating whether the subject has facial palsy or not [37], [46], or if a specific face region is paralyzed [14]. The reported classification accuracy (by comparing the predicted grade against that from the clinician) varies a lot, ranging from 49.9% [12] to 95.5% [47]. As the dataset used for evaluation and the grading scale applied are different in studies, it's difficult to compare solutions from each other. Additionally, although many studies [46], [47] claim that their solutions have been implemented into a computer program or mobile application, only one presents the system prototype [6].

V. DISCUSSION

Although a number of automated facial nerve function assessment instruments have been developed, none of them has gained widespread use in clinical practice to date. The reliability of these instruments lacks sufficient clinical validation, which is the major concern. The instrument's inadequate applicability also remains a big obstacle for it to become widely accessible. According to outcomes discussed above, existing instruments can be broken down into two types – non-semantic instrument (nsINST) and semantic instrument (sINST). nsINST targets at supplying the clinician with objective quantification of facial nerve function. sINST is built on top of a clinical grading scale, which requires a specialized model training on some clinician-labelled data. In the following, we will discuss the limitations of both instruments respectively and envisage the future directions in this field.

A. Limitations of Existing nsINST

Despite the capability of providing high-precision facial function measures comparable against calipers [55], the clinical effectiveness of nsINST remains the primary question as it lacks thorough and rigorous clinical validation. Researchers or clinical practitioners are consistently working on this issue. Bray et al. [35] tested their SMILE system (for measuring lip excursion during smiling with face photographs) on a database of 20 free gracilis transfer procedures with subjectively excellent results and follow-up of 4 to 12 months following single-stage surgery

TABLE II
COMPARISON OF AUTOMATED FACIAL NERVE GRADING SYSTEMS AND CRITERIA

Methods	Static Features	Dynamic Features	3D Features	Deep Features	Grading Scale*	Dataset	Prediction Accuracy
Hsu et al. 2018 [14]	✗	✓	✗	✓	BFR	Hsu et al. 2018 [14]	93%
Sajid et al. 2018 [51]	✗	✓	✗	✓	HBS	Sajid et al. 2018 [51]	92.9%
Guo et al. 2018 [12]	✓	✓	✗	✗	HBS	Guo et al. 2018 [12]	49.9%
Ngo et al. 2016 [9]	✓	✓	✓	✗	YGS	Kihara et al. 2011 [65]	66.5%
Ngo et al. 2016 [44]	✓	✓	✗	✗	YGS	Kihara et al. 2011 [65]	81.2%
Wang et al. 2016 [11]	✓	✓	✗	✗	HBS	Wang et al. 2014 [38]	89.9%
Azoulay et al. 2014 [47]	✓	✓	✗	✗	B	Azoulay et al. 2014 [47]	95.5%
He et al. 2009 [30]	✓	✓	✗	✗	HBS	He et al. 2009 [30]	69.3%

*Grading scale: B – a binary value indicating if the subject has facial palsy or not; BFR – a binary value indicating if a specific face region is paralyzed or not

