

UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE POSTGRADO



“Cuantificación del grado de parálisis facial mediante algoritmos basados en tracking facial”

Rodrigo Eduardo Gálvez Rojas

TESIS PARA OPTAR AL GRADO DE MAGISTER EN INFORMÁTICA MÉDICA

Director de Tesis: Prof. Dr. Mauricio Cerda

Co-Director: Prof. Dr. Lucas Terissi

2017

Agradecimientos

A Mauricio Cerda, por su constante guía y paciencia infinita.

A Lucas Terissi, por su ayuda e interés en este trabajo.

A Carlos Cruz, por facilitar los videos de pacientes chilenos.

A Rodrigo Assar, por su comprensión.

A Slim Ouni, parte de los datos fueron adquiridos gracias al financiamiento del proyecto STIC-AmSud (15STIC04).

A mis padres, que me han apoyado incondicionalmente toda la vida.

Índice

1	Introducción	7
1.1	Evaluación clínica de la parálisis facial	8
1.2	Métodos de graduación de la parálisis facial	10
1.2.1	Métodos de graduación clínicos	10
1.2.2	Métodos de graduación no clínicos	10
1.2.3	Comparación de métodos no clínicos actuales	11
1.3	Métodos de tracking facial	13
1.3.1	Tracking basado en flujo óptico	13
1.3.2	Tracking basado en modelos	15
1.3.3	Modelo Candide-3	18
1.4	Problema	20
2	Hipótesis y objetivos	21
3	Materiales y métodos	22
3.1	Bases de datos experimentales	22
3.1.1	Filmación con el sistema de seguimiento cinemático con marcadores pasivos ..	22
3.1.2	Filmación de pacientes con graduación	23
3.1.3	Filmación de pacientes sin graduación	25
3.2	Evaluación de una herramienta de tracking facial existente	26
3.3	Diseñar un algoritmo de seguimiento facial con soporte de asimetría	27
3.4	Evaluación del método propuesto	28
4	Resultados Obtenidos	30
4.1	Evaluación de herramienta de seguimiento existente	30
4.2	Algoritmo de medición de parálisis facial propuesto	32
4.2.1	Máscara Candide-3 asimétrica	33

4.2.2	Tracking facial	35
4.2.3	Mediciones para la cuantificación	38
4.2.4	Algoritmo de graduación de la parálisis facial	43
4.3	Validación del método propuesto.....	50
4.3.1	Comparación con el sistema de seguimiento cinemático con marcadores pasivos 50	
4.3.2	Hallazgos en filmaciones de sujetos con graduación.....	51
4.3.3	Evaluación en filmaciones de pacientes sin graduación	54
5	Discusión	57
6	Conclusiones	59
7	Referencias	60
	Anexo A: AUV de la máscara Candide-3 asimétrica.....	61
	Anexo B: Errores en el FO.....	63
	Anexo C: Procedimiento de ajuste de marcadores pasivos y el método propuesto	64

Resumen

La parálisis facial afecta entre 15 a 40 casos por cada 100.000 personas en la población mundial. De estos, el 50% corresponde a parálisis de Bell. Esta patología es un tipo de parálisis de origen desconocido (parálisis idiopática) con posibilidad de rehabilitación. Este tipo de parálisis es la causa más frecuente de la parálisis facial parcial. En la actualidad existen métodos clínicos de graduación del nivel de parálisis facial que tienen la desventaja de ser subjetivos, ambiguos y ofrecer una graduación discreta. Se han hecho trabajos para objetivar la graduación con el apoyo computacional constituyendo métodos de graduación no clínicos, que resultan más objetivos y precisos, pero con la desventaja de que requieren marcadores visuales, mucho tiempo de preparación o son de difícil uso, lo que limita su aplicación en entornos clínicos. Esta tesis busca proponer un nuevo método de graduación de la parálisis facial parcial basado en un algoritmo de tracking facial automático sin marcadores ni preparación previa y validarlo computacionalmente.

Se propone un nuevo método de graduación no clínica que detecta expresiones faciales asimétricas utilizando un modelo paramétrico tridimensional modificado para dar soporte a las expresiones faciales asimétricas. Este se basa en un algoritmo diseñado para realizar tracking facial que detecta expresiones faciales simétricas, sin la ayuda de marcadores, basado en un modelo tridimensional paramétrico que codifica rostros y expresiones faciales simétricas. El algoritmo de tracking facial del método propuesto es comparado con el algoritmo original, presentando un tracking de calidad similar. También es comparado con un sistema de seguimiento cinemático con marcadores ópticos pasivos, con una diferencia promedio entre un marcador pasivo y un punto del modelo de 1,03 mm (d.s. 0,91 mm). El método propuesto es validado utilizando vídeos de personas con graduación de parálisis facial en escala de House-Brackmann (una persona por cada grado de parálisis, la escala cuenta con seis grados), proponiendo una equivalencia entre el resultado del tracking facial y la escala House-Brackmann. Finalmente, el método propuesto se evalúa de manera preliminar con tres pacientes chilenos con parálisis facial parcial con graduación desconocida, determinando un grado de parálisis coherente con la literatura. Si bien esta última evaluación no es concluyente estadísticamente, deja abierta la posibilidad de un futuro desarrollo de software y a una validación clínica del algoritmo propuesto.

Abstract

Facial paralysis affects between 15 and 40 cases per 100,000 people in the world population. Of these, 50% corresponds to Bell's palsy. This pathology is a type of paralysis of unknown origin (idiopathic paralysis) with the possibility of rehabilitation. This is the most common cause of unilateral facial paralysis. Currently, there are clinical graduation methods of the level of facial paralysis that have the disadvantages of: being subjective, ambiguous and to deliver a discrete graduation. Previous work has been done to make graduation objective using computer algorithms, the so-called non-clinical graduation methods, which are more objective and precise. Non-clinical methods have the disadvantage that they require visual markers, extensive preparation time, and are difficult to use, all of which limits their application in clinical environments. This thesis aims to propose a new method of graduation of partial facial paralysis based on an automatic facial tracking algorithm without markers or previous preparation and computationally validate it.

A new non-clinical graduation method is proposed. This method modifies an existing previous algorithm designed to perform facial tracking that detects symmetrical facial expressions, without the help of markers, based on a parametric three-dimensional model that encodes faces and symmetrical facial expressions. The proposed method detects asymmetric facial expressions using a modified three-dimensional parametric model to support asymmetric facial expressions. The facial tracking algorithm of the proposed method is compared with the original algorithm, presenting similar tracking quality, and compared to a cinematic tracking system with passive markers, showing an average difference of 1.03 mm (d.s.0.91 mm) between one passive marker and one point of the model. A validation using videos of people with graduation of facial paralysis in House-Brackmann scale is presented (one person for each degree of paralysis, the scale has six degrees), proposing an equivalence of facial tracking with House-Brackmann scale. Finally, the proposed algorithm is preliminarily evaluated with three Chilean patients with partial facial paralysis presenting an unknown graduation, determining a degree of paralysis consistent with the literature. Although the presented evaluations are not statistically conclusive in a clinical context, it leaves open the possibility of future software development and a clinical validation of the proposed algorithm.

1 Introducción

La parálisis facial afecta entre 15 a 40 casos por cada 100.000 personas en la población mundial (Samsudin et al. 2014). La mitad de estos casos corresponden a parálisis de Bell, que es un tipo de parálisis facial parcial de origen desconocido caracterizada por un bajo impulso motriz del nervio facial. Es la causa más frecuente de parálisis facial parcial, que resulta en debilidad muscular de un lado de la cara. Esta patología tiene dos *peak* de incidencia: entre los 30 y 50 años, y entre los 60 a 70 años. En la rehabilitación de esta disfunción se consideran aspectos psicológicos y fisiológicos. La recuperación de la movilidad y la expresión facial, resulta en mejoras de salud, auto estima, aceptación social y calidad de vida.

Para determinar si una persona con parálisis facial padece de parálisis de Bell se realiza una evaluación clínica. Una vez descartada cualquier causa conocida de parálisis, se concluye que es una parálisis de Bell. El diagnóstico incluye una graduación de la parálisis facial. Para realizar la graduación existen diferentes métodos y escalas siendo la más utilizada la escala House-Brackmann (HB) (Figura 1) ya que no utiliza un soporte computacional para la evaluación y que tiene una interpretación directa del médico, estos métodos de graduación se conocen como métodos de graduación clínicos (Samsudin et al. 2014). Debido a la interpretación del médico, los métodos clínicos presentan una falta de objetividad en la graduación. Por otro lado existen métodos que se apoyan en un soporte computacional, conocidos como métodos de graduación no clínicos. Estos entregan una medición objetiva pero difícil de implementar en entornos clínicos debido al tiempo de preparación del paciente, el uso de marcadores y dificultad de uso. Existen también métodos de graduación clínica que no requieren el uso de marcadores visuales, lo que reduce el tiempo de preparación para la medición, haciéndolos más atractivos pero aún insuficientes para el uso clínico debido a que requieren el uso de equipamiento complejo como un casco y un soporte para una cámara (Wang et al. 2004), o en el mejor caso no han sido validados aún en más de un centro (Haase et al. 2015).

En este contexto, el desarrollo de un sistema de cuantificación del grado de parálisis apoyado por un soporte computacional puede ayudar al personal clínico a mejorar la rehabilitación de los pacientes. Este sistema debería ser utilizable por personal no especializado, entregar una medición interpretable por personal no clínico, requerir poco tiempo en la preparación y ser preciso.

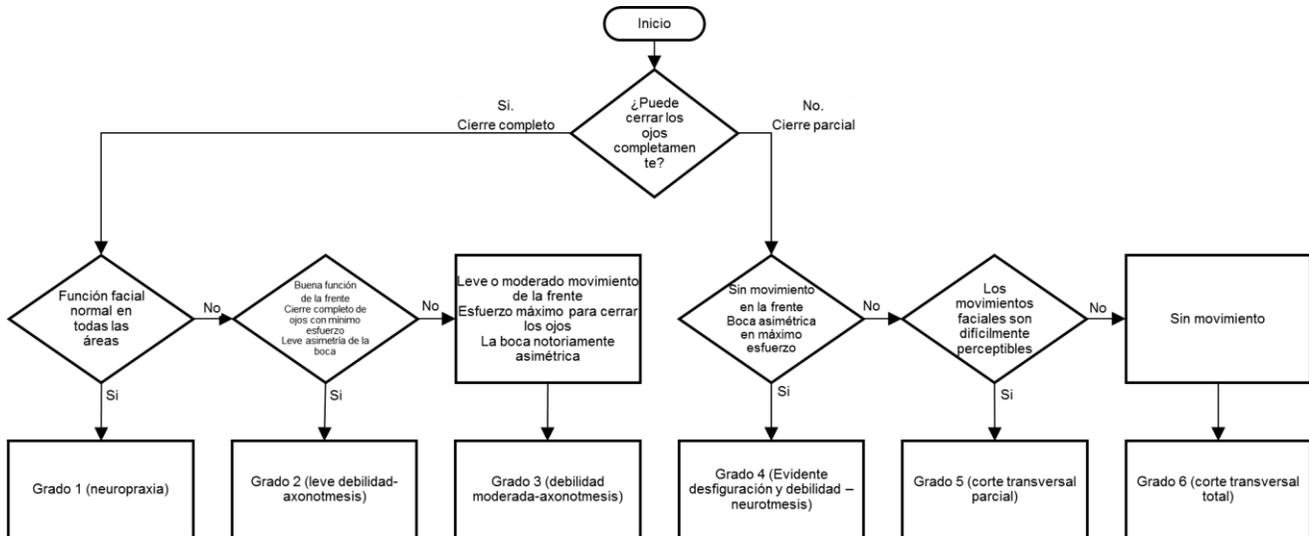


Figura 1: Sistema simplificado de graduación de parálisis del nervio facial de House-Brackmann.

Después de ejecutar cinco expresiones faciales, el médico sigue este árbol de decisión. El movimiento del cierre de los ojos determina la ubicación del daño en el nervio facial. Si el paciente cierra completamente los ojos, su grado de parálisis será 1, 2 o 3. Si el paciente cierra parcialmente los ojos, su grado de parálisis será 4, 5 o 6. Para determinar cuál de estos niveles corresponde al paciente el médico observa las funciones faciales que abarcan las cinco expresiones ejecutadas (adaptado de Aik Kah et al. 2011).

1.1 Evaluación clínica de la parálisis facial

El procedimiento estándar de evaluación incluye historia y examen clínico, investigaciones relevantes y tratamiento. Los diagnósticos de la parálisis del nervio facial son numerosos existiendo más de cincuenta causas documentadas entre ellas: la parálisis neural facial motora superior (Upper Motor Neuron Facial Paralysis – UMNL) causada por patologías intracraneales y la parálisis neural facial motora inferior (Lower Motor Neuron Facial Paralysis – LMNL) causada generalmente por patologías sistémicas o locales (Figura 2) (Aik Kah et al. 2011).

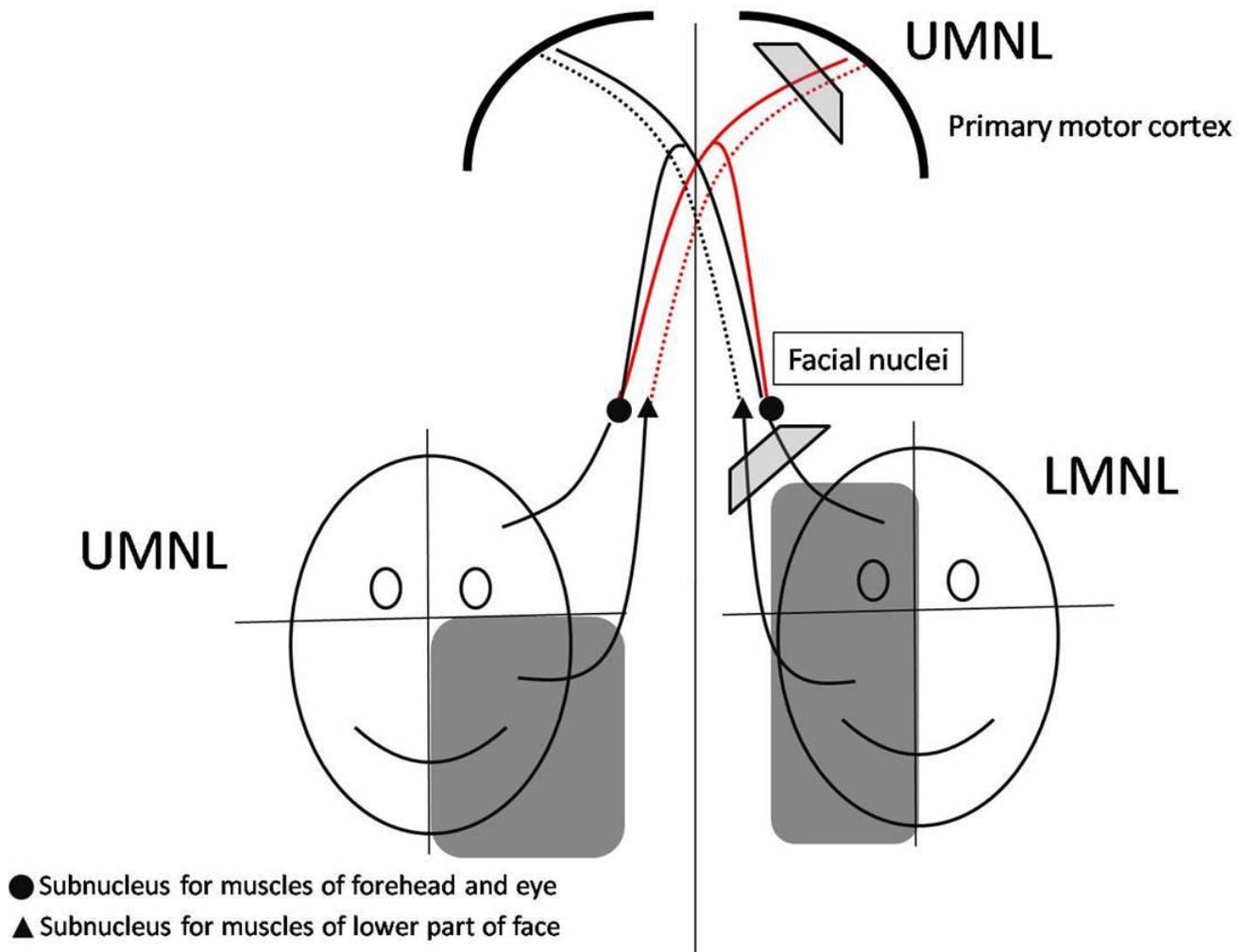


Figura 2. Inervación de la frente y los ojos. Dependiendo del lugar del daño del nervio facial, se manifiestan diferentes zonas afectadas en el rostro. Si existe un daño intracraneal, es una parálisis neural facial motora superior (UMNL), el paciente puede mover la frente debido a que la inervación es bilateral, es decir está conectada a ambos hemisferios cerebrales. Si existe una patología local, es una parálisis neural facial motora inferior (LMNL), el paciente presenta dificultad para mover la mitad del rostro (original de Aik Kah et al. 2011).

El examen clínico debe incluir un examen neurológico y sistémico, otoscopia, examen de piel, nodos linfáticos, y glándula parótida. El diagnóstico de la parálisis de Bell se considera después de excluir cualquier otra posible causa. Existen nueve señales que indican que la parálisis facial no sería parálisis de Bell: (1) si se presenta dolor facial, (2) debilidad en las extremidades, (3) parestesia en las extremidades, (4) signos cerebelosos, (5) otros nervios craneales involucrados, (6) inicio gradual de la debilidad facial, (7) antecedentes de debilidad facial, (8) cáncer de piel, y (9) parálisis facial prolongada por más de seis meses.

1.2 Métodos de graduación de la parálisis facial

Samsudin 2014 presenta una revisión cuyo objetivo es proveer a los investigadores información de los métodos propuestos en varias publicaciones con el fin de desarrollar un sistema estándar de graduación de parálisis facial globalmente aceptado (Samsudin 2014). En esta revisión se comparan más 18 publicaciones entre los años 1996 a 2012, y los métodos de graduación presentados en estas se clasifican en dos categorías según su enfoque: métodos clínicos y métodos no clínicos. A continuación se explican brevemente ambos enfoques.

1.2.1 Métodos de graduación clínicos

Los métodos clínicos no involucran el uso de equipo computacional especializado y se basan en la evaluación subjetiva de personal clínico utilizando escalas para cuantificar el nivel de parálisis. Dependiendo del método, los criterios y puntajes son variables. Por ejemplo el sistema Yanagihara se basa en diez diferentes expresiones faciales que reciben un puntaje de cero a cuarenta puntos, mientras que HB se basa en una escala de seis niveles (House-Brackmann Scale o House-Brackmann Score) para evaluar cinco expresiones faciales (Samsudin 2014). Este último método fue adoptado por la *North American Standard* para evaluar parálisis facial. Si bien es sencillo de usar, es discontinuo (es decir utiliza categorías discretas, ver Sargent et al. 1998), ambiguo y no es fácil de memorizar, por lo que los médicos deben recurrir a textos de referencia para evaluar a sus pacientes. Cabe destacar que Satoh et al. 2000 mostró la correspondencia entre el sistema Yanagihara y HB. En general los métodos clínicos tienen una interpretación directa por el médico, pero tienen importantes desventajas tales como: subjetividad, ambigüedad, difícil uso, y la medición en escalas discretas.