TABLE III
DATASET USED TO DEVELOP sINSTS

Dataset	Descriptions
Hsu et al. 2018 [14]	**Source**: collected from YouTube. **Data**: 32 videos of 21 facial palsy patients. **Label**: Paralyzed face region – eyes/mouth in a video frame was outlined with an average rectangle plotted by three specialists.
Sajid et al. 2018 [51]	**Source:** collected from UCSD, PCDS and online resources**. Data:** 2, 000 real faial palsy images and 5,000 synthetic facial palsy images generated by GANs [66]. **Label:** each image was labelled with a HBS score.
Guo et al. 2018 [12]	**Source:** captured from recruited subjects. **Data:** 480 images (480×640 pixels) selected from 160 facial expression videos captured from 32 subjects (14 males, 18 females). Each subject performed 5 expressions - expressionless, raising eyebrows, closing eyes, bulging cheek and showing teeth. 3 images randomly selected from each video. **Label:** subjects were graded according to HBS - 5 in I (healthy), 2 in II, 5 in III, 4 in IV, 5 in V and 11 in VI.
Wang et al. 2014 [38]	**Source:** captured from recruited subjects. **Data:** 570 facial epression images from 57 facial palsy patients (31 females, 26 males). 2 images per patient for each of 5 facial expressions – raising eyebrows, closing eyes, screwing up nose, plumping cheek and opening mouth. **Label:** each subject was graded with a HBS score.
Azoulay et al. 2014 [47]	**Source:** captured from recruited subjects. **Data:** videos of 9 facial expressions (face at rest, strong eye elosure, weak eye closure, rasied eyebrows, closed mouth smile, big smile, puckering of lips, puff-up cheeks and stretching down lower lip) were recorded from 14 patients and 31 healthy subjects (15 females, 30 males). **Label:** three otolaryngologists independently graded the patients's facial palsy according to HBS, YGS and SGS.
Kihara et al. 2011 [65]*	**Source:** captured from recruited subjects. **Data:** multiview face images captured from 83 subjects (74 patients, 9 healthy subjects) with a multi-camera setup (7 cameras). Each subject performed 10 expressions. Each camera took 60 images ($2,112 \times 2,816$ pixels) for each expression. **Label:** each expression was graded with a YGS score.
He et al. 2009 [30]	**Source:** captured from recruited subjects. **Data:** 197 videos (720×576 pixels, 500-700 frames per video) taken from subjects with Bell's palsy, trauma to the nerve from skull fracture and surgical damage, and normal subjects. Each video presents 5 facial movements. **Label:** each video was graded with HBS by a clinician.

* The dataset was proposed in [65] and then applied in [9] and [44]. However, some key information of the database reported in the three papers are inconsistent, including the number of subjects involved, which is 5 in [65], 83 in [9] and 85 in [44]. As common authors are found in all the three papers, it seems that the database has been extended after it was first reported. We reported here the version with the most details.

or 12 to 18 months following second-stage surgery to evaluate outcomes in facial reanimation. In [8], Tzou et al. reported 241 facial palsy patients were filmed and analyzed with their 3D facial motion capture system, accounting for more than 1,000 videos made to track the rehabilitation progress after each operational therapy. These tests validate the reliability of nsINST to some extent, however the sample size and variety involved in the cohort study yet seem to be insufficient for a medical tool. The instruments hence fail to gain a wider (e.g. international) agreement and are only locally accepted.

Another essential problem is that existing nsINST are highly constrained by ineffective motion capture techniques used. To ensure the accuracy of measures, during motion capture, the patient's head was often required to stay in a stable position relative to the camera [12]. Intrusive markers were normally required to be placed on the face to track facial landmarks [8], [10]. These not only cause discomfort or annoyance to the patient, but also prevent the system from being applied in a broader range of circumstances such as the patient's home.

B. Limitations of Existing sINST

sINST utilizes machine intelligence to grade facial nerve function according to a semantic facial nerve grading scale. However, current sINSTs are still far from satisfying clinical requirements and have apparent limitations. As the performance of a sINST relies on the grading scale applied, the extracted features in the prediction process and the dataset for training, the following discussion will concentrate on these three aspects.

Since a sINST is built on top of a facial nerve grading scale, its reliability highly depends on the robustness of the applied scale. As described in the previous section, clinicians have specified several characteristics for an ideal facial nerve grading scale and find only Sunnybrook meets all the criteria [26]. However, Table II shows that most existing sINSTs were built upon less advanced grading scales such as HBS and YGS, which divide the overall facial nerve function into a few discrete levels with only general explanations. The potential effect of such sINSTs is therefore limited. The reason that previous sINSTs preferred to use less sophisticated grading scales is supposed to have two folds: 1) Sophisticated grading scales such as Sunnybrook require accurate sub-scores for different facial regions and facial expressions, which is more arduous for the clinician to grade. This makes the training data more expensive to acquire. 2) Modelling the grade consisting of semantic sub-grades will introduce more complexity to the machine learning algorithm. Therefore, to develop a sINST, it is important to find a good trade-off between the grading scale's robustness and the machine learning model's complexity.