1.2.2 Métodos de graduación no clínicos

En contraste, los métodos no clínicos utilizan software computacional que permite una evaluación de parálisis facial objetiva (Samsudin et al. 2014). Desde 1977 (Samsudin et al. 2014) se han desarrollado múltiples métodos basados en detección de puntos destacados (Dong J et al. 2008), distancia entre puntos de interés (Anguraj et al. 2012) o reconocimiento de regiones (Liu et al. 2010). También existen métodos basados en expresiones faciales que son reconocidas automáticamente por un software (Haase et al. 2015). Algunos métodos

requieren el uso de marcadores visuales lo que demora el proceso de diagnóstico (Anguraj et al. 2012). Los marcadores visuales pueden ser activos o pasivos (Tzou 2012). Los métodos de graduación no clínicos basados en procesamiento de imágenes sin marcadores superan esta desventaja, pero requieren un soporte adicional, como un casco y una cámara adosada a este (Wang et al. 2004) o requieren un tiempo extenso para pre-procesar los vídeos o fotografías antes de que puedan ser utilizados por un software (McGrenary 2005, Dong J et al. 2008). Existe software avanzado para analizar vídeos de pacientes con parálisis facial que ha sido validado contra mediciones manuales y que permite un resultado en un tiempo promedio de 1,3 minutos (Hadlock 2012). También se pueden utilizar fotografías de un paciente realizando un ejercicio facial y comparando las diferencias entre puntos destacados del rostro (Dong J et al. 2008). Algunos métodos buscan una equivalencia con la escala de HB (Wang et al. 2004), mientras que otros proponen su propia medida sin procurar una equivalencia (Dong J et al. 2008, Hadlock 2012). Si bien los métodos no clínicos son objetivos y precisos, estos presentan desventajas como: requerir marcadores visuales, tiempo de preparación, pre-procesamiento de los vídeos y fotografías, y/o difícil uso. Lo que tiene como consecuencia su bajo uso en un entorno clínico.

1.2.3 Comparación de métodos no clínicos actuales.

Sobre la base del estudio de Samsudin 2014 y Haase et al. 2015, se comparan los métodos no clínicos basados en procesamiento de imágenes con diferentes enfoques. Algunos requieren equipo adicional, como un casco con una cámara adosada (Wang et al. 2004) o requieren un tiempo extenso para pre-procesar los vídeos o fotografías antes de que puedan ser utilizados por un software (McGrenary 2005, Dong J et al. 2008). Existe software para analizar vídeos de pacientes con parálisis facial que ha sido validado contra mediciones manuales y que permite un resultado en un tiempo promedio de 1,3 minutos (Hadlock 2012). También se pueden utilizar fotografías de un paciente realizando un ejercicio facial y comparando las diferencias entre puntos destacados del rostro (Dong J et al. 2008). Algunos métodos buscan una equivalencia con la escala de HB (Wang et al. 2004), mientras que otros proponen su propia medida sin procurar una equivalencia (Dong J et al. 2008, Hadlock 2012). La Tabla 1 muestra una comparación entre los métodos mencionados, sus enfoques, pacientes evaluados, mímicas utilizadas y resultados obtenidos.

Publicación	Wang 2004	McGrenary 2005	Dong J et al 2008	Tzou 2012	Hadlock 2012
Muestra	30	22	10	222	50
Requiere marcadores	NO	NO	NO	SI	NO
Preproceso	REDUCCIÓN	N.S.I.	SEGMENTACIÓN	N.S.I.	N.S.I.
Accesorios	CASCO	NO	NO	ESPEJOS	NO
Método	Factores de densidad y expresión	Red neuronal artificial	Análisis de imágenes	Estereoscopia	Análisis de imágenes
Ejercicios	4	6	5	9	5
Medidas	Izquierda vs Derecha	Sustracción de imágenes	Desplazamientos de puntos	Seguimiento 3D	Características faciales
Resultado	Mimica/Reposo ojos megillas y boca	ECM máximo de 26%.	Promedio de todas las distancias.	Discusión sobre el uso clínico del software	Software vs manual
Equivalencia HB	SI	NO	NO	N.A.	N.A.

Tabla 1: Comparación de publicaciones sobre métodos no clínicos de diagnóstico de parálisis facial. Cada columna corresponde a una publicación. Cada fila es una característica destacada de la publicación. La fila "Muestra" indica la cantidad de vídeos de pacientes sobre los que se hizo el estudio. A continuación si requiere marcadores o no, el tipo de pre proceso requerido (por ejemplo, Wang 2004 requiere que la imagen sea reducida a 64x64 pixeles y Dong J et al 2008 realiza una segmentación), cuáles accesorios son requeridos, breve descripción del método utilizado, cantidad de ejercicios realizados por los pacientes, medidas o cálculos realizados (por ejemplo, Wang 2004 realiza una resta de la mitad izquierda y la derecha de la imagen del paciente), resultados obtenidos y si tienen equivalencia en escala HB. Abreviaciones: N.S.I.: No se indica, N.A.: No aplica. (Elaboración propia).

En resumen, si bien se han reportado en la literatura métodos no clínicos objetivos y precisos, no existe un método que no requiera marcadores o accesorios y que tenga una equivalencia en la escala clínica HB. Lo que probablemente tiene como consecuencia su bajo uso en un entorno clínico. El desarrollo de un sistema basado en algoritmos que combine las mejores características de los métodos descritos ayudaría a los clínicos a medir

objetivamente el progreso de la rehabilitación. Este sistema debería ser utilizable por personal no especializado siendo de fácil interpretación, requerir poco tiempo en la preparación, ser preciso y ayudar a varios tipos de afecciones faciales (Samsudin et al. 2014).

1.3 Métodos de tracking facial

El tracking facial consiste en determinar la posición, la orientación y la expresión facial del rostro de un paciente en una filmación (en adelante, el término filmación hace referencia tanto a un vídeo almacenado como a una captura en vivo, y el término “pose” se refiere a la posición y orientación del rostro). Sus aplicaciones principales son vigilancia, interacción humano-computador, análisis de expresión facial entre otros (Terissi et al. 2010). El principal problema del tracking facial es que, en una filmación, el rostro humano cambia constantemente de posición y de expresión.

En términos generales, existen dos enfoques para hacer tracking en una filmación: i) una manera es por medio del tracking de características (principalmente flujo óptico), y ii) basada en modelos deformables. A continuación, se describen ambos enfoques

1.3.1 Tracking basado en flujo óptico

El flujo óptico (FO) es el patrón del movimiento aparente de un objeto en una filmación entre dos o más frames consecutivos. El patrón de movimiento es un campo vectorial donde cada vector representa el desplazamiento de un punto de interés entre dos frames (Figura 3). Entre sus aplicaciones están: estimar la estructura del movimiento, compresión de vídeo y estabilización de vídeos (OpenCV, 2017). El FO se basa en los siguientes supuestos:

- La intensidad de un pixel de un objeto no cambia entre dos frames consecutivos.
- Los pixeles cercanos al punto que se está siguiendo tienen un movimiento similar al punto de interés.

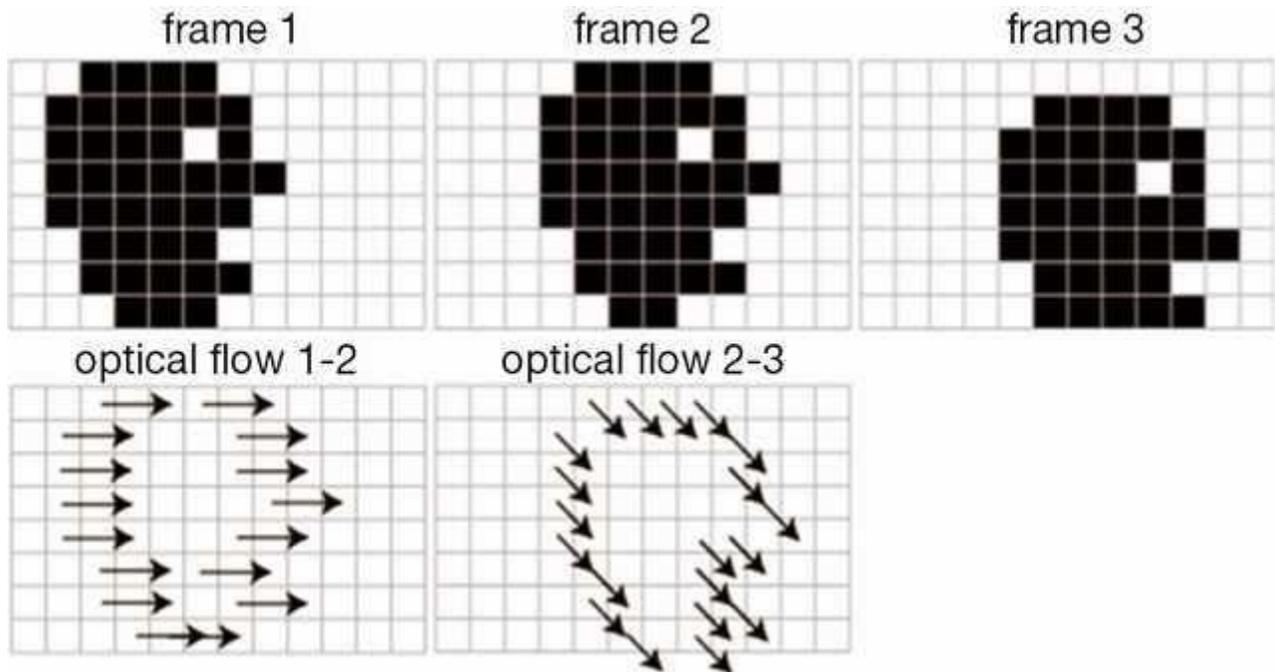


Figura 3: Ejemplo simple de flujo óptico entre tres frames. La figura en el frame 1 está desplazada un pixel a la derecha en el frame 2. En el frame 3, está desplazada un pixel a la derecha y uno hacia abajo. El flujo óptico entre el frame 1 y 2 es un campo vectorial que apunta a la derecha. El flujo óptico entre el frame 2 y 3 es un campo vectorial que apunta hacia abajo y a la derecha (imagen original de <http://tcr.amegroups.com/article/view/3200/html>).

La formulación matemática es la siguiente: considerando un pixel en la posición (x, y) de intensidad $I(x, y, t)$ que depende del tiempo t , que se desplaza una distancia (dx, dy) en el siguiente frame después de un tiempo dt , y asumiendo que la intensidad de este pixel no cambia, se puede describir el desplazamiento como en la Ecuación 1.

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (1)$$

Asumiendo desplazamientos pequeños y haciendo una aproximación por series de Taylor, se obtiene la Ecuación 2. Los gradientes de imagen I_x , I_y y I_t (Ecuación 3) están multiplicados por un vector (u, v) (Ecuación 4) que indican la dirección y magnitud del desplazamiento del pixel. A las Ecuaciones 2, 3 y 4 se les conoce como restricción de la conservación del brillo.

$$I_x u + I_y v + I_t = 0 \quad (2)$$

$$I_x = \frac{\partial I}{\partial x}; I_y = \frac{\partial I}{\partial y}; I_t = \frac{\partial I}{\partial t} \quad (3)$$

$$u = \frac{\partial x}{\partial t}; v = \frac{\partial y}{\partial t} \quad (4)$$

El vector (u, v) es desconocido y no es posible determinarlo usando una única ecuación. Para resolver este problema existen varios métodos, uno de ellos es Lucas-Kanade. Este método, apoyado en los supuestos descritos anteriormente (misma intensidad del pixel y comportamiento similar de su vecindad), considera los nueve puntos alrededor del punto de interés, incluyéndolo. Es posible determinar I_x , I_y y I_t para los nueve puntos. Luego el vector (u, v) queda sobredeterminado. La Ecuación 5 muestra el desarrollo del método Lucas-Kanade para determinar el vector (u, v) (el índice i de las sumatorias señala cada uno de los nueve pixeles alrededor del punto de interés).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i I_{x_i}^2 & \sum_i I_{x_i} I_{y_i} \\ \sum_i I_{x_i} I_{y_i} & \sum_i I_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_{x_i} I_{t_i} \\ -\sum_i I_{y_i} I_{t_i} \end{bmatrix} \quad (5)$$

Este método funciona muy bien si el desplazamiento del pixel de interés está dentro de una ventana pequeña de 3x3 pixeles. Para detectar un desplazamiento mayor, la imagen analizada es reducida a una escala menor. Esto elimina los desplazamientos pequeños de las escalas mayores y hace evidentes en la ventana de 3x3 los desplazamientos grandes. En OpenCV, este método es llamado FO piramidal y está implementado por la función *calcOpticalFlowPyrLK*.

1.3.2 Tracking basado en modelos

El objetivo de este método, es estimar la pose de la cámara con el fin de seguir un objeto ajustando seis grados de libertad dada una imagen en dos dimensiones y un modelo en tres dimensiones. Si se utiliza un modelo deformable, los grados de libertad aumentan incluyendo los parámetros de deformación.

En visión computacional estimar la pose de una cámara a partir de la correspondencia de puntos en tres dimensiones que describen el modelo y puntos en dos dimensiones que

describen su proyección en la imagen es un problema fundamental. La versión más general del problema requiere estimar los seis grados de libertad de la pose (tres componentes para indicar la posición y tres para indicar la orientación) y los parámetros de calibración (distancia focal horizontal, relación de aspecto y distorsión) (ver Figura 4, OpenCV 2015). Notar que la pose puede determinarse asumiendo que la cámara está fija y el objeto observado se desplaza o viceversa.

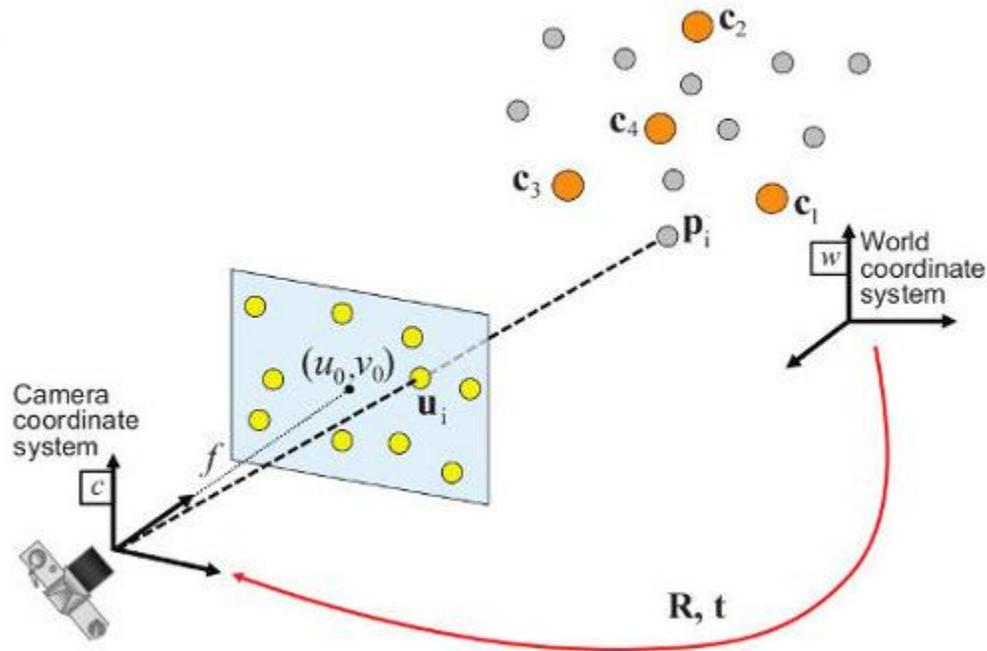


Figura 4: Estimar pose de una cámara. Dado un conjunto de puntos en tres dimensiones p_i expresados en un sistema de referencia global (*World coordinate system*) y sus proyecciones correspondientes u_i en dos dimensiones en una imagen, se busca determinar la pose de una cámara R y t con respecto al *World coordinate system* y la distancia focal f (original de OpenCV 2015).

Asumiendo que la cámara es una cámara “pinhole”, la correspondencia de los puntos del modelo y su proyección se determinan por la Ecuación 6. Los parámetros de calibración son representados por la matriz A y la pose por la matriz de rotación-traslación $[R|t]$. Esta matriz está compuesta por la matriz de rotación R que describe la orientación y el vector de traslación t . M es la coordenada de un punto en tres dimensiones. Esta se denota M' porque es un vector de cuatro componentes, con cuarta componente fija en 1. m es la coordenada

de la proyección en píxeles en dos dimensiones. Se denota m' porque es un vector de tres componentes con la tercera componente fija en 1. s es un factor de escala.

$$sm' = A[R|t]M' \quad (6)$$

La forma de las matrices está detallada en la Ecuación 7. Donde (X, Y, Z) es la coordenada de un punto en tres dimensiones, (u, v) es la coordenada de la proyección de este punto en píxeles, (c_x, c_y) es la coordenada del punto principal (generalmente es el centro de la imagen) y, (f_x, f_y) es la distancia focal expresada en píxeles.

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (7)$$

El problema de estimación de la pose se divide en tres etapas. La primera etapa consiste en determinar los parámetros de la cámara, es decir, calibrar la cámara determinando la distancia focal, relación de aspecto y la distorsión propia de la cámara. La segunda etapa consiste en determinar la pose de la cámara (posición y orientación) respecto del objeto observado. La tercera etapa consiste en determinar los parámetros de deformación del modelo facial para realizar el ajuste fino del modelo deformado y su proyección.

Para un sujeto en una filmación, Terissi et al. 2010 utiliza una funcionalidad disponible en OpenCV, llamada *solvePnP*, para determinar la pose del rostro, utilizando los puntos del modelo, los puntos seguidos usando FO y los parámetros de la cámara (ver Figura 5). En dicho trabajo se utiliza el algoritmo de Levenberg-Marquardt (Levenber-Marquardt 2017) para determinar la orientación y desplazamiento de los puntos del modelo en relación a la proyección de estos, que idealmente debe coincidir con el FO. El resultado de *solvePnP* son dos vectores, uno de rotación y uno de traslación. La magnitud del vector de rotación representa el ángulo de rotación sobre el vector de rotación normalizado. El vector de traslación representa el desplazamiento en los ejes cartesianos del modelo.

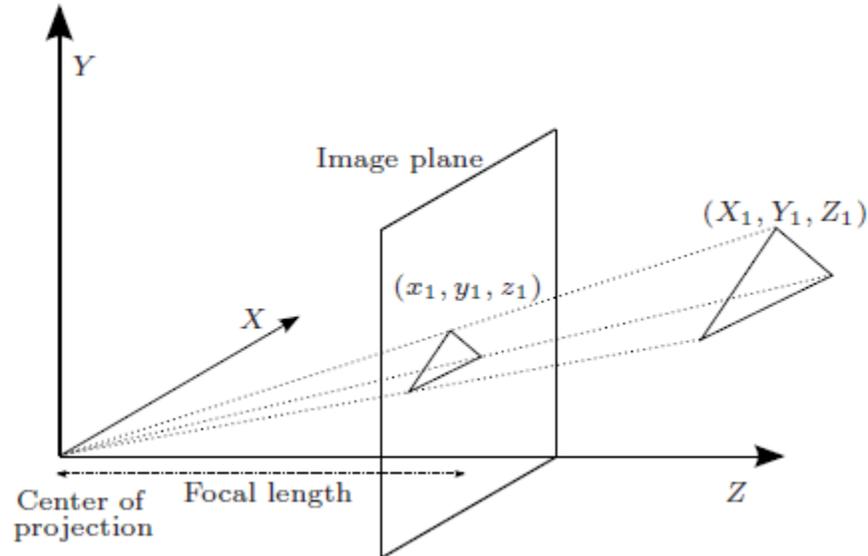


Figura 5: Esquema de proyección de perspectiva. El modelo de coordenadas (X_i, Y_i, Z_i) es proyectado sobre un plano de imagen con coordenadas (x_i, y_i, z_i) (con $z_i = constante$ ya que es una imagen en dos dimensiones). La determinación de los parámetros de la cámara consiste en buscar aquella combinación de distancia focal, relación de aspecto y distorsión que entregue para cada coordenada (X_i, Y_i, Z_i) la proyección (x_i, y_i, z_i) que más se asemeje a una imagen de referencia (original de Terissi et al. 2010).