An ideal feature is supposed to contain critical information of facial nerve function, mainly including resting symmetry, symmetry of voluntary facial movement and synkinesis. As introduced in the previous section, these features can be acquired from static, dynamic and 3D measures. We thus summarize the representative sINST according to the measures they performed. As shown in Table II, almost all sINST conducted static and dynamic measures. In [14] and [51], deep learning methods were applied to extract high-level features that output a promising prediction accuracy rate. Meanwhile, rare sINST utilized 3D measures. As discussed in Section IV, 3D measures have shown to be superior against 2D measures, hence should cause more attention. It can also be noticed that the prediction accuracy rates of sINST vary a lot, from 49.9% to 95.5%. Since datasets (see Table III), facial nerve function grading scales and evaluation protocols (e.g. what were input to the instrument, images or videos? How many samples were for training and testing?) adopted in these sINST are different from each other, the accuracy value actually cannot fully reflect the instrument's true performance.

As shown in Table III, datasets applied in studies are different from each other. The biggest concern is that the subject cohort involved in existing datasets seems to be insufficient. For example, in [12], most HBS grades contain less than 5 subjects. Meanwhile, most datasets [11], [12], [65] only include subjects from an identical ethnic background. Their applicability to other ethnic groups needs to be further verified. Another issue is none of these datasets is publicly accessible, causing no benchmark available to develop a widely accepted sINST and further push it to the clinical use.

C. Prospect

Overall, for both nsINST and sINST, a widely acceptable benchmark database for evaluation is urgently needed. The constraint is mainly due to the high complexity and expense of data collection which could be alleviated by more extensive collaborations among practitioners across the world. It is worth pointing out that, in [51], the authors proposed to augment the original training dataset by automatically synthesizing facial palsy images (see Fig. 9) with a cutting-edge deep learning method – GANs [66], which is cost-effective and highly efficient. Although the synthetic facial palsy images in [51] still need significant improvements, it inspires us to introduce a novel theory – Parallel Vision [67] to solve the data problem.

Parallel Vision emphasizes the importance of photorealistic image synthesis in addressing the problems of visual perception and understanding. It comprises three stages: 1) building artificial (virtual) scenes by synthesizing diverse photorealistic data samples to simulate natural physical scenes that occur in real life; 2) conducting computational experiments on the pre-built artificial data to develop vision models (algorithms); 3) executing the vision model on the artificial data and real data concurrently to realize virtual/real interaction. Consequently, the vision model could be continuously optimized. The theory has been successfully applied in many facial analysis tasks, e.g. monocular 3D face reconstruction [68], facial expression synthesis [69], 3D gaze estimation [70] and facial expression
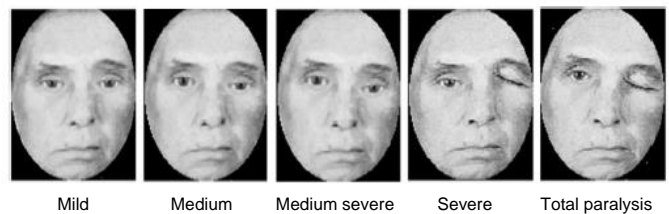

Fig. 9. Facial palsy images synthesized in [51] with various severity level.

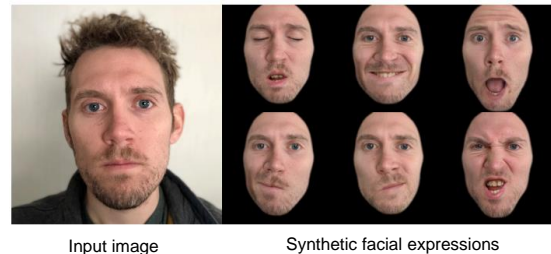
Input image   Synthetic facial expressions

Fig. 10. Facial expressions synthesized in [72].

recognition [71]. In [71], the authors trained their facial expression recognition model on a dataset consisting of synthetic face images rendered from 3D facial scans and real images captured from movies, and achieved a very promising recognition rate which outperformed the state-of-the-art by an average of 11.13% for seven basic facial expressions. Meanwhile, with only a single face image in arbitrary poses, existing face synthesis techniques can generate natural-looking face images [72] even for those with extreme facial expressions such as asymmetric facial expressions (see Fig. 10). This provides solid technical supports for synthesizing photorealistic facial palsy images. We therefore believe that the Parallel Vision theory has a huge potential to fill the data gap discussed in this paper and worth to be further investigated.

On the other side, unconstrained monocular 2D/3D face reconstruction and tracking approaches [73]-[77], and integrated mobile RGB-D sensors such as TrueDepth™, are suggested to be incorporated for more flexible and portable nsINST and sINST. All these future work call for profound interdisciplinary collaborations. Specifically, advancements in other areas, such as computer vision and deep learning should be incorporated much more to promote the instrument development.

VI. CONCLUSION

Effective and objective assessment of facial nerve function in facial palsy patients is essential to gauge severity of disease, monitor progression over time, evaluate the outcomes of therapeutic interventions and facilitate communications among practitioners, however still remains unresolved. Automated instrument working on biomedical visual face capture utilizes image processing, computer vision and machine learning techniques to carry out computational measures on facial nerve function in a highly efficient and widely accessible way, is appearing as a promising solution. By reviewing principal studies related to this topic, this review finds that though many automated instruments have been developed, they are still at a preliminary stage far from meeting clinical requirements. These instruments are severely limited by the lack of a rigorously

validated benchmark database and insufficient incorporation of advancements in other areas such as monocular 3D face tracking and deep learning. To eliminate these obstacles, broader and deeper interregional and interdisciplinary collaborations are necessary and highly anticipated. Advancements in computer vision and deep learning areas such as the Parallel Vision theory [67], unconstrained monocular 2D/3D face reconstruction and tracking techniques [73]-[76] should be incorporated much more to further develop the instrument.

REFERENCES

[1] A. Fattah, G. H. Borschel, R.T. Manktelow, M. Bezuhly and R. M. Zuker, "Facial palsy and reconstruction", Plastic and Reconstructive Surgery, vol. 129, no. 2, pp. 340-352, 2012.

[2] J.M. VanSWEARINGEN, J. F. Cohn, J. Turnbull, T. Mrzai and P. Johnson, "Psychological distress: linking impairment with disability in facial neuromotor disorders", Otolaryngology—Head and Neck Surgery vol. 118, no. 6, pp. 790-796, 1998.

[3] E. T. Bradbury, W. Simons and R. Sanders, "Psychological and social factors in reconstructive surgery for hemi-facial palsy", Journal of Plastic, Reconstructive & Aesthetic Surgery, vol. 59, no. 3, pp. 272-278, 2006.

[4] L. Fu, C. Bundy and S. A. Sadiq, "Psychological distress in people with disfigurement from facial palsy", Eye, vol. 25, no. 10 pp.1322, 2011.

[5] D.T. Walker, M. J. Hallam, S. Mhurchadha, P. McCabe and C. Nduka, "The psychosocial impact of facial palsy: our experience in one hundred and twenty six patients", Clinical Otolaryngology, vol. 37, no. 6, pp. 474-477, 2012.

[6] B. O'reilly, J. J. Soraghan, S. McGrenary and S. He, "Objective method of assessing and presenting the House-Brackmann and regional grades of facial palsy by production of a facogram", Otology & Neurotology, vol. 31, no. 3 pp. 486-491, 2010.

[7] T. A. Hadlock and L. S. Urban. "Toward a universal, automated facial measurement tool in facial reanimation", Archives of Facial Plastic Surgery, vol. 14, no. 4, pp. 277-282, 2012.

[8] C. H. J. Tzou, I. Pona, E. Placheta, A. Hold, M. Michaelidou, N. Artner, W. Kropatsch, H. Gerber and M. Frey, "Evolution of the 3-dimensional video system for facial motion analysis: ten years' experiences and recent developments", Annals of plastic surgery, vol. 69, no. 2, pp. 173-185, 2012.

[9] T. H. Ngo, Y.-W. Chen, M. Seo, N. Matsushiro and W. Xiong, "Quantitative analysis of facial paralysis based on three-dimensional features." In International Conference on Image Processing (ICIP), pp. 1319-1323, 2016.

[10] A. Gerós, R. Horta and P. Aguiar, "Facegram–Objective quantitative analysis in facial reconstructive surgery", Journal of Biomedical Informatics, vol. 61, pp. 1-9, 2016.

[11] T. Wang, S. Zhang, J. Dong, L. Liu, and H. Yu. "Automatic evaluation of the degree of facial nerve paralysis." Multimedia Tools and Applications, vol. 75, no. 19, pp. 11893-11908, 2016.