Una vez determinada la pose del modelo tridimensional, el ajuste de su proyección no es exacto ya que el sujeto filmado puede estar haciendo algún gesto. El modelo tridimensional utilizado por Terissi et al. 2010 tiene parámetros de deformación. El ajuste de estos parámetros se realiza utilizando el algoritmo de Levenberg-Marquardt, cambiando estos parámetros en cada iteración para que la proyección actualizada coincida con los puntos del FO.

1.3.3 Modelo Candide-3

La máscara Candide-3 (Ahlberg 2001) fue creada para codificar rostros humanos basados en un modelo. Para ello, es configurada por unidades de deformación (“SU” por las siglas en inglés de “*Shape Units*”) que ponderadas por un valor escalar, deforman el modelo básico moviendo los puntos que lo describen en diferentes direcciones. La máscara es controlada por unidades de acción (“AU” por las siglas en inglés de “*Action Units*”) globales y locales. Las unidades de acción globales corresponden a la posición y rotación (también llamada “pose”) respecto de los planos cartesianos. Las unidades de acción locales controlan las expresiones faciales al ser ponderadas por un valor escalar, y deforman el modelo

desplazando los puntos que lo describen en diferentes direcciones de manera similar a las SU. Las unidades de acción locales están agrupadas en vectores (“AUV” por las siglas en inglés de “*Action Units Vector*”). El algoritmo de Terissi et al. 2010 utiliza una versión modificada de esta máscara, configurada y controlada de manera similar. Ambos modelos pueden reflejar expresiones simétricas, es decir, las AUV deforman el modelo simétricamente respecto del eje medio de la cara. La Figura 6 muestra una representación del modelo Candide-3 hecha en una planilla de cálculo para su estudio al inicio de este proyecto.

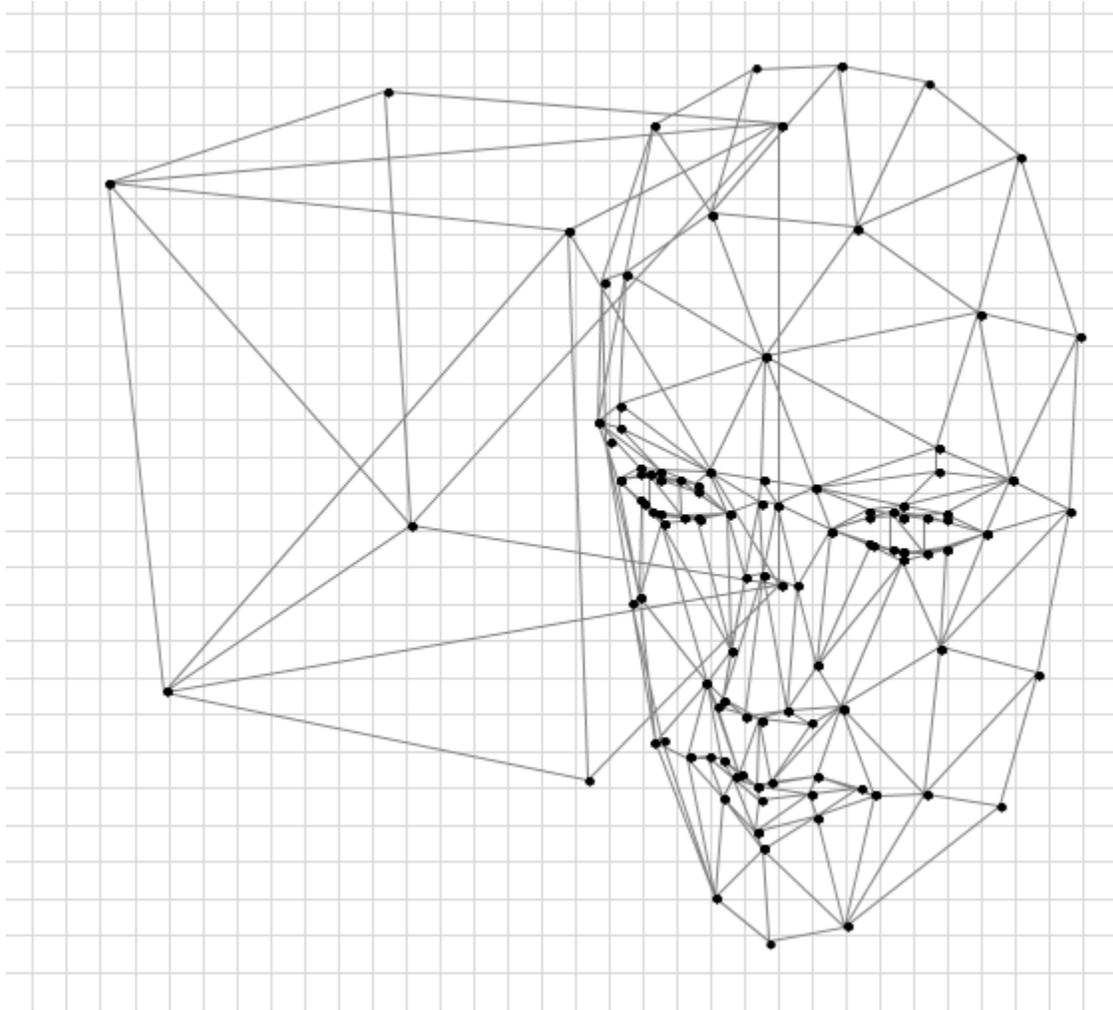


Figura 6: Máscara Candide-3 representada en una planilla de cálculo para su estudio al principio de este proyecto. El cubo dibujado al lado de la máscara es una referencia. Mientras un vértice del cubo está en el origen de coordenadas, el vértice opuesto está en coordenadas (1,1,1) (elaboración propia).

1.4 Problema

En resumen, para cuantificar el grado de parálisis facial parcial, la principal desventaja de los métodos clínicos es su subjetividad. Por otra parte, los métodos no clínicos si bien reducen la subjetividad, resultan difíciles de aplicar en la práctica clínica. El desarrollo de un sistema basado en algoritmos que combine las mejores características de ambos métodos ayudaría a los profesionales clínicos a medir objetivamente el progreso de la rehabilitación. Este sistema debería ser utilizable por personal no especializado siendo de fácil interpretación, requerir poco tiempo en la preparación, ser preciso y ayudar a varios tipos de afecciones faciales (Samsudin et al. 2014).

2 Hipótesis y objetivos

Hipótesis: La cuantificación del grado de parálisis mediante cámaras convencionales y algoritmos de tracking facial adaptados permite una evaluación objetiva, fácil de usar, y en escalas continuas.

Objetivo general: Diseñar un método no clínico, basado en un algoritmo de tracking facial automático para la cuantificación del grado de parálisis facial parcial, validado computacionalmente.

Objetivos específicos:

1. Comparar métodos actuales de cuantificación del grado de parálisis facial y determinar ventajas y desventajas (ver Introducción).
2. Evaluar una herramienta de tracking existente.
3. Diseñar un algoritmo de medición de parálisis facial parcial, con soporte de asimetría en los movimientos de la cara.
4. Validar el algoritmo computacionalmente.

3 Materiales y métodos

3.1 Bases de datos experimentales

Se cuentan con tres tipos de filmaciones diferentes. La primera es una filmación hecha con un sistema de seguimiento cinemático con marcadores pasivos. La segunda es una filmación disponible en Internet que presenta diferentes pacientes con parálisis facial y su graduación respectiva en escala HB. La tercera son filmaciones de tres pacientes con parálisis facial parcial chilenos, de los que no se cuenta con su grado de parálisis en escala HB. Los tres tipos de filmaciones se describen a continuación.

3.1.1 Filmación con el sistema de seguimiento cinemático con marcadores pasivos

La filmación con el sistema de seguimiento cinemático con marcadores pasivos fue adquirida gracias al financiamiento del proyecto STIC-AmSud (15STIC04) (en adelante, *marcadores pasivos* o *marcador pasivo* hace referencia a esta filmación). Esta filmación se utiliza para comparar la precisión del seguimiento por FO con el seguimiento utilizando el marcador. Se cuenta con información precisa de la posición de estos marcadores gracias a que el registro se realiza a 110 hz con un sistema fabricado por Vicon Motion System Ltd. El vídeo fue filmado con una cámara Intel® Realsense™ a 120 frames por segundo y muestra una sola expresión facial: un gesto de sorpresa. La Figura 7 muestra los marcadores pasivos en el rostro de un sujeto.



Figura 7: Rostro con marcadores ópticos pasivos. Los marcadores pasivos se ven como puntos blancos en la fotografía (la filmación fue adquirida gracias al financiamiento del proyecto STIC-AmSud (15STIC04)).

3.1.2 Filmación de pacientes con graduación

Hasta el momento de la redacción de este documento, no ha sido posible acceder a una base de datos de pacientes con parálisis facial parcial graduados en escala de HB. En Youtube, existe el vídeo *House Brackmann - Facial Nerve Grading* en el canal del Dr. Murali Chand Nallamothe. El vídeo es presentado por el Facial Nerve Committee of A.A.O. – H.N.S. & The House Ear Institute, Los Angeles, CA. (Chand 2016). El vídeo presenta un paciente con parálisis facial parcial por cada uno de los seis grados de la escala HB. La Figura 8 muestra los pacientes y sus graduaciones. La leyenda con la graduación está superpuesta al rostro del paciente dificultando algunas mediciones. El vídeo está almacenado a 640x480 pixeles por frame, a 31 frames por segundo. La Figura 8 muestra los rostros de los sujetos con sus graduaciones en escala HB extraídas el vídeo.

El vídeo fue dividido en seis secciones, una por cada paciente. Se buscó un key frame donde el rostro esté aproximadamente relajado y no se superpongan las etiquetas con la graduación para utilizarlo como referencia. Si bien la graduación está en escala de HB, los pacientes no realizan todos los ejercicios requeridos por la graduación. El vídeo no está segmentado para saber exactamente qué ejercicio se está ejecutando.



Figura 8: Pacientes del vídeo House Brackmann - Facial Nerve Grading en el canal del Dr. Murali Chand Nallamothu. Cada paciente representa un grado de parálisis facial específico. Las imágenes corresponden a un momento del vídeo en que figura la glosa con la graduación y realizan una expresión facial en que la parálisis es relativamente evidente. El sujeto 1 (a) presenta una parálisis facial de grado 1. La asimetría del movimiento facial es difícil de notar en el vídeo. El sujeto 2 (b) tiene las cejas cubiertas por el cabello, pero es evidente que puede levantar ambas cejas. En la fotografía se aprecia asimetría en la boca, ligeramente torcida al lado derecho del paciente. De hecho, presenta parálisis hipertónica del lado derecho, lo que explica la demarcación de las mejillas. El sujeto 3 (c) también presenta parálisis del lado derecho. El sujeto 4 (d) presenta parálisis del lado izquierdo con una notable asimetría en el movimiento. Si bien puede levantar la ceja derecha completamente, la ceja izquierda ejecuta aparentemente el movimiento contrario (se mueve hacia abajo). El sujeto 5 (e) presenta parálisis flácida del lado derecho, con la ausencia de movimiento de la ceja derecha. El sujeto 6 (f) presenta una parálisis flácida del lado derecho con notoria asimetría del rostro (imágenes extraídas de Chan 2016).

3.1.3 Filmación de pacientes sin graduación

Los pacientes que participan en los vídeos han firmado consentimiento informado y están aprobados por el comité de ética del Servicio de Salud Norte (parte del proyecto de investigación por Carlos Cruz, Kinesiólogo del Hospital San José). Cada vídeo está dividido en secciones con cada paciente realizando diferentes ejercicios: arrugar la frente, cerrar los ojos, arrugar la nariz, sonreír, expresar sorpresa y lanzar un beso.

Los vídeos están grabados a una resolución de 1280x720 píxeles, a una tasa de 120 frames por segundo. Sin embargo, al extraer los frames para analizar, la tasa resultante es de 31 frames por segundo por limitaciones de la librería utilizada para la lectura del vídeo. La Figura 9 muestra los pacientes sin graduación.

Los vídeos de los pacientes no incluyen la evaluación del grado de parálisis facial, por lo que se aplica el método descrito por Aik Kah et. al 2011 para obtener un punto de comparación.

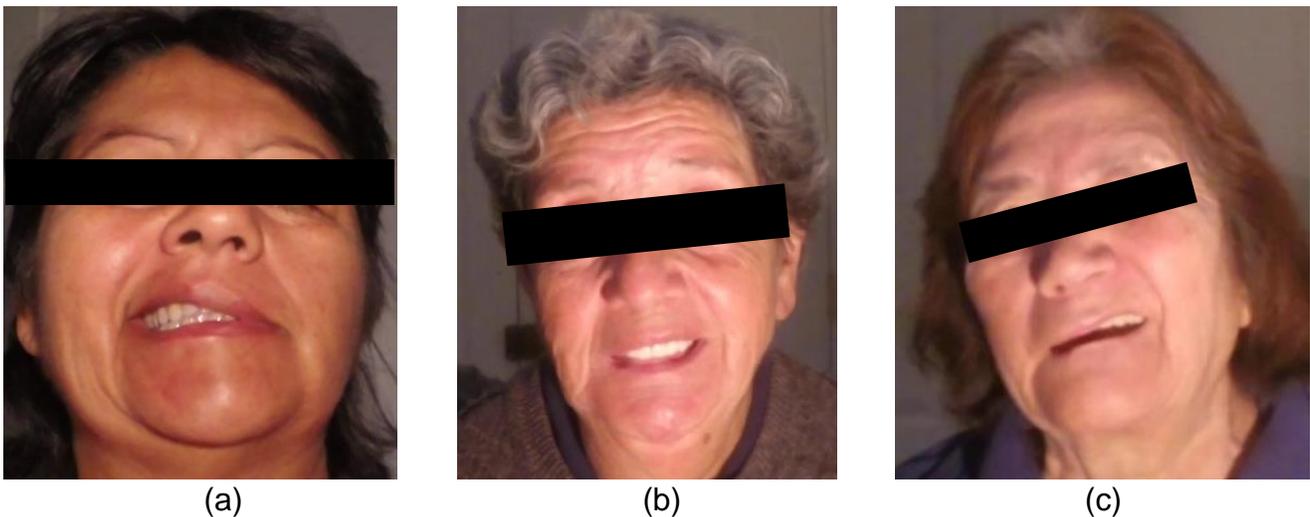


Figura 9: Pacientes sin graduación, numerados como 7 (a), 8 (b) y 9 (c) respectivamente.

Para efectos de diferenciar los sujetos de la filmación con graduación y estos pacientes en el entorno de memoria de R, los tres pacientes sin graduación son numerados como 7, 8 y 9 respectivamente.

3.2 Evaluación de una herramienta de tracking facial existente

Terissi et al. 2010 propone un algoritmo de tracking facial tridimensional usando una cámara (en adelante, *algoritmo original*). Se basa en una combinación de enfoques de tracking basado en FO y basado en modelos. El enfoque basado en FO rastrea zonas de interés como puntos clave, curvas y color de la piel. El enfoque basado en modelos utiliza un modelo 3D de la cara (basado en el modelo Candide-3) para estimar la postura y ubicación del rostro en la escena (Terissi et al. 2010).

Para dar estabilidad al tracking mediante FO, utiliza un frame de referencia (key frame) en el que la posición de las zonas de interés es conocida. Respecto de un frame actual, el algoritmo calcula el FO entre el frame anterior y el actual y el FO entre el key frame y el frame actual. Con esta información calcula la pose del rostro (ver Figura 10). Luego utilizando una librería que implementa el algoritmo de Levenberg-Marquardt, realiza el ajuste fino del nivel de activación de los AUV en el modelo usando como entrada ambos FO recién descritos (ver Figura 11). Este algoritmo ya fue probada con la base de datos de La Cascia et al. 2000.

Como salida del algoritmo, se genera un archivo que sigue a 22 puntos específicos, relevantes para un fisiatra para determinar algún problema en el movimiento de la cara de un paciente. También genera un vídeo con la superposición de la máscara Candide-3 superpuesta al rostro del paciente grabado.

El algoritmo original es probado con dos vídeos sobre una persona realizando gestos simétricos y asimétricos voluntariamente. Los vídeos tienen una resolución de 320x240 pixeles debido a una limitación técnica de la implementación del algoritmo. Se compara el movimiento de dos puntos ubicados sobre las cejas del sujeto para determinar cómo estos son reportados.

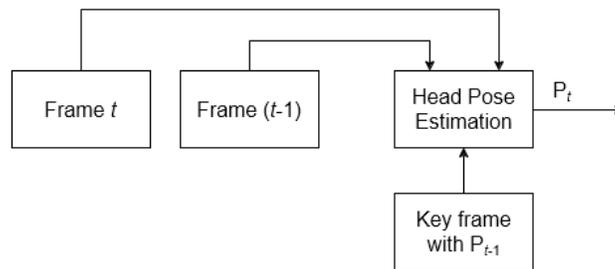


Figura 10: Algoritmo de Terissi et al. 2010 para determinar la pose. Utiliza como entrada el frame actual (Frame t), el frame anterior (Frame $(t-1)$) y el key frame anterior (Key frame with P_{t-1}), calcula el FO entre los dos frames consecutivos y el key frame anterior y el frame actual. Luego, calcula el vector de pose usando ambos flujos ópticos. El resultado es el vector de pose P_t (adaptado de Terissi et al. 2010).

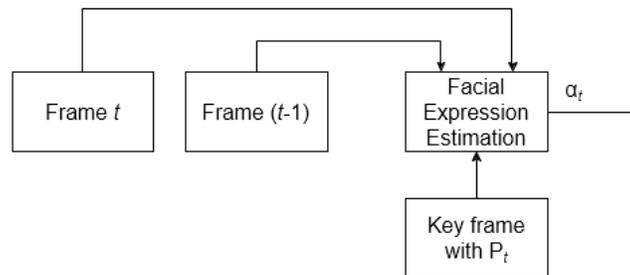


Figura 11: Algoritmo de Terissi et al. 2010 para determinar la expresión facial. Utiliza como entrada el frame actual (Frame t), el frame anterior (Frame $(t-1)$) y el key frame actualizado (Key frame with P_t). Calcula el FO entre ambos frames. Luego, calcula el vector de niveles de activación de los AUV, α_t , usando ambos flujos (adaptado de Terissi et al. 2010).

3.3 Diseñar un algoritmo de seguimiento facial con soporte de asimetría

Se propone la creación de un algoritmo para medir el grado de parálisis facial mediante una extensión del trabajo de Terissi et al. 2010. El seguimiento facial será modificado para soportar la asimetría del rostro de personas con parálisis facial. Para validar el algoritmo de seguimiento, se utilizan los videos de pacientes con parálisis facial parcial con graduación en escala de HB descritos anteriormente. Para determinar la graduación correspondiente en la escala HB, se propone utilizar los niveles de activación de los AUV de una máscara Candide-3 con soporte de asimetría.

Al igual que en Terissi et al. 2010, el algoritmo propuesto requiere una configuración inicial, sobre la base de un key frame donde el rostro del sujeto idealmente está mirando directamente a la cámara. Los parámetros físicos se ajustan una sola vez por sujeto, pero la

expresión facial inicial se ajusta en cada vídeo según el ejercicio que estuviese realizando. Frame a frame, las posiciones de los 22 puntos de interés quedan capturadas en un archivo de texto junto con el nivel de activación de los AUV. Finalmente estos archivos de texto son analizados usando R, en búsqueda de relaciones entre los niveles de activación de los AUV y el grado de parálisis de cada paciente.

En resumen, se propone un **método de graduación no clínico** (en adelante, **método propuesto**) que se compone de tres partes: un algoritmo de tracking facial, un modelo deformable con soporte de asimetría y un algoritmo de decisión que se apoya en los datos reportados por el tracking facial y la deformación del modelo (Figura 12).

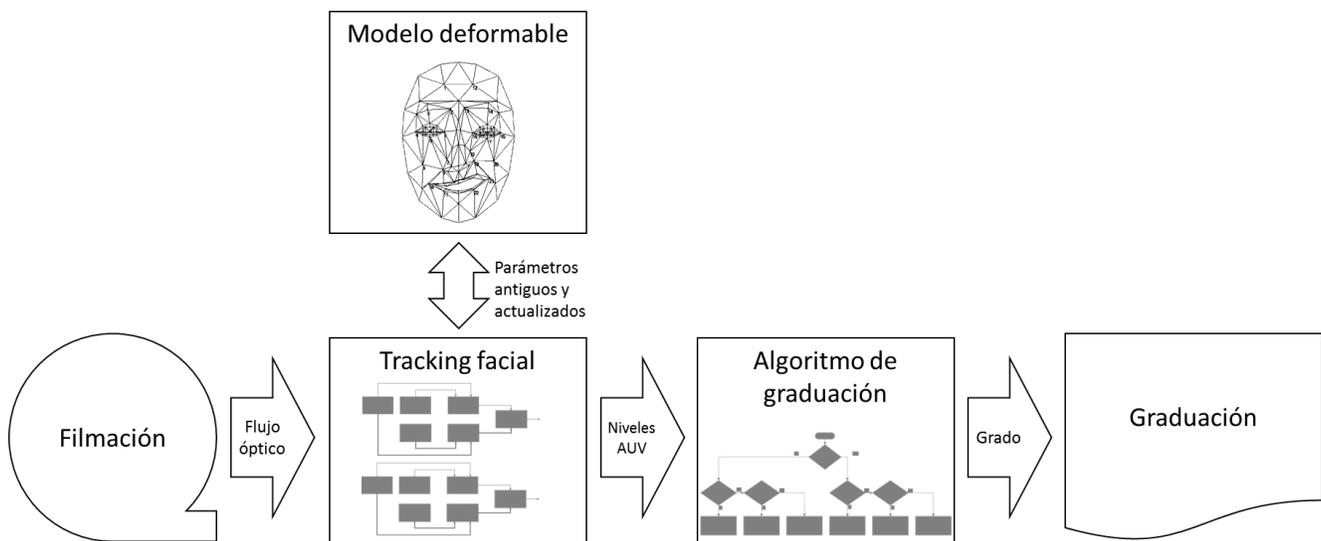


Figura 12: Diagrama general del método propuesto. El método de graduación no clínico propuesto se compone de tres partes: un algoritmo de tracking facial, un modelo deformable con soporte de asimetría y un algoritmo de decisión que se apoya en los datos reportados por el tracking facial y la deformación del modelo. Como entrada recibe una filmación y su salida es la graduación en escala HB.

3.4 Evaluación del método propuesto

El método propuesto es probado usando las bases de datos descritas en 3.1 y comparado con el método propuesto por Terissi et al. 2010. La filmación con marcadores pasivos permite una medición de la precisión del tracking en mm. La filmación de pacientes con graduación permite una caracterización de los niveles de activación de los AUV (recordar que estos niveles de activación son magnitudes sin dimensión, es decir, representan la magnitud de una deformación). Finalmente la filmación de pacientes sin graduación permite una

comparación del algoritmo contra un diagnóstico basado en la escala HB descrita en Aik Kah et al. 2011.

El resultado de la comparación con Terissi et al. 2010 es presentado en la sección 4.1 y la evaluación usando las filmaciones de pacientes graduados es presentada en la sección 4.2.4.

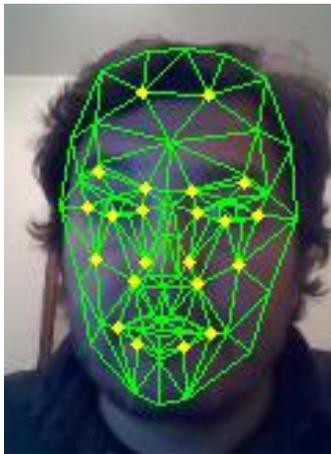
4 Resultados Obtenidos

4.1 Evaluación de herramienta de seguimiento existente

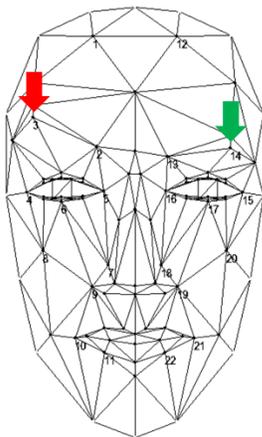
El algoritmo original fue probado con dos vídeos sobre una persona realizando gestos simétricos y asimétricos voluntariamente. Los vídeos tienen una resolución de 320x240 píxeles debido a una limitación técnica de la implementación del algoritmo.

El algoritmo entrega un registro con la posición de 22 puntos del rostro. La Figura 13 muestra la proyección vertical del punto 3 y punto 14, correspondiente al punto central de las cejas del rostro del sujeto, en un vídeo con movimientos asimétricos. Las curvas Puntos 3 y 14 corresponden a la proyección de este punto en el eje Y se superponen, demostrando que el algoritmo no entrega información distinta para los puntos señalados.

El algoritmo original demostró un seguimiento estable y robusto, pero limitado al seguimiento de gestos simétricos. Al ser probada con el vídeo de gestos asimétricos, el algoritmo original interpreta la asimetría como cambios en la posición del rostro del sujeto.

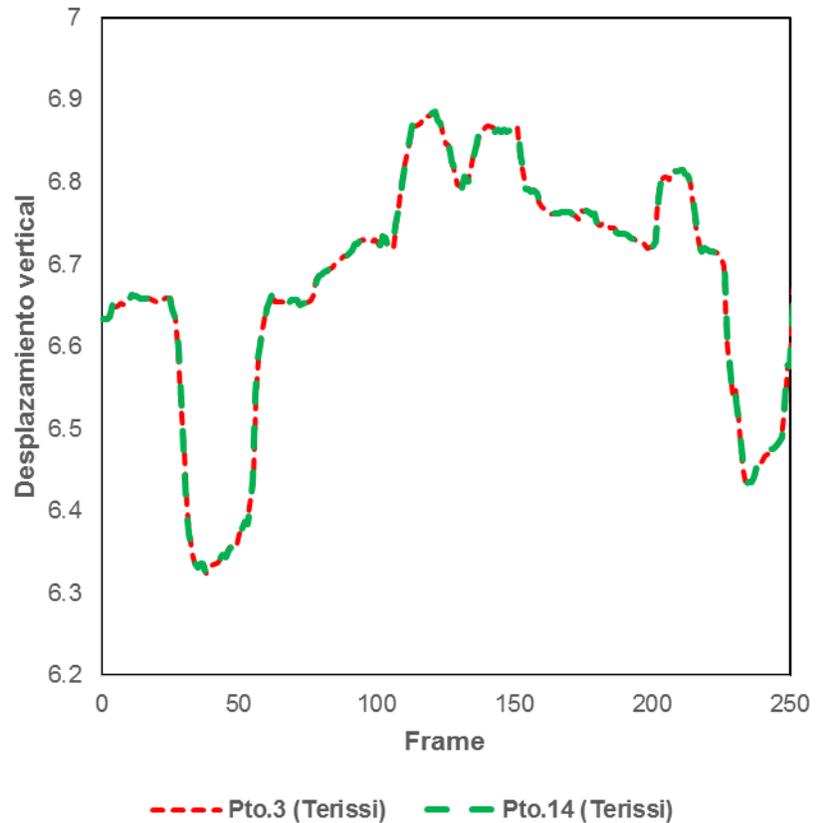


(a)



(b)

Movimiento vertical
punto 3 (ceja derecha) y
punto 14 (ceja izquierda)



(c)

Figura 13: Proyección en el eje vertical de un punto de las cejas. (a) Sujeto levantando la ceja derecha. (b) El punto 3 correspondiente a la ceja derecha, señalado con la flecha roja, se mueve verticalmente en la filmación. El punto 14 correspondiente a la flecha verde, se mantiene más o menos estático en el gesto. (c) En el gráfico se compara el registro del movimiento vertical del punto 3 y 14. El punto 3 capturado con el algoritmo original en color rojo. El punto 14 capturado con algoritmo original en color verde. El algoritmo original detecta el movimiento de manera simétrica, por lo que la curva roja y la verde se superponen.

4.2 Algoritmo de medición de parálisis facial propuesto

La idea principal que se propone en este trabajo es que si la máscara Candide-3 es modificada, dividiendo los AUV en dos grupos: un grupo que afecte al lado izquierdo de la máscara y otro que afecte al lado derecho independientemente, entonces esta puede representar expresiones faciales asimétricas. También se proponen mejorar la estabilidad del algoritmo mismo de tracking. Luego para determinar el grado de parálisis facial parcial en escala HB, se analiza la primera derivada (en adelante llamada *velocidad*) del nivel de activación de los AUV entre cada frame, comparando el lado izquierdo y derecho usando, y caracterizándolas según la graduación en la filmación cuando está disponible.

En estricto rigor, se propone un ***método de graduación no clínico (método propuesto)*** que se compone de tres partes: un modelo deformable con soporte de asimetría, un algoritmo de tracking facial y un algoritmo de decisión que se apoya en los datos reportados por el tracking facial y la deformación del modelo. A continuación, se detallan las partes de este método, comenzando por el modelo, el algoritmo de tracking facial, las mediciones necesarias para la cuantificación y el algoritmo para la graduación de la parálisis facial.

4.2.1 Máscara Candide-3 asimétrica

La máscara Candide-3 en principio codifica expresiones faciales simétricas, es decir, ambos lados del rostro del paciente se mueven en igual magnitud. Si los parámetros que codifican las expresiones faciales se dividen en dos grupos, uno que afecta al lado izquierdo de la máscara y otro al lado derecho de manera independiente, entonces es posible representar expresiones faciales asimétricas. Esta máscara con posibilidad de representar expresiones faciales asimétricas será llamada en adelante *Candide-3 asimétrica*. La Figura 14(a) muestra los AUV definidas por Ahlberg 2001. La Figura 14(b) muestra los AUV propuestas después de dividir las originales en versiones izquierda (terminadas en “L”) y derecha (terminadas en “R”). Los AUV que controlan el maxilar inferior no son divididos ya que produciría una expresión facial donde ambas quijadas se moverían independientemente.

Los AUV controlan la máscara Candide-3 asimétrica al ser ponderadas por un valor escalar, un número real, entre -2 y 2 para la mayoría de estos (no existe un límite definido para estos valores, pero valores muy extremos producen, por ejemplo, que los párpados se solapen o que el labio inferior suba hasta la frente). En adelante, el número escalar recién mencionado será llamado *el nivel de activación*. La Figura 15 muestra la máscara Candide-3 asimétrica con algunos AUV propuestos con nivel de activación 1 para el lado derecho del paciente. Los AUV asimétricos completos se encuentran en el Anexo A.

Para graduar la parálisis facial, se propone usar un sistema simplificado de graduación en escala de HB (Aik Kah et al. 2011), midiendo el nivel de activación de los AUV de ambos lados del rostro del paciente. De acuerdo al algoritmo descrito en la Figura 1, los AUV observados según el grado de parálisis quedan descritos en la Tabla 2.

Bajo el supuesto que un paciente afectado de parálisis facial parcial puede mover un lado del rostro con dificultad, el diagnóstico del grado de parálisis puede realizarse comparando el nivel de activación de los AUV del lado paralizado contra el lado sano cuando el paciente realiza los ejercicios siguientes: cerrar los ojos, fruncir el ceño, arrugar la nariz, sonreír, mostrar los dientes, dar un beso y hacer un puchero (Haase 2015).

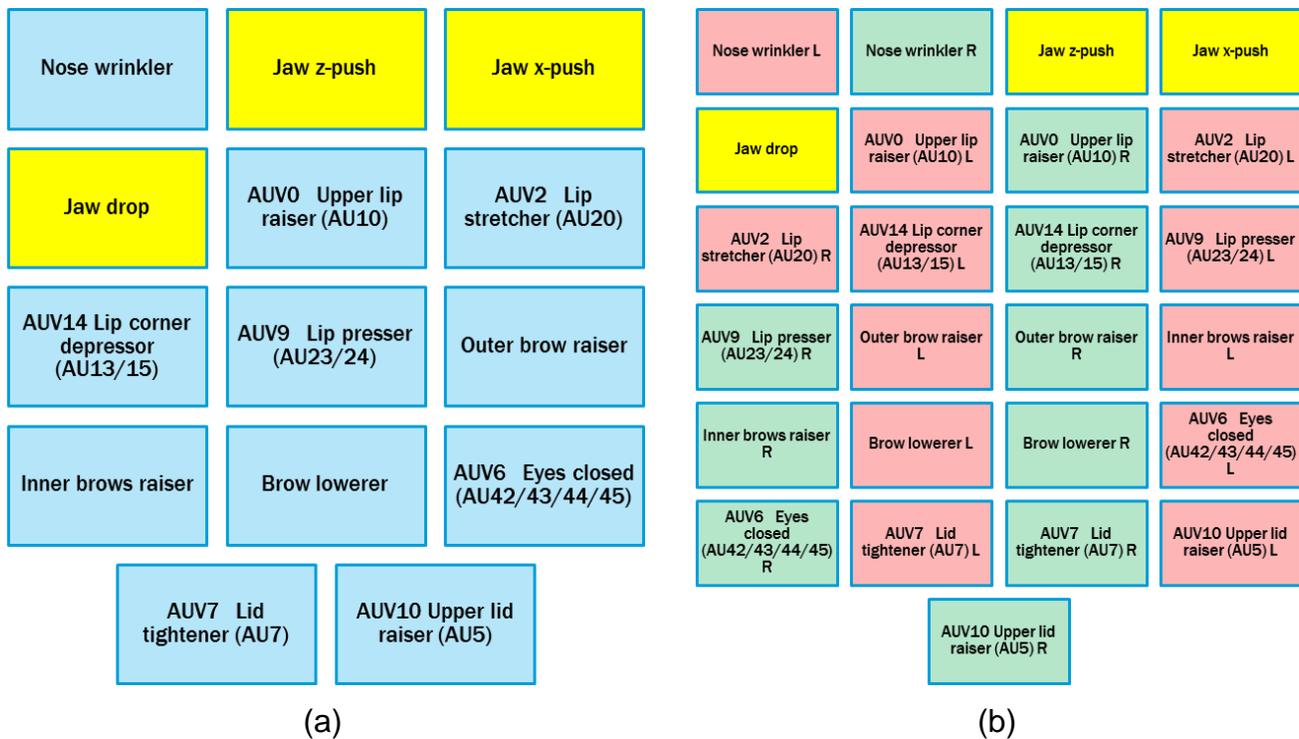


Figura 14: AUV originales y propuestas. (a) Nombre de los AUV Candide-3 según Ahbelrg 2001. (b) AUV propuestas después de dividir las originales en AUV izquierda (terminadas en "L") y AUV derecha (terminadas en "R"). Los AUV Jaw z-push, Jaw x-push y Jaw drop no fueron divididas ya que representan el movimiento del maxilar inferior.

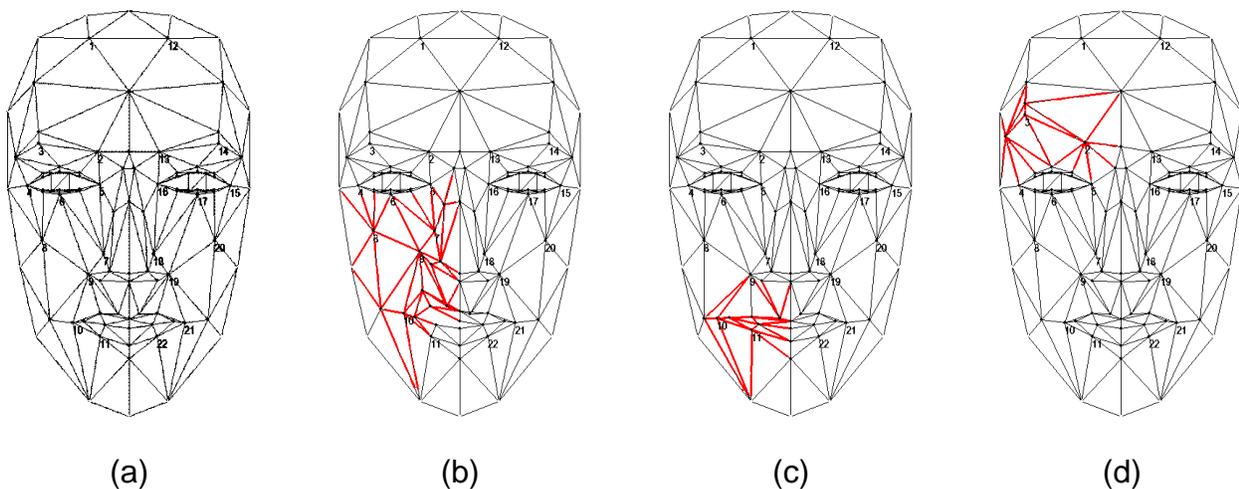


Figura 15: Máscara Candide-3 asimétrica con nivel de activación en 1.0 para algunos AUV específico del lado derecho del paciente. (a) Máscara con todos los AUV en nivel de activación 0. (b) *Nose wrinkler R* (con nivel 1 de activación). (c) *AUV2 Lip stretcher R*. (d) *Outer brow raiser R*.

Porción del rostro	AUV	Grado 1 neuropraxia	Grado 2 axonotmesis	Grado 3 axonotmesis	Grado 4 neurotmesis	Grado 5 Corte parcial	Grado 6 Corte total
Frente	Brow lowerer	Simétrico	Simétrico	Movimiento leve	Sin movimiento	Sin movimiento	Sin movimiento
	Inner brows raiser						
	Outer brow raiser						
Ojos	AUV6 Eyes closed	Simétrico	Simétrico	Levemente asimétrico	Asimétrico	Asimétrico	Asimétrico
	AUV7 Lid tightener						
	AUV10 Upper lid raiser						
Boca	AUV14 Lip corner depressor	Simétrico	Levemente asimétrico	Notoriamente asimétrico	Notoriamente asimétrico	Movimiento leve	Sin movimiento
	AUV0 Upper lip raiser						
	AUV2 Lip stretcher						
	AUV9 Lip presser						

Tabla 2: Comparación subjetiva del nivel de activación de los AUV observados según el grado de parálisis diagnosticado. El criterio de simétrico, levemente asimétrico, y asimétrico es subjetivo. Es necesario crear una medida objetiva para cuantificar el nivel de simetría/asimetría (elaboración propia).