[12] Z. Guo, G. Dan, J. Xiang, J. Wang, W. Yang, H. Ding, O. Deussen and Y. Zhou, "An unobtrusive computerized assessment framework for unilateral peripheral facial paralysis", IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 3, pp. 835-841, 2018.

[13] D. L., Guarin, J. Dusseldorp, T. A. Hadlock and Nate Jowett, "A machine learning approach for automated facial measurements in facial palsy", JAMA Facial Plastic Surgery, 2018.

[14] G. J. Hsu, W.-F. Huang and J-H Kang, "Hierarchical Network for Facial Palsy Detection", In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 693-6936, 2018.

[15] N. Yanagihara, "Grading of facial nerve palsy". Third International Symposium on Facial Nerve Surgery. pp. 533–535, 1977.

[16] J.W. House, "Facial nerve grading systems", The Laryngoscope, vol. 93, no. 8, pp. 1056-1069, 1983.

[17] B. G. Ross, G Fradet and J. M. Nedzelski, "Development of a sensitive clinical facial grading system", Otolaryngology-Head and Neck Surgery, vol. 114, no. 3, pp. 380-386, 1996.

[18] Facial Nerve Disorders Committee, J. T. Vrabec, D. D. Backous, H. R. Djalilian, P. W. Gidley, J. P. Leonetti, S. J. Marzo et al, "Facial nerve grading system 2.0", Otolaryngology-Head and Neck Surgery, vol.140, no. 4, pp.445-450, 2009.

[19] C. A. Banks, P. K. Bhama, J. Park, C. R. Hadlock and T. A. Hadlock, "Clinician-graded electronic facial paralysis assessment: the eFACE", Plastic and Reconstructive Surgery, vol. 136, no. 2, pp. 223-230, 2015.

[20] M. D. Samad, N. Diawara, J. L. Bobzien, J. W. Harrington, M. A. Witherow and K. M. Iftekharuddin, "A Feasibility Study of Autism Behavioral Markers in Spontaneous Facial, Visual, and Hand Movement Response Data", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, no. 2, pp. 353-361, 2018.

[21] S. Gupta, F. Mends, M. Hagiwara, G. Fatterpekar and P. C. Roehm, "Imaging the facial nerve: a contemporary review", Radiology Research and Practice, 2013.

[22] E. Peitersen, "Bell's palsy: the spontaneous course of 2,500 peripheral facial nerve palsies of different etiologies", Acta Oto-Laryngologica, vol. 122, no. 7, pp. 4-30, 2002.

[23] U. R. Newadkar, L. Chaudhari and Y. K. Khalekar, "Facial palsy, a disorder belonging to influential neurological dynasty: Review of literature", North American Journal of Medical Sciences, vol.8, no. 7, pp. 263, 2016.

[24] D. Jayatilake, T. Isezaki, Y. Teramoto, K. Eguchi and Kenji Suzuki, "Robot assisted physiotherapy to support rehabilitation of facial paralysis", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 22, no. 3 pp. 644-653, 2014.

[25] J. T. Heaton, C. J. Knox, J. S. Malo, J. B. Kobler and T. A. Hadlock, "A system for delivering mechanical stimulation and robot-assisted therapy to the rat whisker pad during facial nerve regeneration", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 21, no. 6, pp. 928-937, 2013.

[26] A. Y. Fattah, A. D. Gurusinghe, J. Gavilan, T. A. Hadlock, J. R. Marcus, H. Marres, C. C. Nduka, W. H. Slattery and A. K. Snyder-Warwick, "Facial nerve grading instruments: systematic review of the literature and suggestion for uniformity", Plastic and Reconstructive Surgery, vol. 135, no. 2, pp. 569-579, 2015.

[27] R. Niziol, F.P. Henry, J. I. Leckenby and A. O. Grobbelaar, "Is there an ideal outcome scoring system for facial reanimation surgery? A review of current methods and suggestions for future publications", Journal of Plastic, Reconstructive & Aesthetic Surgery, vol. 68, no. 4, pp.447-456, 2015.

[28] I. J. Kleiss, M. H. Hohman, O. E. Quatela, H. A. Marres and T. A. Hadlock, "Computer - Assisted Assessment of Ocular Synkinesis: A Comparison of Methods", The Laryngoscope, vol. 123, no. 4, pp. 879-883, 2013.