4.2.2 Tracking facial

Se realizó una nueva implementación de software basada en el trabajo de Terissi et al. 2010, utilizando la librería OpenCV para el análisis de imágenes. A diferencia de aquel trabajo, el método propuesto utiliza la máscara Candide-3 asimétrica descrita en la sección anterior en vez de un modelo Candide-3 simplificado. Las AU globales de posición y pose están calculadas en milímetros y radianes respectivamente. La máscara Candide-3 asimétrica es escalada al tamaño del rostro del paciente filmado. El reporte de los puntos seguidos es entregado en puntos de la imagen mientras que el reporte del software original es sobre la base de un modelo virtual en una escala arbitraria.

El método propuesto utiliza un key frame para dar estabilidad a la implementación, pero se han presentado problemas de estabilidad al movimiento y sensibilidad a los cambios de iluminación, y errores en los puntos que están en el borde de la máscara y alrededor de la boca, debido al que el key frame representa la máscara texturizada sobre un fondo negro (en el software original se representa sobre un fondo gris). Para reducir el error del FO en este escenario, la máscara ajustada al key frame es superpuesta al fondo que corresponde a la imagen original del key frame. Además, la imagen es separada en sus tres canales de color (rojo, verde y azul), a cada canal se aplica un filtro SOBEL y se selecciona el valor máximo que resulta en cada canal. La Figura 16 muestra el resultado para el sujeto 1 de la aplicación de este filtro. El Anexo B muestra un ejemplo de error de FO.



Figura 16: Filtro SOBEL aplicado a cada canal, preservando el valor máximo por canal (imagen extraída de Chan 2016, procesada por el método propuesto).

El ajuste de las AU globales y de los AUV es realizado usando el algoritmo de Levenberg-Marquardt implementado en OpenCV 3.0.0, a diferencia de la implementación original que utiliza una librería externa.

Ambos softwares utilizan un key frame para dar estabilidad a la implementación. En el software original, se da un peso fijo de 0,3 al error del FO basado en el key frame y de 0,7 al FO entre fotogramas, es decir, se obtiene un ajuste de la expresión facial con dos flujos ópticos de entrada. En el método propuesto, el FO del key frame es utilizado para determinar una expresión facial y el FO entre imágenes es utilizado para determinar otra expresión facial. Posteriormente se determina el error de la proyección de la máscara Candide-3 asimétrica deformada por cada una de las expresiones faciales determinadas y se seleccionan los AUV de cada una que representan la menor diferencia para los puntos en que afectan. La Figura 17 muestra la diferencia entre el algoritmo de determinación de pose de Terissi et al 2010 y el método propuesto. La Figura 18 muestra la diferencia entre el algoritmo de determinación de expresión facial original y el propuesto.

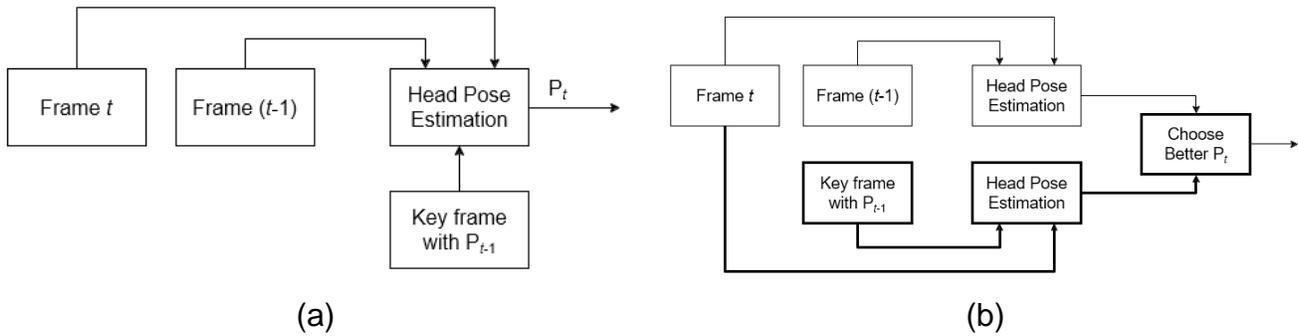


Figura 17: Comparación entre los algoritmos para determinar la pose. (a) Algoritmo de Terissi et al. 2010 para determinar pose, utiliza como entrada el frame actual (Frame t), el frame anterior (Frame $(t-1)$) y el key frame anterior (Key frame with P_{t-1}), calcula el FO entre los dos frames consecutivos y el key frame anterior y el frame actual. Luego, calcula el vector de pose usando ambos flujos ópticos. El resultado es el vector de pose P_t . (b) Método propuesto para determinar la pose (se han dibujado en líneas gruesas los bloques que difieren del algoritmo de Terissi et al. 2010) utiliza como entrada el frame actual (Frame t) y el frame anterior (Frame $(t-1)$), calcula el FO entre ambos frames y produce como resultado un vector de pose P'_t . Luego, usando como entrada el frame actual (Frame t) y el key frame anterior (Key frame with P_{t-1}), calcula el FO entre ambos frames y produce como resultado un vector de pose P''_t . Finalmente se elige el vector P'_t o P''_t que mejor coincide con el FO.

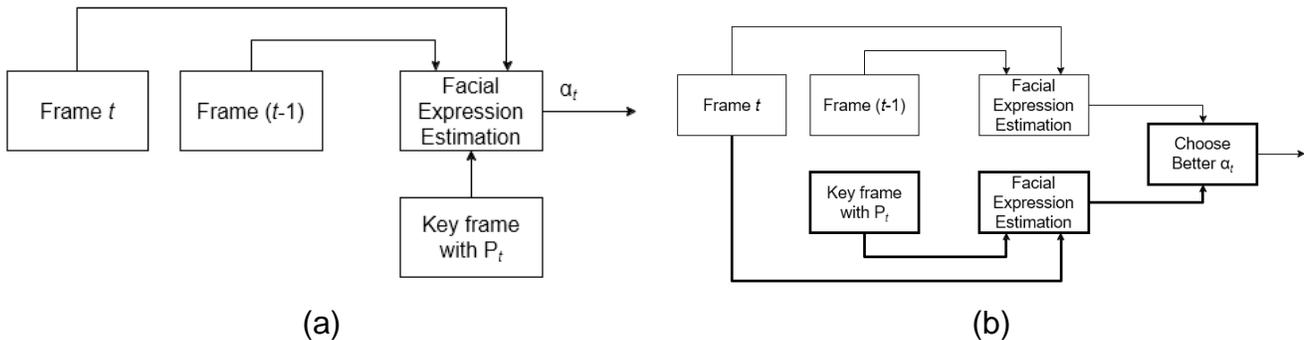


Figura 18: Comparación entre los algoritmos para determinar la expresión facial. (a) Algoritmo de Terissi et al. 2010 para determinar la expresión facial, utiliza como entrada el frame actual (Frame t), el frame anterior (Frame $(t-1)$) y el key frame actualizado (Key frame with P_t). Calcula el FO entre ambos frames. Luego, calcula el vector de niveles de activación de los AUV, α_t , usando ambos flujos. (b) Método propuesto para determinar la expresión facial (se han dibujado en líneas gruesas los bloques que difieren del algoritmo de Terissi et al. 2010) utiliza como entrada el frame actual (Frame t) y el frame anterior (Frame $(t-1)$), produce como resultado un vector α'_t , luego, el frame actual (Frame t) y el key frame actualizado (Key frame with P_t) produce como resultado un vector de pose α''_t . Finalmente se elige el vector α'_t o α''_t que mejor coincide con el FO.

4.2.3 Mediciones para la cuantificación

La cuantificación del movimiento utilizando los parámetros de activación de los AUV de la máscara Candide-3 asimétrica es diferente entre el lado izquierdo y derecho tanto en personas sanas como afectadas por la parálisis facial parcial. Pero esta diferencia de movimiento es menor en las personas sanas cuando realizan ejercicios simétricos. Si el nivel de activación de los AUV del lado derecho e izquierdo se grafican en un plano cartesiano, donde la abscisa representa el nivel de activación del lado derecho y la ordenada representa el nivel de activación del lado izquierdo, el gráfico en una persona sana debería concentrar los puntos sobre la diagonal.

Por diseño, la máscara Candide-3 original y asimétrica interpreta los niveles de activación de los AUV como un factor que multiplica a un vector de deformación. Debido a esto, el nivel máximo de activación de un AUV depende de la fisionomía de la persona que participa en alguna filmación analizada. Esto implica que el nivel de activación de los AUV no es una medida de comparación apropiada entre sujetos.

Algunos pacientes presentan una parálisis flácida, es decir, la musculatura del lado paralizado está totalmente relajada. Otros presentan una parálisis hipertónica, al contrario de la anterior, la musculatura del lado paralizado está tensa. Esto significa que en algunos pacientes la máscara Candide-3 asimétrica debe ser ajustada a una expresión facial inicial, y tal vez esta no cambie en el tiempo (en la Figura 8 el sujeto 2 (b) presenta una parálisis hipertónica, mientras que los sujetos 5 (e) y 6 (f) presentan parálisis flácida).

Para superar este problema y considerar los momentos en las filmaciones en que las expresiones faciales cambian, se propone calcular la velocidad del nivel de activación de las AUV entre cada frame. Esto reduce la sensibilidad a las expresiones faciales fijas. Suponiendo que el lado paralizado se mueve menos que el lado conservado, habrá una diferencia de velocidad destacada entre ambos lados. Las velocidades se comparan entre sí en un gráfico derecha versus izquierda donde se analizan los siguientes AUV:

- *AUV6 Eyes closed*
- *AUV7 Lid tightener*
- *AUV10 Upper lid raiser*
- *Nose wrinkler*
- *AUV0 Upper lip raiser*
- *AUV2 Lip stretcher*

- *AUV14 Lip corner depressor*
- *Outer brow raiser*
- *Inner brows raiser*

Los movimientos de mandíbula no son considerados en la evaluación del grado de parálisis, por lo tanto, los niveles de activación de los AUV *Jaw z-push*, *Jaw x-push* y *Jaw drop* son descartados.

Los movimientos de ambos lados del rostro, deberían ser coherentes en todo momento en una persona sana, por lo que se calcula el ángulo promedio respecto del eje diagonal de los gráficos. Esto significa que si una persona puede mover ambos lados del rostro al mismo tiempo, los niveles de activación de los AUV deben quedar graficados en la diagonal, por lo que el ángulo de desviación respecto de la diagonal debiese ser cero.

Para determinar el ángulo respecto de la diagonal, se utiliza la definición del producto cruz entre dos vectores definido por $a \times b = (|a||b| \sin \alpha)\hat{n}$, que lo describe en función del ángulo α entre los dos vectores. Calculando la norma del producto cruz se puede despejar el ángulo α . Igualando este desarrollo con la magnitud del producto cruz en función de las componentes de cada vector $|a \times b| = \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix}$ y considerando el ángulo a como un ángulo unitario que corresponde a la diagonal, y $b = (r', l')$ como el vector de la velocidad del nivel de activación de un AUV analizado, se obtiene una fórmula para el seno del ángulo α . La Figura 19 presenta tres velocidades de activación para tres sujetos hipotéticos a los que se ha medido la velocidad del nivel de activación de un mismo AUV cualquiera.

El ángulo de desviación recién descrito se puede relacionar con el grado de parálisis en escala HB. Básicamente, entre mayor es la desviación significa que el movimiento de un lado del rostro respecto del otro es más diferente. Sin embargo, existe una limitación importante con este enfoque: en diferentes ejercicios, el ángulo de desviación es diferente también dependiendo del ejercicio. Además, durante la filmación, un sujeto realiza varios gestos voluntarios e involuntarios, más allá de los ejercicios requeridos para la graduación usando la escala HB. Esto se traduce en una serie de expresiones faciales que no guardan relación con la graduación. Las filmaciones tienen diferente duración. Por otro lado, en reposo, el tracking basado en FO detecta pequeñas variaciones de iluminación que pueden afectar las mediciones.

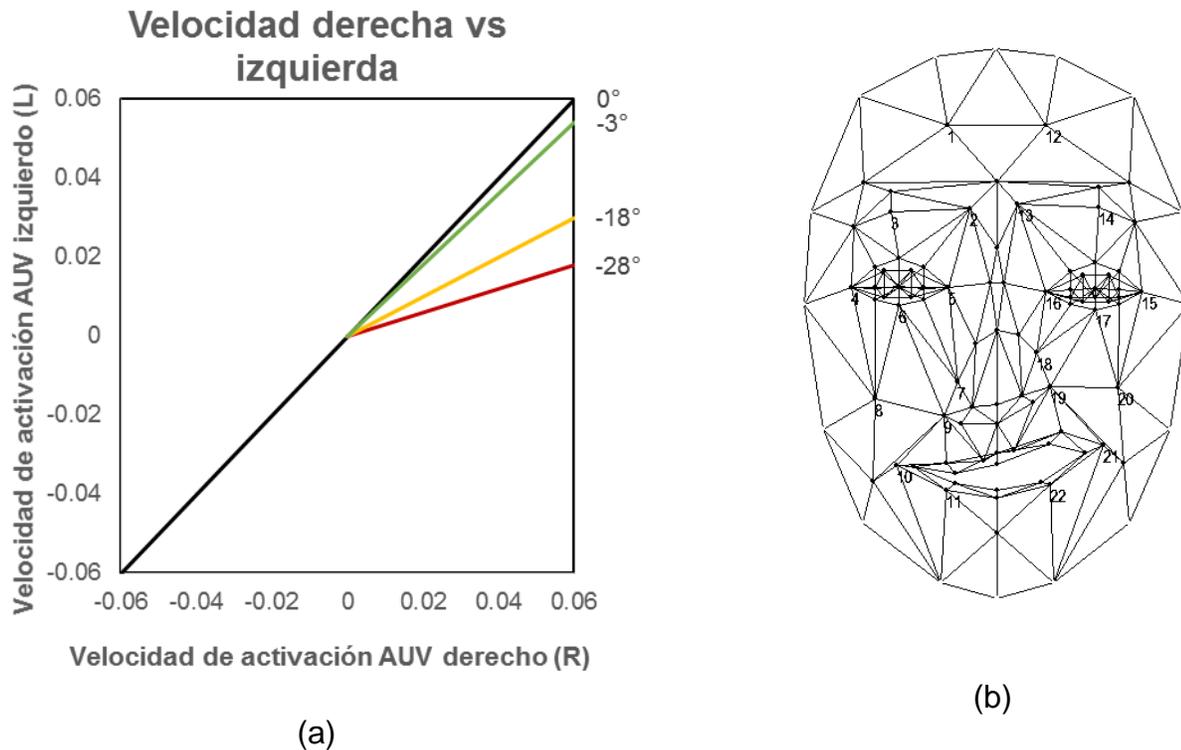


Figura 19: Comparación del ángulo de la velocidad del nivel de activación entre tres sujetos para un AUV cualquiera. (a) Un sujeto sano o con una parálisis muy leve, presentará una velocidad de activación cercana a la diagonal, como sería la línea verde con ángulo -3° (negativo es bajo la diagonal). Un sujeto con una diferencia de velocidad mayor, presentará una velocidad de activación más parecida a la línea amarilla con ángulo -18° para el mismo AUV. Finalmente, un sujeto con una parálisis de grado elevado, presentará una velocidad de activación parecida a la línea roja con ángulo -28° para el mismo AUV. (b) Ejemplo de parámetros asimétricos para una configuración específica de un rostro, bajo el supuesto que el frame anterior a esta configuración está en reposo, la expresión facial corresponde a la línea amarilla de -18° (gráfico y figura elaboración propia).

Para superar estos problemas, se proponen las siguientes restricciones:

- i) Se incluye un punto de corte para la velocidad, tal que si la magnitud de la velocidad es menor o igual a 0,01 (1/frame), esta no es considerada en las mediciones.
- ii) En vez de medir el ángulo de desviación respecto del eje central, se calcula la frecuencia acumulada de los ángulos en valor absoluto en intervalos discretos de 1° a partir de 0° hasta llegar a los 90° . Así un sujeto sano presentaría una

frecuencia acumulada alta en torno al ángulo 0° , mientras que un sujeto con parálisis facial parcial, presentaría una acumulación en torno a un ángulo superior.

La distribución puede variar entre los AUV activados, por ejemplo, si el sujeto sonríe mucho y levanta pocas veces las cejas, habrá más frecuencia de los ángulos para los AUV asociados a una sonrisa que para el movimiento de cejas. Por esto se grafica la frecuencia acumulada dividida por la suma de todas las frecuencias por cada AUV.

Al graficar la distribución de probabilidad se obtiene una curva que se acerca al valor 1 muy cerca del ángulo 0° para un sujeto sano, mientras que para un sujeto con parálisis facial parcial se acerca en algún ángulo superior (ver Figura 20).

Para comparar las mediciones entre curvas (entre frecuencias acumuladas de los ángulos de velocidad y entre sujetos), se considera el área bajo la curva normalizada (en adelante, AUC). Este valor para un AUV en un paciente sano será un número cercano a 1. En AUV que presentan alguna asimetría, será un valor mejor. En caso de que no haya movimiento en absoluto, será un valor cercano a cero. Finalmente, para determinar la graduación en escala HB, se deben definir rangos de valores para las AUC que describan el nivel de parálisis facial. Recordando que HB requiere ciertos ejercicios para diferenciar los grados de parálisis I, II y III versus IV, V y VI, la decisión inicial es comparar el movimiento de los ojos o las cejas. En caso de que el AUC de del movimiento de los ojos o las cejas sea inferior a un umbral, el grado de parálisis es IV, V o VI. De lo contrario, es I, II o III. La escala de HB requiere comparar luego la zona de la boca para terminar la graduación. Si el paciente tiene grado de parálisis I, II o III, entonces el AUC del movimiento de la boca es a penas cercano a 1 en el grado I y va disminuyendo hacia el grado III. Por otro lado, si el paciente tiene grado de parálisis IV, V o VI, el movimiento de la boca será dispar en todo momento por lo que el AUC podría estar cercado a 0,5 o inferior. Los valores de corte para tomar estas decisiones son calculados con la filmación de los pacientes graduados descrita en el capítulo anterior.

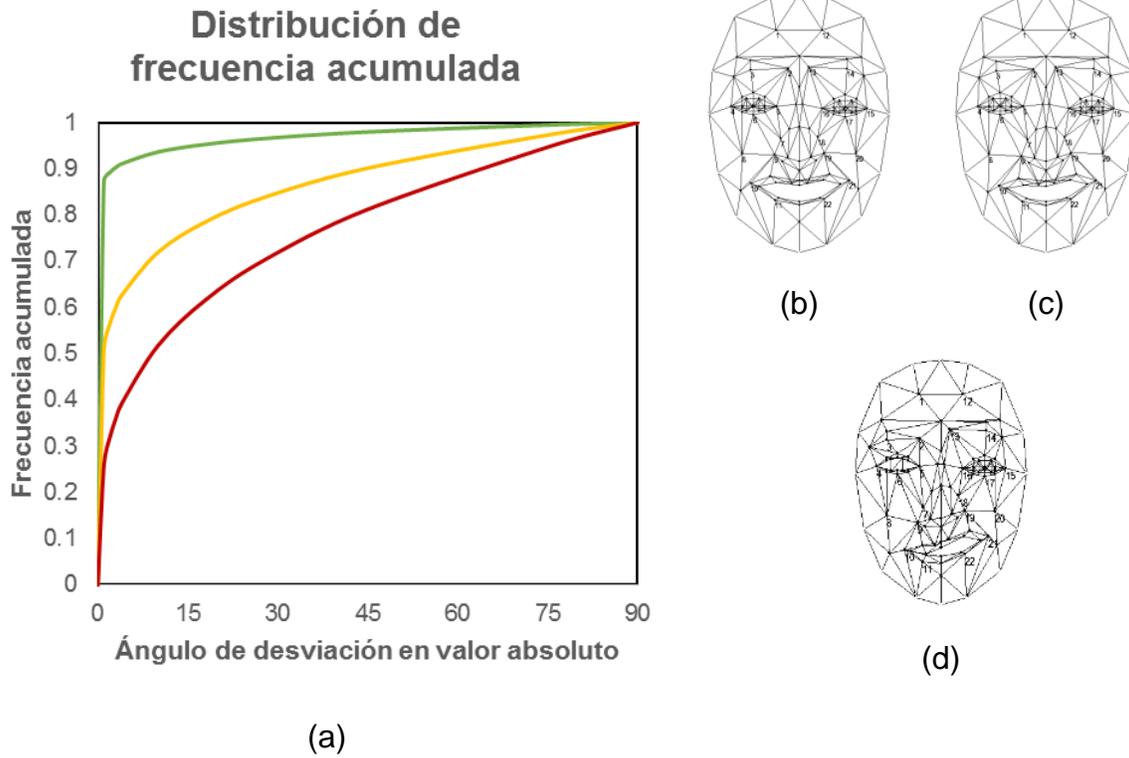


Figura 20: Esquema de la distribución de las frecuencias acumuladas para la velocidad del ángulo respecto de la diagonal de un AUV específico en diferentes sujetos. El gráfico (a) presenta en la abscisa, los ángulos discretos de 0° a 90° . En la ordenada, la frecuencia acumulada, con máximo en 1. La línea verde representa a un sujeto sano (b). La línea amarilla a un sujeto con un grado bajo de parálisis facial parcial (c). La línea roja, un sujeto con un grado mayor de parálisis facial parcial (d) (gráficos y figuras elaboración propia).

4.2.4 Algoritmo de graduación de la parálisis facial

En principio, de acuerdo a lo descrito en la introducción de este documento, la graduación en escala HB compara el lado paralizado con el lado conservado, independiente de cuál de los dos lados es el que está paralizado o conservado. Considerando la habilidad de mover ambos lados del rostro. La capacidad máxima de movimiento depende de en gran medida de las características fisiológicas de los pacientes. De acuerdo a (Aik Kah et al. 2011), es posible considerar la asimetría del movimiento de la frente como un indicador de la separación de niveles 1, 2 y 3 de los niveles 4, 5 y 6 de la escala HB. Esto se traduce en que la primera decisión es “si existe ausencia de movimiento de las cejas, entonces el paciente tiene graduación 4, 5 o 6” (de lo contrario, su graduación es en niveles 1, 2 o 3).

No fue posible realizar el seguimiento de las AUV correspondientes al movimiento ocular, por lo que se analizan las AUV siguientes:

- *Nose wrinkler (nw)*
- *AUV0 Upper lip raiser (ulr)*
- *AUV2 Lip stretcher (ls)*
- *AUV14 Lip corner depressor (cd)*
- *Outer brow raiser (obr)*
- *Inner brows raiser (ibr)*
- *AUV Brow Loweler (bl)*

En algunos casos, la parálisis hipertónica deja fijo un nivel de activación en un valor diferente de cero. La Figura 21 muestra el nivel de activación del AUV correspondiente a “bajar el borde del labio” (cd) para algunos casos representativos. El sujeto con graduación 2 (Figura 21b) presenta una diferencia aparentemente constante del nivel de activación debido a su parálisis hipertónica. Esto queda resuelto al ocupar la velocidad del nivel de activación (Figura 22). La comparación por histograma de ángulos de desviación descrita en 4.2.3 es insensible a esta situación. Se ha construido un histograma de la probabilidad de que el ángulo α sea inferior a un valor, considerando aquellas velocidades cuya magnitud vectorial sea superior a 0.01 (Figura 23).

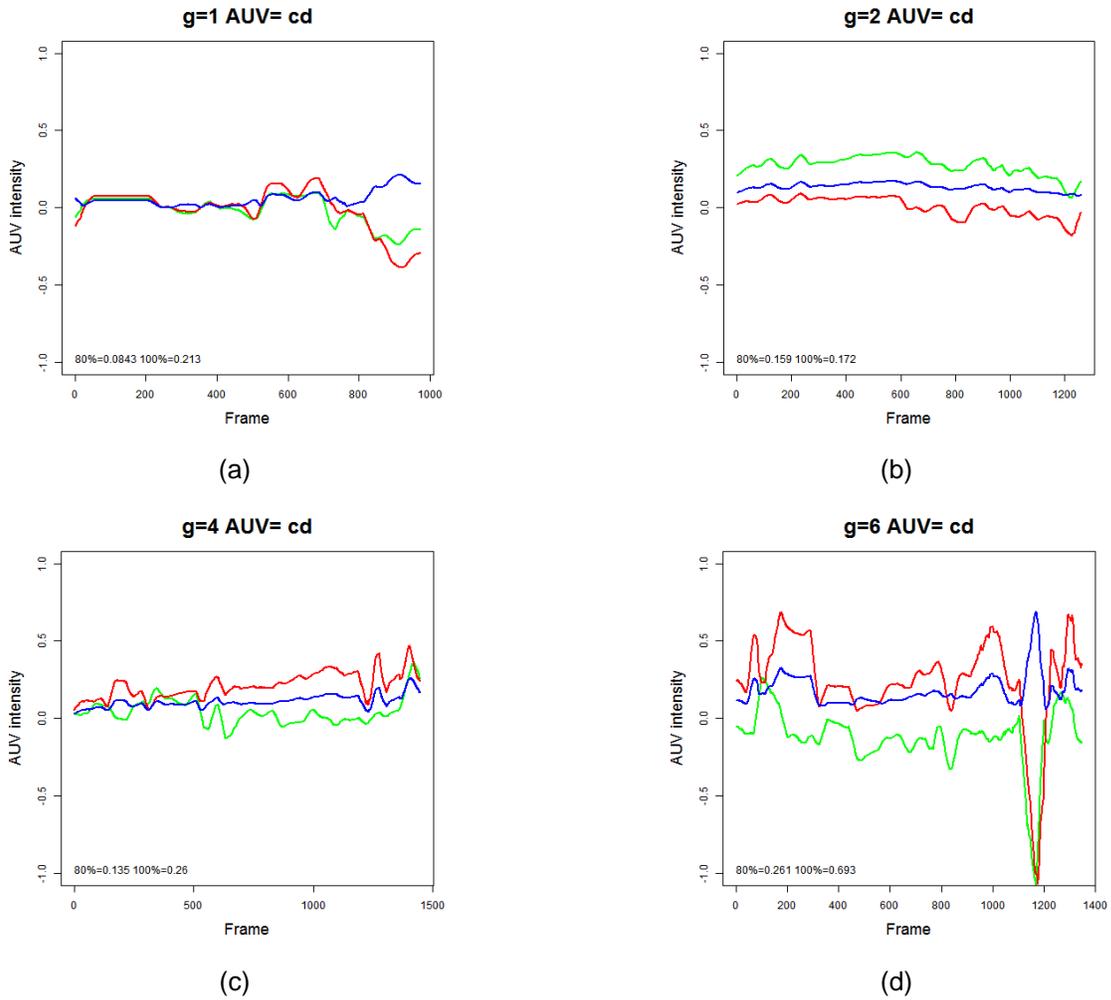


Figura 21: Niveles de activación del *AUV14 Lip corner depressor*. En rojo el lado derecho, verde el lado izquierdo. En azul, el valor absoluto de la diferencia de ambos niveles. (a) Sujeto con graduación 1, (b) sujeto con graduación 2, (c) sujeto con graduación 4, (d) sujeto con graduación 6. Todos los gráficos están a la misma escala vertical (de -1,0 a 1,0) (elaboración propia).

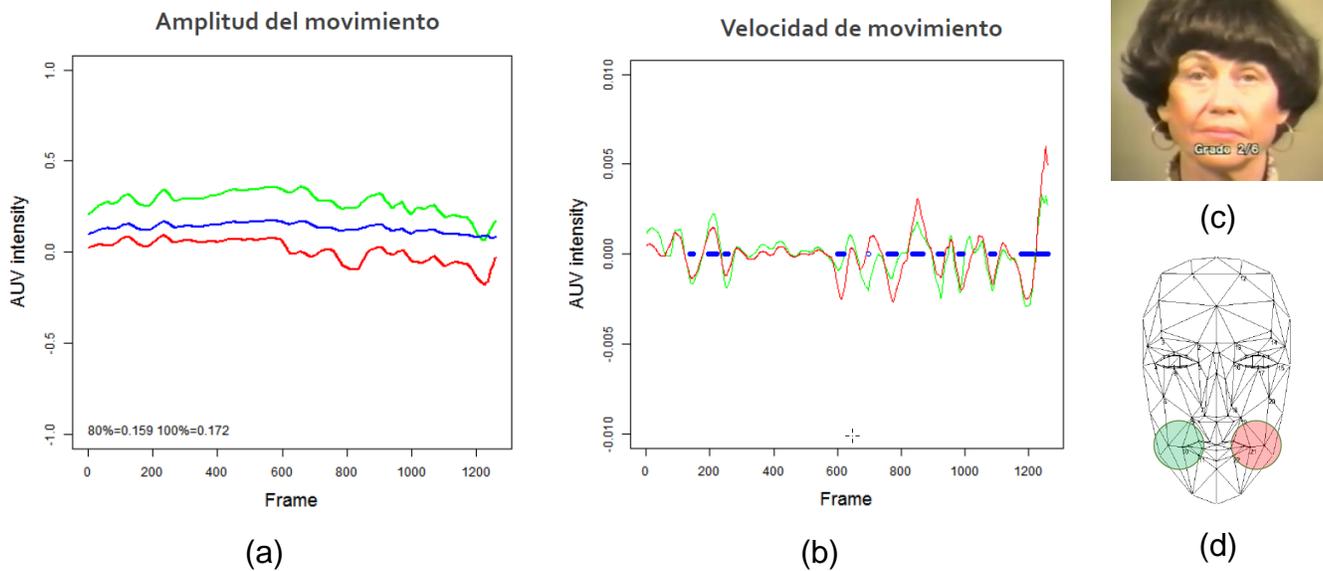


Figura 22: Conversión del intensidad de AUV a velocidad. La conversión de la intensidad del nivel de activación del AUV a velocidad de activación es la diferencia entre el nivel de activación de un AUV en un frame y en el frame anterior. (a) Es el nivel de activación del *AUV14 Lip corner depressor*. En color rojo el lado derecho y en verde el lado izquierdo. El color azul representa la diferencia entre ambos niveles izquierdo y derecho en valor absoluto. (b) Corresponde a la velocidad de activación del mismo AUV. En color rojo la velocidad de activación del lado derecho y en verde la del lado izquierdo. En azul, los movimientos con amplitud mayor a 0,01 como se describió en la sección anterior. (c) Imagen de referencia del sujeto analizado, descrita en 3.1.2. (d) A modo de referencia, los círculos representan la zona que afecta este AUV a cada lado de la boca del sujeto. El lado derecho en color rojo y el lado izquierdo en color verde.

Las AUC son más cercanas a 1 entre mayor simetría del movimiento presenten el paciente, es decir, una movilidad conservada (Figura 23a). Un sujeto con graduación 4 presenta AUC más cercanas a 0,5 (Figura 23b). Un sujeto con graduación 6 presenta AUC muy diferentes (Figura 23c). Observando la citada figura, se puede notar que el comportamiento del AUV *Brow Loweler* (bl) es muy diferente del resto de los niveles de activación (línea roja en los tres gráficos). Es muy similar en el sujeto con graduación 1 y 4, mientras que el resto de los AUV se comportan de manera diferente en cada caso.

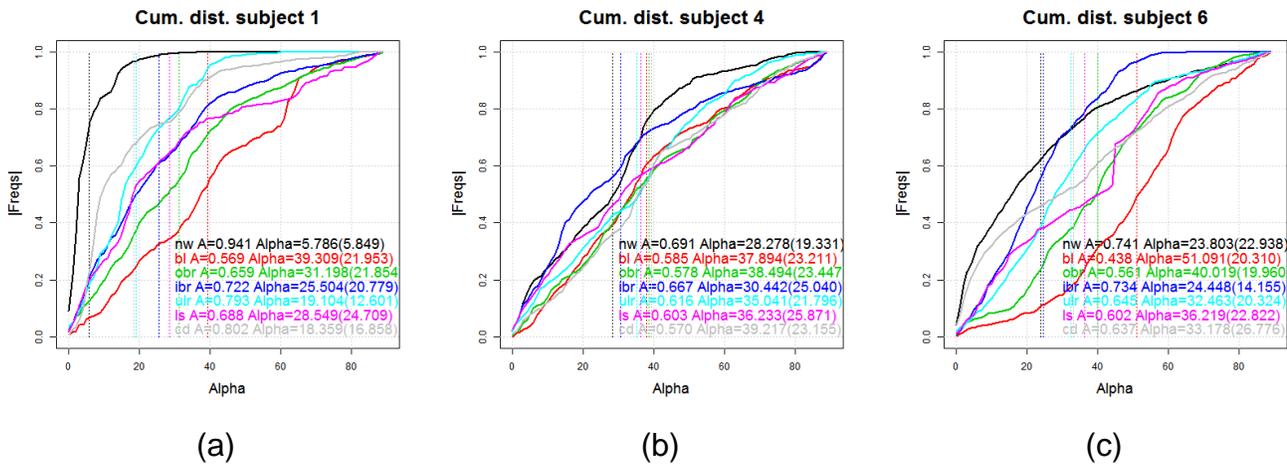


Figura 23: Distribución de probabilidad acumulada del ángulo de desviación respecto de la diagonal (α). Un sujeto con graduación 1 (a) presenta un movimiento muy simétrico del movimiento de nariz (línea negra). Un sujeto con graduación 4 (b) presenta una distribución homogénea de sus movimientos faciales. Un sujeto con graduación 6 (c) presenta una distribución más dispersa. Destacado en línea punteada está el ángulo promedio por cada nivel de activación de AUV. En la leyenda, se indica el área bajo la curva (“A”), α promedio (“Alpha”) y su desviación estándar (entre paréntesis). (Elaboración propia).

Las AUC para todos los sujetos están disponibles en la Tabla 3. Se han agrupado los resultados de arriba abajo en AUV relacionados con nariz, frente y boca respectivamente. Numéricamente, los valores de AUC para los AUV nw, obr y cd son más altos en los grados 1 a 3 que en los grados 4 a 6, mientras que el ángulo de desviación α no tiene un comportamiento coherente con la graduación salvo en el AUV nw. La desviación estándar de α tampoco evidencia alguna graduación.

Grado	Nose wrinkler (nw)		
	A.U.C.	Prom.(α)	D.S.(α)
1	0.94	5.79	5.85
2	0.91	8.22	5.47
3	0.85	13.61	14.24
4	0.69	28.28	19.33
5	0.71	26.41	23.16
6	0.74	23.80	22.94

Grado	Brow lowerer			Outer brow raiser (obr)			Inner brow raiser (ibr)		
	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)
1	0.57	39.31	21.95	0.66	31.20	21.85	0.72	25.50	20.78
2	0.64	32.68	15.28	0.71	26.62	14.88	0.87	11.78	7.83
3	0.41	53.52	21.55	0.71	26.27	22.07	0.70	27.50	22.26
4	0.58	37.89	23.21	0.58	38.49	23.45	0.67	30.44	25.04
5	0.33	60.74	12.92	0.55	40.64	17.95	0.62	34.36	26.33
6	0.44	51.09	20.31	0.56	40.02	19.96	0.73	24.45	14.16

Grado	AUV0 Upper lip raiser (ulr)			AUV2 Lip stretcher (ls)			AUV14 Lip corner depressor (cd)		
	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)
1	0.79	19.10	12.60	0.69	28.55	24.71	0.80	18.36	16.86
2	0.78	20.40	16.25	0.60	36.63	24.25	0.86	12.92	8.81
3	0.57	39.45	28.12	0.68	29.41	23.16	0.80	18.29	14.81
4	0.62	35.04	21.80	0.60	36.23	25.87	0.57	39.22	23.16
5	0.76	22.34	17.33	0.62	34.77	21.21	0.80	18.37	16.32
6	0.64	32.46	20.32	0.60	36.22	22.82	0.64	33.18	26.78

Tabla 3: Área bajo la curva (AUC) de la distribución de probabilidad del ángulo (α). Desviación respecto de la diagonal promedio de α (en grados sexagesimales) y su desviación estándar, para cada sujeto con graduación 1 a 6 en escala HB. Para facilitar la lectura, la tabla ha sido dividida en tres grupos de AUV, de arriba abajo, nariz, frente y boca respectivamente (elaboración propia).

Si se promedian las AUC de los niveles de activación de los AUV correspondientes a la frente (ibr, obr en la Tabla 3) y los correspondientes a la boca (ls, ulr, cd), se puede hacer una comparación sencilla para determinar el grado de parálisis facial en escala HB. El AUV *Brow Lowerer* (bl) es muy diferente del resto de los niveles de activación, por lo que no es considerado en el cálculo del promedio del movimiento de la frente (Tabla 4). Se puede presumir en general que un AUC superior a 0,65 es un punto de corte para decir que existe simetría. Como alternativa, se puede estimar un área menor (un peor caso), si se resta del

AUC promedio general su desviación estándar, se obtiene un punto de corte de 0.60. Esto queda representado en la última columna de la Tabla 4.

Grado	AUV ls, ulr, cd		AUV ibr, obr		Todos los AUV		General - SD
	Boca	DS(boca)	Frente	DS(frente)	General	DS(general)	
1	0.76	0.06	0.69	0.04	0.73	0.06	0.67
2	0.75	0.13	0.79	0.12	0.76	0.11	0.65
3	0.68	0.12	0.71	0.01	0.69	0.08	0.61
4	0.60	0.02	0.62	0.06	0.61	0.04	0.57
5	0.73	0.10	0.59	0.05	0.67	0.10	0.57
6	0.63	0.02	0.65	0.12	0.64	0.06	0.57

Tabla 4: Promedio de las AUC de los niveles de activación de los AUV de la frente (ibr, obr) y de la boca (ls, ulr, cd). La tabla incluye también sus desviaciones estándar para los sujetos graduados (grados 1 a 6). También se incluye el promedio de activación general de los AUV de la boca y frente (ibr, obr, ls, ulr y cd). La columna “General - SD” muestra la resta del promedio general y su desviación estándar (elaboración propia).

En cualquiera de los dos criterios, se puede determinar fácilmente si el paciente tiene grados de parálisis 1, 2 o 3, o grados 4, 5 o 6. De acuerdo a cómo se describe la graduación de HB en (Aik Kah et al. 2011), se propone un procedimiento de decisión descrito brevemente en la Tabla 5. La Figura 24 es una propuesta para un algoritmo simplificado de graduación de parálisis del nervio facial de House-Brackmann (HB).