[29] A. Song, G. Xu, X. Ding, J. Song, Gang. Xu and W. Zhang, "Assessment for facial nerve paralysis based on facial asymmetry", Australasian Physical & Engineering Sciences in Medicine, vol. 40, no. 4, pp. 851-860, 2017.

[30] S. He, J. J. Soraghan, B. F. O'Reilly, and D. Xing, "Quantitative analysis of facial paralysis using local binary patterns in biomedical videos", IEEE Transactions on Biomedical Engineering, vol 56, no. 7, pp. 1864-1870, 2009.

[31] S. A. Burres, "Facial biomechanics: the standards of normal", The Laryngoscope, vol. 95, no. 6, pp. 708-714, 1985.

[32] S. Miyazaki, A. Ishida and A. Komatsuzaki, "A clinically oriented video-based system for quantification of eyelid movements", IEEE Transactions on Biomedical Engineering, vol. 47, no. 8, pp. 1088-1096, 2000.

[33] C. J. Linstrom, "Objective facial motion analysis in patients with facial nerve dysfunction." The Laryngoscope, vol. 112, no. 7, pp. 1129-1147, 2002.

[34] J. Jorge, J.J. Pialarissi, P.R. Borges, G.C. Squella, S.A.F. Gouveia, M.D.F.D. Saragiotto, J.C. Jr and V.R. Gonçalves, "Objective computorized evaluation of normal patterns of facial muscles contraction", Brazilian Journal of Otorhinolaryngology, vol. 78, no. 2, pp. 41-51, 2012.

[35] D. Bray, D. K. Henstrom, M. L. Cheney and T. A. Hadlock, "Assessing outcomes in facial reanimation: evaluation and validation of the SMILE system for measuring lip excursion during smiling", Archives of Facial Plastic Surgery, vol. 12, no. 5, pp. 352-354, 2010.

[36] J. Dong, L. Ma, Q. Li, S. Wang, L. Liu, Y. Lin and M. Jian, "An approach for quantitative evaluation of the degree of facial paralysis based on salient point detection", In Intelligent Information Technology Application Workshops (IITAW'08), pp. 483-486, 2008.

[37] J. Barbosa, K. Lee, S. Lee, B. Lodhi, J.-G. Cho, W.-K. Seo and J. Kang, "Efficient quantitative assessment of facial paralysis using iris segmentation and active contour-based key points detection with hybrid classifier", BMC Medical Imaging, vol. 16, no. 1, pp. 23, 2016.

[38] T. Wang, J. Dong, X. Sun, S. Zhang and S. Wang, "Automatic recognition of facial movement for paralyzed face", Bio-medical Materials and Egineering, vol. 24, no. 6, pp. 2751-2760, 2014.

[39] P. Dulguerov, D. Wang, T. V. Perneger, F. Marchal and W. Lehmann, "Videomimicography: the standards of normal revised", Archives of Otolaryngology–Head & Neck Surgery, vol. 129, no. 9, pp. 960-965, 2003.

[40] F. Rüfer, A. Schröder and C. Erb, "White-to-white corneal diameter: normal values in healthy humans obtained with the Orbscan II topography system", Cornea, vol. 24, no. 3, pp. 259-261, 2005.

[41] S. Wang, H. Li, F. Qi and Y. Zhao, "Objective facial paralysis grading based onP face and eigenflow", Medical and Biological Engineering and Computing, vol. 42, no. 5, pp. 598-603, 2004.

[42] S. Wang and F. Qi, "Compute aided diagnosis of facial paralysis based on pface", In 27th IEEE Annual International Conference on Engineering in Medicine and Biology Society (EMBS), pp. 4353-4356, 2005.

[43] L. Liu, G. Cheng, J. Dong, S. Wang and H. Qu, "Evaluation of facial paralysis degree based on regions", In 2010 Third International Conference on Knowledge Discovery and Data Mining, pp. 514-517, 2010.

[44] T. H. Ngo, M. Seo, N. Matsushiro, W. Xiong and Y.-W. Chen, "Quantitative analysis of facial paralysis based on limited-orientation modified circular Gabor filters", In 23rd International Conference on Pattern Recognition (ICPR), pp. 349-354, 2016.