Porción del rostro	AUV	Grado 1 neuropraxia	Grado 2 axonotmesis	Grado 3 axonotmesis	Grado 4 neurotmesis	Grado 5 Corte parcial	Grado 6 Corte total	
Frente	Brow lowerer	AUC promedio frente > 0.65			AUC promedio general < 0.65 y desv estándar < 0.05	AUC promedio frente < 0.65		
	Inner brows raiser							
	Outer brow raiser							
Ojos	AUV6 Eyes closed	---Sin datos---				AUC promedio general < 0.65 y desv estándar < 0.05	--Sin datos--	
	AUV7 Lid tightener							
	AUV10 Upper lid raiser							
Boca	AUV14 Lip corner depressor	AUC promedio boca > 0.75	AUC promedio bocca entre 0.70 y 0.75	AUC promedio boca < 0.70	AUC promedio general < 0.65 y desv estándar < 0.05		AUC promedio boca > 0.65	AUC promedio boca < 0.65
	AUV0 Upper lip raiser							
	AUV2 Lip stretcher							
	AUV9 Lip presser							

Tabla 5: Determinación del grado de parálisis en escala HB de acuerdo al área bajo la curva (AUC) de la distribución de probabilidades de los niveles de activación de los AUV. Al carecer de información sobre el movimiento ocular, se propone usar el AUC del movimiento de la frente como referencia (promedio de las AUC ibr y obr), con un umbral bajo 0,65 se considera que es parálisis de grados 1, 2 o 3 y sobre este valor, 4, 5 o 6. Luego para determinar el grado de parálisis en el primer grupo (1, 2 o 3), se compara el AUC promedio de la boca (promedio de las AUC ls, ulr, cd), sobre el umbral 0,75 es grado 1, entre 0,70 y 0,75 es grado 2 y bajo 0,70 es grado 3. Para distinguir los grados 4, 5 y 6, se observa que el promedio general de las AUC es menor a 0,65 con una desviación estándar bajo 0,05. Si esto no se cumple, se compara el AUC promedio de la boca con 0,65. Si es superior, es grado 5 y si es inferior es grado 6 (elaboración propia).

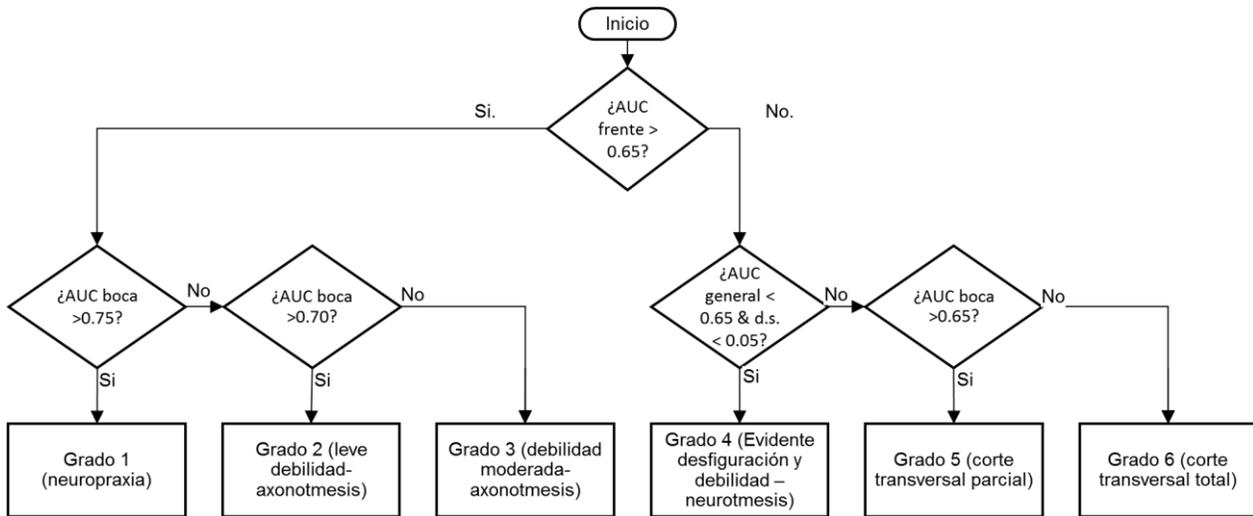


Figura 24: Algoritmo propuesto de graduación de parálisis del nervio facial de House-Brackmann (HB). Se consideran los valores promedio de las AUC de la distribución de probabilidades de los niveles de activación de los AUV.

4.3 Validación del método propuesto

A continuación se presentan los resultados de la validación del método propuesto. Comenzando por la comparación con el sistema de seguimiento cinemático con marcadores pasivos, los hallazgos en las filmaciones de sujetos con graduación y la evaluación de pacientes sin graduación.

4.3.1 Comparación con el sistema de seguimiento cinemático con marcadores pasivos

El método propuesto fue comparado con una grabación en vídeo de una persona haciendo gestos con su rostro, en particular un gesto de sorpresa. La Figura 25 muestra el desplazamiento según el marcador pasivo en la curva color azul. Debido al ajuste que se realiza para igualar el sistema de coordenadas del software y del marcador pasivo, no es posible comparar exactamente si las coordenadas son iguales, pero si el desplazamiento total. En los aproximadamente cuatro segundos que se visualizan, la diferencia promedio entre las coordenadas Y de ambos marcadores es de 1,03mm (desviación estándar de 0,91mm). La distancia total recorrida según el marcador pasivo es de 8,90mm, versus el registro del software de 8,67mm. El Anexo C describe el procedimiento de ajuste la información de los marcadores pasivo y del método propuesto.

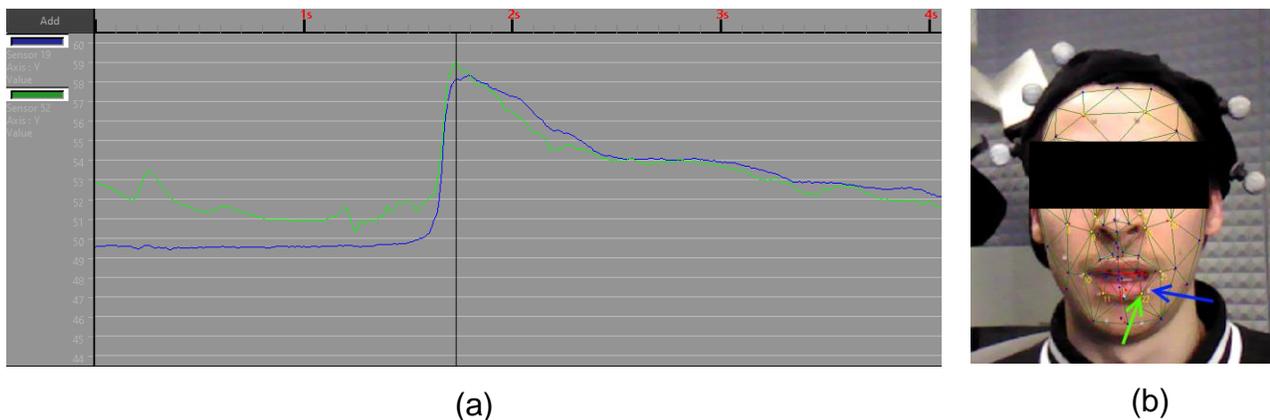


Figura 25: Comparación de la apertura de la boca. (a) Se compara la coordenada Y del punto 19 correspondiente a un marcador pasivo (en color azul) y el punto 22 que sigue el método propuesto (en color verde). La comparación se hizo mediante utilizando VisArtico (Ouni 2012). Los números a la izquierda de la captura de pantalla representan un desplazamiento en mm respecto de un sistema de coordenadas arbitrario. El gráfico representa casi cuatro segundos de grabación. (b) El marcador pasivo 19 está señalado por la flecha azul y el punto 22 del modelo que reporta el método propuesto está señalado por la flecha verde (la filmación fue adquirida gracias al financiamiento del proyecto STIC-AmSud (15STIC04)).

4.3.2 Hallazgos en filmaciones de sujetos con graduación

La máscara Candide 3 asimétrica fue configurada usando un key frame de cada paciente, igualando tanto su postura y características faciales como su expresión facial.

El método propuesto fue ejecutado sobre cada vídeo de manera independiente. Frame a frame, las posiciones de los 22 puntos de interés quedaron capturadas en un archivo de texto junto con el nivel de activación de los AUV. Finalmente estos archivos de texto fueron analizados usando R, en búsqueda de relaciones entre los niveles de activación de los AUV y el grado de parálisis de cada paciente.

El método propuesto limita la detección del nivel de activación al rango de -1,5 a 1,5 de cada AUV, debido a que en las pruebas de desarrollo algunas activaciones extremas producen resultados fuera de rango, como levantar los labios más arriba de la nariz. Si bien a nivel del seguimiento óptico esto es correcto, en términos anatómicos, no lo es. Debido a esto, el análisis en R contempla corregir mediante una interpolación de media móvil los niveles de activación de los AUV que estén en 1,5 o -1,5. La interpolación de media móvil se realiza considerando una ventana de 31 frames centrados en el punto que se desea interpolar (31 frames corresponden aproximadamente a un segundo de vídeo.).

El registro de los niveles de activación es ruidoso, es decir, presenta peaks de manera aleatoria entre frames. El método propuesto guarda una imagen de la superposición de la máscara Candide 3 asimétrica sobre el frame del vídeo original. En este log se pueden apreciar desviaciones en un único frame o en dos frames consecutivos. Estas se deben a cambios en las condiciones de iluminación que afectan levemente el seguimiento mediante FO. Para reducir este ruido, se aplica una media móvil centrada de 31 cuadros a los niveles de activación originales o a los que fueron corregidos con el método de interpolación recién descrito. La Figura 26 muestra un ejemplo de esta aplicación para el *AUV2 Lip stretcher R*. Suponiendo que esta aplicación se utilice en vídeos en vivo, la media móvil debería contemplar algunos frames hacia atrás a partir del frame actual. Esto se traduce en que los valores de los niveles de activación suavizados por la media móvil quedan desplazados más adelante en el tiempo, lo que no representa una diferencia en el método de análisis empleado.

La detección del movimiento de los párpados no fue posible utilizando el método propuesto. Sin embargo, la detección del movimiento de otras partes del rostro, resulta ser a simple vista

bastante precisa pero no exenta de problemas. La Figura 27 muestra varios problemas de tracking en ellos: con el movimiento de párpados, y artefactos de grabación.

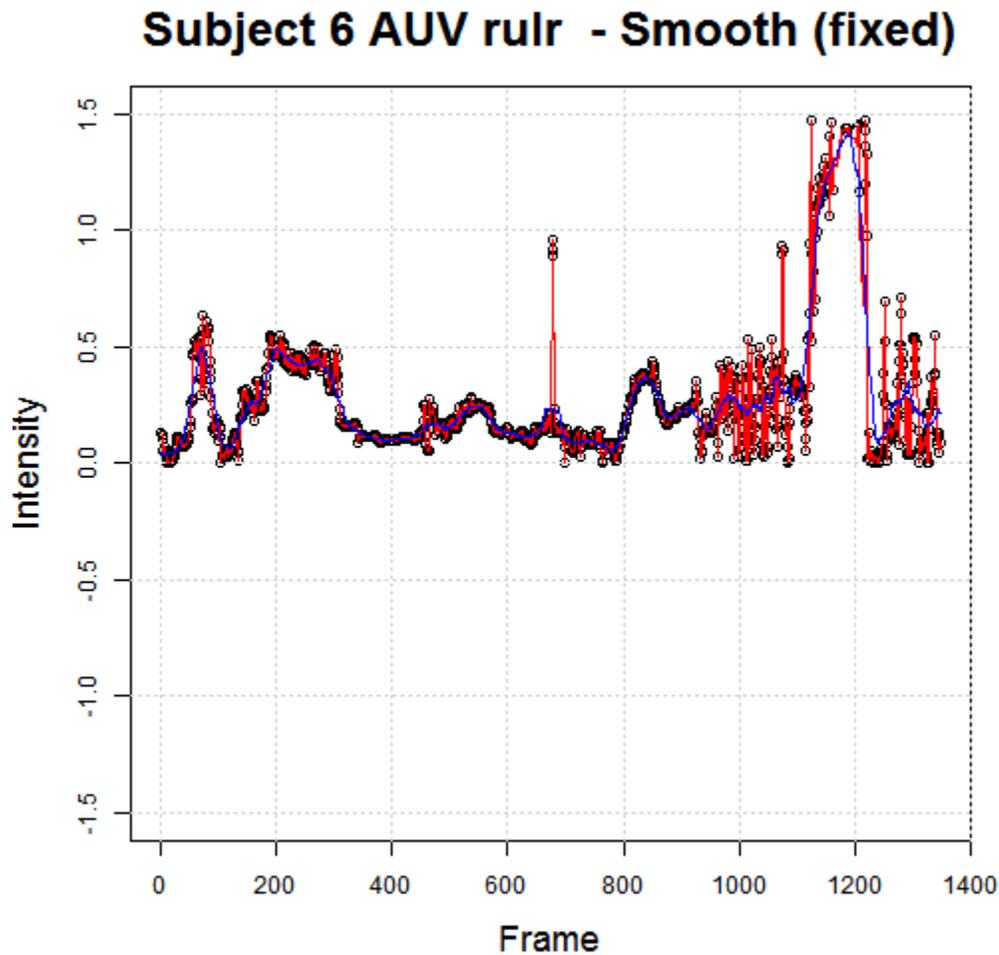
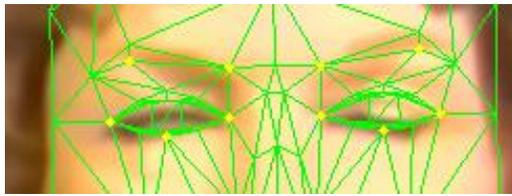
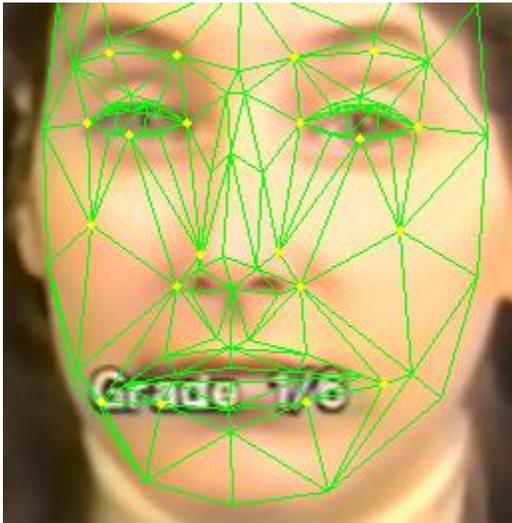


Figura 26: Ejemplo de interpolación y suavizado de los niveles de activación del AUV0 Upper lip raiser R (derecho), del paciente 6. Los puntos negros corresponden a las lecturas efectivas del nivel de activación. La línea roja es la unión de estos puntos después de interpolar los niveles que quedan en 1,5 o -1,5 usando una media móvil de 31 frames centrada sobre el punto por interpolar. Después de corregir la curva, se aplica una media móvil de 31 frames para suavizar.



(a)



(b)



(c)

Figura 27: Diferentes problemas de seguimiento. (a) No es posible detectar el movimiento de los párpados. El paciente 3 tiene los ojos cerrados, pero la máscara Candide 3 asimétrica no se adapta a este gesto. (b) El paciente 1, tiene la boca en reposo mientras se visualiza la etiqueta con su graduación. El algoritmo de FO detecta esto como un cambio de lugar de los puntos de la máscara produciendo un ajuste extraño. (c) El paciente 5 es seguido con bastante precisión en general, pero la máscara no se ajusta al labio inferior en este gesto (imágenes extraídas de Chan 2016, procesadas con el método propuesto).

4.3.3 Evaluación en filmaciones de pacientes sin graduación

El método propuesto fue probado con tres vídeos de pacientes con parálisis facial parcial. La Figura 28 muestra el paciente 7 en el punto de máxima expresión para un ejercicio que consiste en mostrar los caninos. La asimetría del gesto es evidente, ya que solo muestra los caninos del lado derecho de su rostro.



Figura 28: Activación máxima para el paciente 7, ejercicio mostrar caninos (imagen procesada con el método propuesto).

Se graficó la distribución de probabilidad del ángulo de desviación de la velocidad respecto de la diagonal para efectos de comparación visual entre los sujetos (Figura 29). Se ha omitido el AUV *Brow Loweler* (b) y *Nose wrinkler* (nw).

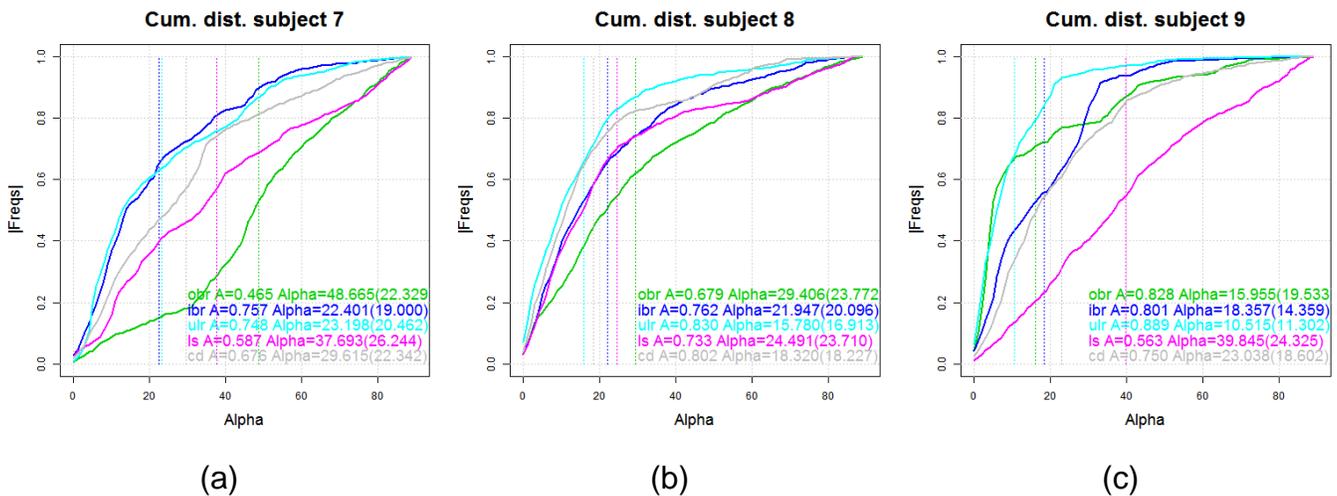


Figura 29: Distribución de probabilidad del ángulo de desviación respecto de la diagonal (α) para los pacientes sin graduación. Destacado en línea punteada está el ángulo promedio por cada nivel de activación de AUV. En la leyenda, se indica el área bajo la curva (AUC), α promedio y su desviación estándar (elaboración propia).

El AUC, α promedio y desviación estándar de α promedio para cada caso están disponibles en la Tabla 6. Estos datos son relevantes para realizar una graduación en escala HB de acuerdo al método propuesto en la sección anterior. Como un cálculo intermedio, los promedios de las AUC para estos pacientes están en la Tabla 7.