[45] S. A. Burres, "Objective grading of facial paralysis." Annals of Otology, Rhinology & Laryngology, vol. 95, no. 3, pp. 238-241, 1986.

[46] H. S. Kim, S. Y. Kim, Y. H. Kim and K. S. Park, "A smartphone-based automatic diagnosis system for facial nerve palsy", Sensors, vol. 15, no. 10, pp. 26756-26768, 2015.

[47] O. Azoulay, Y. Ater, L. Gersi, Y. Glassner, O. Bryt and D. Halperin, "Mobile Application for Diagnosis of Facial Palsy", 2014.

[48] V. Meier-Gallati, H. Scriba and U. Fisch, "Objective scaling of facial nerve function based on area analysis (OSCAR)", Otolaryngology—Head and Neck Surgery, vol. 118, no. 4, pp. 545-550, 1998.

[49] S. McGrenary, B. F. O'Reilly and J. J. Soraghan, "Objective grading of facial paralysis using artificial intelligence analysis of video data", In 18th IEEE Symposium on Computer-Based Medical Systems, pp. 587-592, 2005.

[50] T. H. Ngo, M. Seo, N. Matsushiro and Y.-W. Chen, "Evaluation of facial paralysis based on spatial features of filtered images", International Journal of Bioscience, Biochemistry and Bioinformatics, vol. 6, no. 1, pp. 1, 2016.

[51] M. Sajid, T. Shafique, M. Baig, I. Riaz, S. Amin and S. Manzoor, "Automatic Grading of Palsy Using Asymmetrical Facial Features: A Study Complemented by New Solutions", Symmetry, vol. 10, no. 7, p. 242, 2018.

[52] N. Sawai, N. Hato, N. Hakuba, H. Takahashi, M. Okada and K. Gyo, "Objective assessment of the severity of unilateral facial palsy using OKAO Vision® facial image analysis software", Acta oto-laryngologica vol. 132, no. 9, pp.1013-1017, 2012.

[53] M. M. Gross, C.-A. Trotman and K. S. Moffatt, "A comparison of three-dimensional and two-dimensional analyses of facial motion", The Angle Orthodontist, vol. 66, no. 3, pp. 189-194, 1996.

[54] V. M. Mendes, J. Lasudry, L. Vandermeeren and S. D. Fontaine, "Computerised 3D evaluation of the functional eyelid deficit in facial palsy", Journal of Plastic, Reconstructive & Aesthetic Surgery, vol. 67, no. 2, pp. 178-182, 2014.

[55] B. Hontanilla and C. Aubá, "Automatic three-dimensional quantitative analysis for evaluation of facial movement", Journal of Plastic, Reconstructive & Aesthetic Surgery, vol. 61, no. 1, pp. 18-30, 2008.

[56] R. P. Mehta, S. Zhang and T. A. Hadlock, "Novel 3-D video for quantification of facial movement", Otolaryngology—Head and Neck Surgery, vol. 138, no. 4, pp. 468-472, 2008.

[57] A. Gaber, M. F. Taher and M. A. Wahed, "Quantifying facial paralysis using the kinect v2", In 37th IEEE Annual International Conference on Engineering in Medicine and Biology Society (EMBS), pp. 2497-2501, 2015.

[58] N. Vinokurov, D. Arkadir, E. Linetsky, H. Bergman and D. Weinshall, "Quantifying hypomimia in Parkinson patients using a depth camera", In International Symposium on Pervasive Computing Paradigms for Mental Health, pp. 63-71, 2015.

[59] S. Katsumi, S. Esaki, K. Hattori, K. Yamano, T. Umezaki and S. Murakami, "Quantitative analysis of facial palsy using a three-dimensional facial motion measurement system", Auris Nasus Larynx, vol. 42, no. 4, pp. 275-283, 2015.

[60] U. Özsoy, R. Sekerci, A. Hizay, Y. Yildirim and H. Uysal. "Assessment of reproducibility and reliability of facial expressions using 3D handheld scanner", Journal of Cranio-Maxillofacial Surgery, vol. 47, no. 6, pp. 895-901, 2019.

[61] C. Sforza, E. Ulaj, D. M. Gibelli, F. Allevi, V. Pucciarelli, F. Tarabbia, D. Ciprandi, G. Dell'Aversana Orabona, C. Dolci and F. Biglioli. "Three-dimensional superimposition for patients with facial palsy: an innovative method for assessing the success of facial reanimation procedures", British Journal of Oral and Maxillofacial Surgery, vol. 56, no. 1, pp. 3-7, 2018.

[62] D. Gibelli, M. Codari, V. Pucciarelli, C. Dolci, and C. Sforza, "A quantitative assessment of lip movements in different facial expressions through 3-dimensional on 3-dimensional superimposition: a cross-sectional study", Journal of Oral and Maxillofacial Surgery, vol. 76, no. 7, pp. 1532-1538, 2018.

[63] A. Patel, S. M. S. Islam, K. Murray and M. S. Goonewardene, "Facial asymmetry assessment in adults using three-dimensional surface imaging", Progress in Orthodontics, vol. 16, no. 1 pp. 36, 2015.

[64] H. O. Taylor, C. S. Morrison, O. Linden, B. Phillips, J. Chang, M. E. Byrne, S. R. Sullivan and C. R. Forrest, "Quantitative facial asymmetry: using three-dimensional photogrammetry to measure baseline facial surface symmetry", Journal of Craniofacial Surgery, vol. 25, no. 1, pp. 124-128, 2014.

[65] Y. Kihara, G. Duan, T. Nishida, N. Matsushiro, and Y.-W. Chen, "A dynamic facial expression database for quantitative analysis of facial paralysis", In the 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT), pp. 949-952, 2011.

[66] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, "Generative adversarial nets", In Advances in Neural Information Processing Systems, pp. 2672-2680, 2014.

[67] K. Wang, C. Gou, N. Zheng, J. M. Rehg, and F.-Y. Wang, "Parallel vision for perception and understanding of complex scenes: methods, framework, and perspectives", Artificial Intelligence Review, vol. 48, no. 3, pp.299-329, 2017.

[68] E. Richardson, M. Sela and R. Kimmel, "3D face reconstruction by learning from synthetic data", in Proceedings of the 4th IEEE International Conference on 3D Vision (3DV), pp. 460-469, 2016.

[69] B. Gecer, B. Bhattarai, J. Kittler and T.-K. Kim, "Semi-supervised Adversarial Learning to Generate Photorealistic Face Images of New Identities from 3D Morphable Model", arXiv preprint arXiv:180403675, 2018.

[70] F. Lu, Y. Gao and X. Chen, "Estimating 3D gaze directions using unlabeled eye images via synthetic iris appearance fitting", IEEE Transactions on Multimedia, vol. 18, no. 9, pp. 1772-1782, 2016.

[71] Y. Wang, H. Yu, J. Dong, M. Jian and H. Liu, "Cascade support vector regression-based facial expression-aware face frontalization", In IEEE International Conference on Image Processing (ICIP), pp. 2831-2835, 2017.

[72] K. Nagano, J. Seo, J. Xing, L. Wei, Z. Li, S. Saito, A. Agarwal, J. Fursund and H. Li, "paGAN: real-time avatars using dynamic textures", ACM Transactions on Graphics (TOG), vol. 37, no. 6, pp. 258 , 2018.

[73] C. Cao, D. Bradley, K. Zhou and T. Beeler, "Real-time high-fidelity facial performance capture", ACM Transactions on Graphics (ToG), vol. 34, no. 4, p. 46, 2015.

[74] J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt and M. Nießner. "Face2face: Real-time face capture and reenactment of rgb videos", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2387-2395, 2016.

[75] C. Gou, Y. Wu, F.-Y. Wang and Q. Ji, "Coupled cascade regression for simultaneous facial landmark detection and head pose estimation", In IEEE International Conference on Image Processing (ICIP), pp. 2906-2910, 2017.

[76] S. Yamaguchi, S. Saito, K. Nagano, Y. Zhao, W. Chen, K. Olszewski, S. Morishima and H. Li, "High-fidelity facial reflectance and geometry inference from an unconstrained image", ACM Transactions on Graphics (TOG), vol. 37, no. 4, pp.162, 2018.

[77] H.-G. Ren, W.-M. Liu, T. Shi and F.-J. Li, "Compressive tracking based on online Hough forest", International Journal of Automation and Computing, vol. 14, no. 4, pp. 396-406, 2017.