Nose wrinkler (nw)			
Sujeto	A.U.C.	Prom.(α)	D.S.(α)
7	0.71	26.90	20.72
8	0.84	14.81	13.14
9	0.75	23.29	14.77

Sujeto	Brow lowerer			Outer brow raiser (obr)			Inner brows raiser (ibr)		
	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)
7	0.50	45.58	23.66	0.46	48.67	22.33	0.76	22.40	19.00
8	0.46	48.93	26.76	0.68	29.41	23.77	0.76	21.95	20.10
9	0.25	68.20	11.55	0.83	15.96	19.53	0.80	18.36	14.36

Sujeto	AUV0 Upper lip raiser (ulr)			AUV2 Lip stretcher (ls)			AUV14 Lip corner depressor (cd)		
	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)	A.U.C.	Prom.(α)	D.S.(α)
7	0.75	23.20	20.46	0.59	37.69	26.24	0.68	29.61	22.34
8	0.83	15.78	16.91	0.73	24.49	23.71	0.80	18.32	18.23
9	0.89	10.52	11.30	0.56	39.84	24.32	0.75	23.04	18.60

Tabla 6: Área bajo la curva (AUC) de la distribución de probabilidad del ángulo (α) de desviación respecto de la diagonal. Promedio de α y su desviación estándar, para cada paciente 7, 8, y 9 (no se conoce a priori el grado de parálisis de esos pacientes en escala HB). Para facilitar la lectura, la tabla ha sido dividida en tres grupos de AUV, de arriba abajo, nariz, frente y boca respectivamente (elaboración propia).

Grado	AUVls, ulr, cd		AUVibr, obr		Todos los AUV		General - SD
	Boca	DS(boca)	Frente	DS(frente)	General	DS(general)	
Sujeto 7	0.67	0.08	0.61	0.21	0.65	0.12	0.52
Sujeto 8	0.79	0.05	0.72	0.06	0.76	0.06	0.70
Sujeto 9	0.73	0.16	0.81	0.02	0.77	0.12	0.64

Tabla 7: Promedio de las AUC de los niveles de activación de los AUV de la frente (ibr, obr) y de la boca (ls, ulr, cd). La tabla incluye también sus desviaciones estándar para los pacientes sin graduación. También se incluye el promedio de activación general de los AUV de la boca y frente (ibr, obr, ls, ulr y cd). La columna "General - SD" muestra la resta del promedio general y su desviación estándar (elaboración propia).

De acuerdo al método propuesto y los datos de la Tabla 7, la graduación de estos pacientes sería la siguiente:

- Paciente 7: promedio AUC de la frente menor a 0,65. El promedio del AUC de la boca es mayor a 0,65, por lo tanto su graduación es 5.
- Paciente 8: promedio AUC de la frente mayor a 0,65. El promedio del AUC de la boca es mayor a 0,75 por lo tanto su graduación es 1.
- Paciente 9: promedio AUC de la frente mayor a 0,65. El promedio del AUC de la boca está entre 0,70 y 0,75, por lo tanto su graduación es 2.

La Figura 30 muestra el seguimiento del sistema simplificado propuesto para el paciente 7. El ejercicio es similar para el paciente 8 y 9. Si se aplica el sistema simplificado de graduación de parálisis del nervio facial de House-Brackmann (Aik Kah et al. 2011), la graduación entregada por el método propuesto es coherente con lo observado en las filmaciones.

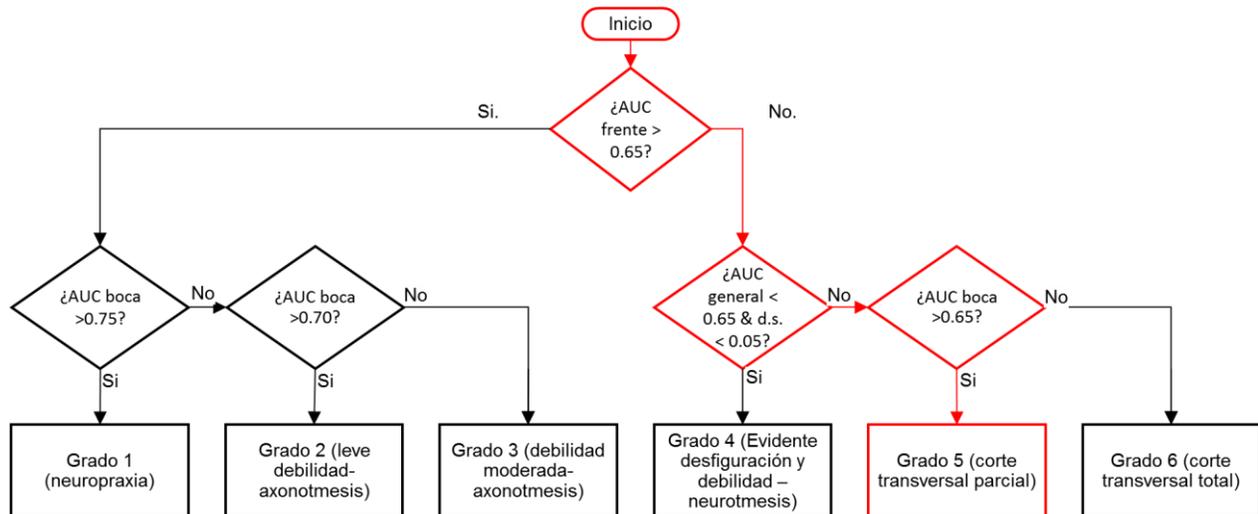


Figura 30: Aplicación del sistema propuesto al paciente 7 (elaboración propia).

5 Discusión

En esa tesis se ha propuesto un algoritmo de tracking facial, sin marcadores usando una cámara de uso común, para expresiones asimétricas y a partir de estos datos una equivalencia con una escala de parálisis clínica. El tracking está basado en la máscara Candide-3 asimétrica propuesta que es escalada al tamaño del rostro del paciente filmado de manera aproximada. Al comparar el tracking por FO y el seguimiento utilizando marcadores pasivos, la distancia entre el punto de interés y el marcador se mantuvo en 1,03 mm (con desviación estándar de 0,91mm), lo que indica un buen seguimiento. Al evaluar el resultado propuesto con 6 videos graduado clínicamente el criterio de graduación propuesto entrega una graduación correcta en todos los casos debido a que la calibración del método propuesto se hizo utilizando estas filmaciones y se cuenta con un ejemplo de cada clase. De acuerdo al método propuesto, los pacientes 7, 8 y 9 (sin graduación previa disponible) tendrían grados de parálisis facial 5, 1 y 2 respectivamente siendo este resultado acorde a la literatura.

A continuación se describen algunas observaciones rescatadas de la implementación del método propuesto:

Sinquinesia: El método propuesto supone que un paciente afectado de parálisis facial parcial puede mover un lado del rostro con dificultad, y por tanto que el diagnóstico del grado de parálisis puede realizarse comparando el nivel de activación de los AUV del lado paralizado contra el lado sano. Esto no es del todo correcto, ya que existe sinquinesia sobre todo en el centro del rostro del paciente. Esto significa que, si el paciente intenta sonreír, la fisura palpebral se desplaza por el movimiento de los músculos del lado no paralizado del rostro. También significa que, si al paciente se le pide cerrar los ojos con fuerza, probablemente apriete los labios fuertemente también.

Asimetría del rostro: Por otra parte, el lado paralizado del rostro puede manifestarse entre dos estados: relajación completa (parálisis flácida) a contracción completa (parálisis hipertónica) (Haase 2015). También puede mostrar asimetría debido a la tonicidad de la musculatura paralizada respecto de la conservada. La máscara Candide-3 asimétrica no considera estas deformaciones al momento de configurarse al rostro del paciente. Se trabajó con el supuesto que los movimientos pueden ser asimétricos, pero la estructura del rostro es simétrica.

Iluminación: El rostro de los pacientes cambia de iluminación según el gesto que realizan. Esto afecta al FO indicando desplazamientos que no son coherentes con la característica

que se desea seguir. Por ejemplo, en algunos casos al sonreír los puntos que representan los bordes de los labios se confunden con el brillo de los dientes, deformando la máscara Candide-3 asimétrica sin concordancia con el gesto. El paciente 7 en particular, tiene el rostro inclinado hacia atrás en la grabación, lo que dificulta el ajuste de la máscara Candide-3 asimétrico.

Sujetos con graduados: El método propuesto entrega la graduación de la parálisis en escala HB en una escala no discreta. Los puntos de corte para determinar cada grado de parálisis podrían ser calibrados a partir de la acumulación de experiencias. De todas maneras, podría ocurrir que entre dos grados, por ejemplo, grado 5 y 6, exista un traslape entre los puntos de corte. Es decir, para distinguir un grado de parálisis de otro en vez de utilizar un umbral AUC promedio de 0,65, podría definirse que menor a 0,7 es un grado y mayor a 0,63 es otro grado, quedando una ventana de ambigüedad entre ambos grados cuando el promedio está entre 0,63 y 0,7.

Artefactos: En la sección 4.2, se indicó que los AUV *Jaw z-push*, *Jaw x-push* y *Jaw drop* fueron descartados debido a que los sujetos y pacientes no realizan gestos que correspondan con estos movimientos. Durante la implementación del método propuesto, se observó que al considerar estos AUV, el ajuste de la expresión facial del modelo basado en la información del FO reporta niveles de activación irreales. Por ejemplo, si bien ningún sujeto en las filmaciones empuja su mandíbula hacia adelante (*Jaw z-push*), el método propuesto indicaba que la mandíbula estaba varios centímetros hacia adelante o hacia atrás del rostro. Esto tiene sentido si se considera que el modelo da estabilidad al tracking, pero posee muchos grados de libertad. En comparación, Candide-3 asimétrica tiene 11 grados de libertad más que el modelo original Candide-3.

6 Conclusiones

En este trabajo de tesis se muestra que la cuantificación del grado de parálisis facial mediante cámaras convencionales y algoritmos adaptados es factible. Después de una configuración del rostro de un sujeto, el algoritmo es fácil de utilizar y permite el seguimiento entre varias filmaciones diferentes. La herramienta de tracking existente (algoritmo original de Terissi et al. 2010) presenta un buen seguimiento de gestos simétricos. El método propuesto extiende su funcionalidad sobre gestos asimétricos y supera algunos problemas de flujo óptico. Se ha propuesto un método no clínico basado en información que es capturada en tiempo real o a partir de vídeos digitalizados. Se ha hecho una validación con un grupo de pacientes del Hospital San José. De acuerdo a la literatura, el método propuesto entrega una graduación coherente con el grado de parálisis aparente en estos pacientes.

El seguimiento facial no está exento de problemas. Comenzando por el ajuste de la máscara Candide-3 Asimétrica, problemas de iluminación en los vídeos, y el ángulo de la cámara respecto al sujeto. Estos problemas se resuelven teniendo especial cuidado al momento de realizar la filmación, asegurando una iluminación homogénea del rostro, manteniendo la frente del sujeto despejada y asegurando que el sujeto mire de frente a la cámara.

La máscara Candie-3 Asimétrica puede ser mejorada considerando la forma asimétrica del rostro. Como trabajo futuro, los parámetros de forma se pueden dividir en lado izquierdo y derecho tal como se hizo con los AUV de expresiones faciales. La configuración de un rostro inicial debe tener en cuenta la parálisis hipertónica donde un lado del rostro tiene una expresión facial fija. Actualmente esto se puede corregir ajustando el parámetro de la expresión facial a un estado inicial. Sin embargo, para su mejor uso debiera ser automático.

Actualmente existe tecnología disponible en computadoras y teléfonos móviles, que podrían correr una aplicación de las características descritas en este documento. Estos dispositivos cuentan con cámara, memoria y capacidad de proceso suficiente para correr los algoritmos descritos y las librerías de software utilizadas (OpenCV en otras) están disponible para varios sistemas operativos. Estos dispositivos también podrían desplegar estadísticas y gráficos históricos y en tiempo real que muestren el progreso del grado de parálisis. Todo lo anterior *en conjunto con una validación clínica* a mayor escala deja abierta la posibilidad de desarrollar aplicaciones móviles en el futuro para el apoyo del diagnóstico y tratamiento clínico de la parálisis facial.

7 Referencias

- Wang S, Li H, Qi F, Zhao Y. (2004). Objective facial paralysis grading based on Pface and eigenflow. *Med Biol Eng Comp* 2004;42(5):598–603.
- McGrenary S, O'Reilly BF, Soraghan JJ. (2005). Objective grading of facial paralysis using artificial intelligence analysis of video data. *Proc 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05)*, IEEE Computer Society Washington, DC, USA.
- Dong J., Ma L., Li Q., Wang S., Liu L., & Jian M. (2008) An approach for quantitative evaluation of the degree of facial paralysis based on salient point detection. *Proc. Intelligent Information Technology Application Workshops (IITAW08) Dec 21–22* , Shanghai, China.
- Tzou CH, Pona I, Placheta E, Hold A, Michaelidou M, Artner N, Kropatsch W, Gerber H, Frey M. (2012). Evolution of the 3-dimensional video system for facial motion analysis; ten years' experiences and recent developments. *Ann Plast Surg* 2012;69(2):173–85.
- Hadlock TA, Urban LS. (2012). Toward a universal: automated facial measurement tool in facial reanimation. *Arch Facial Plast Surg* 2012;14(4):277–82.
- Terissi L. D., & Gómez J. (2010). 3D Head Pose and Facial Expression Tracking using a Single Camera, *Journal of Universal Computer Science*, 16(6):903-920.
- Ahlberg J. (2001), CANDIDE-3 -- an updated parameterized face, Report No. LiTH-ISY-R-2326, Dept. of Electrical Engineering, Linköping University, Sweden.
- Haase D., Minnigerode L., Volk G. F., Denzler J., & Guntinas-Lichius O. (2015). Automated and objective action coding of facial expressions in patients with acute facial palsy. *Eur Arch Otorhinolaryngology*, 272(5), 1259-1267. <http://doi.org/10.1007/s00405-014-3385-8>
- Aik Kah, T., Hanom Annuar, F. (2011). *A systemic approach to facial nerve paralysis*. *Webmed Central Ophthalmology*, 2(4):WMC001856.
- Chand N. (Dr.), House Brackmann - Facial Nerve Grading, video publicado en Youtube (2016).
- OpenCV, (2017). Optical Flow, Obtenido desde http://docs.opencv.org/trunk/d7/d8b/tutorial_py_lucas_kanade.html#gsc.tab=0
- OpenCV, (2015). Real time pose estimation of a textured object. Obtenido desde http://docs.opencv.org/3.1.0/dc/d2c/tutorial_real_time_pose.html

Ouni S., Mangeonjean L., Steiner I. (2012). VisArtico: a visualization tool for articulatory data, *13th Annual Conference of the International Speech Communication Association – InterSpeech 2012*, Sep 2012, Portland, OR, United States.

Anexo A: AUV de la máscara Candide-3 asimétrica

La máscara Candide-3 asimétrica corresponde a la máscara Candide-3 original con los AUV de expresiones faciales divididas en lado izquierdo y derecho. La Figura A1 muestra las activaciones de los nuevos AUV para un lado, en nivel 1.

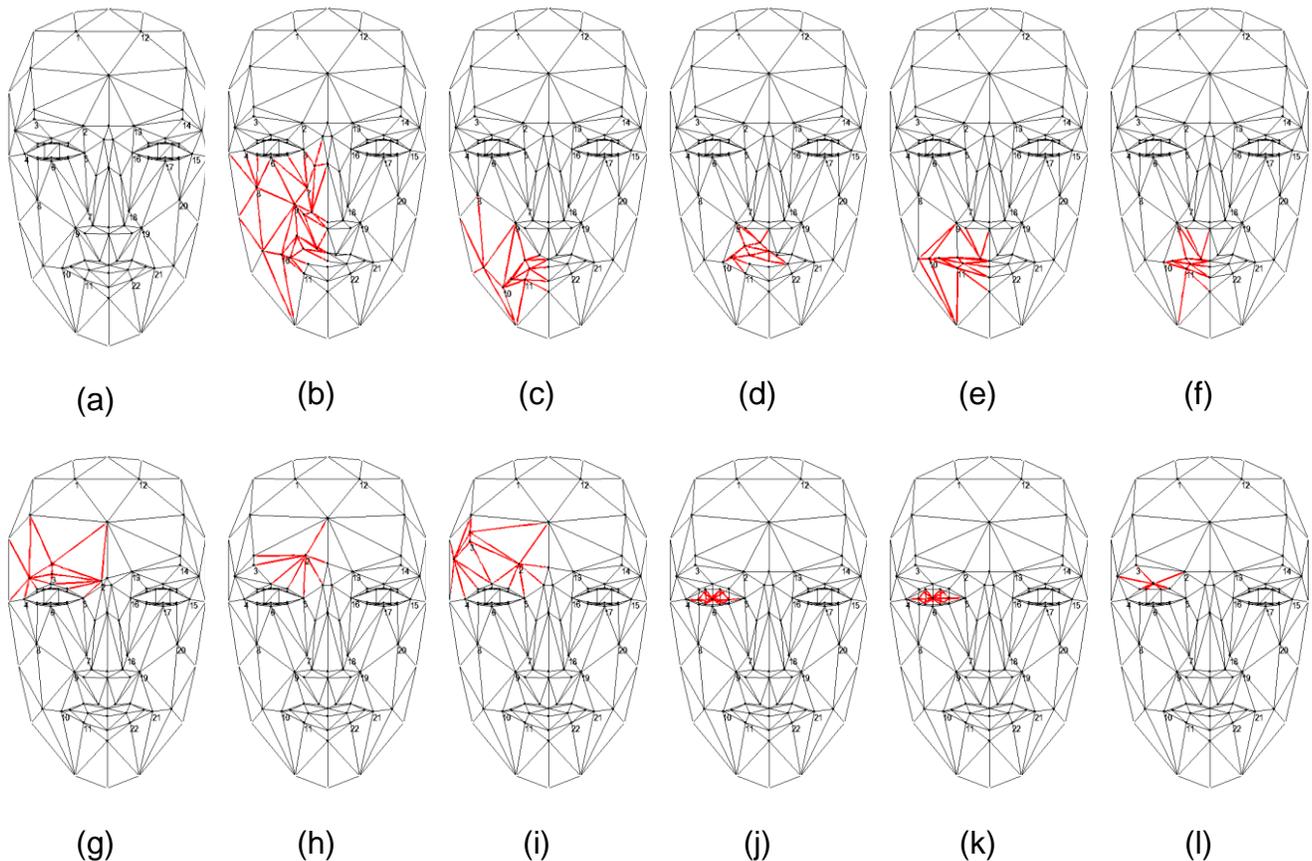


Figura A1: Máscara Candide-3 asimétrica con nivel de activación en 1.0 para un AUV específico del lado derecho del sujeto. Las líneas rojas indican las partes afectadas por la activación de cada AUV. (a) Máscara con todos los AUV en nivel de activación 0. (b) *Nose wrinkler R* (con nivel 1 de activación). (c) *AUV14 Lip corner depressor R*. (d) *AUV0 Upper lip raiser R*. (e) *AUV2 Lip stretcher R*. (f) *AUV9 Lip presser R*. (g) *Brow lowerer R*. (h) *Inner brows raiser R*. (i) *Outer brow raiser R*. (j) *AUV6 Eyes closed R*. (k) *AUV7 Lid tightener R*. (l) *AUV10 Upper lid raiser R*. Notar que los AUV *AUV7 Lid tightener (AU7)* y *AUV 6 Eyes closed (AU42/43/44/45)*, son muy similares. El AUV *AUV14 Lip corner depressor AU13/15* con valor positivo mueve el borde del labio hacia abajo, valor negativo mueve el borde del labio hacia arriba (elaboración propia).

Anexo B: Errores en el FO

Debido a los cambios de iluminación del rostro del sujeto o paciente, la librería utilizada para calcular el FO entrega, a veces, resultados extraños. Por ejemplo, en el vídeo del paciente 1, en el frame 105, se producen errores de FO tanto por el key frame como por el FO frame a frame (Figura B1).

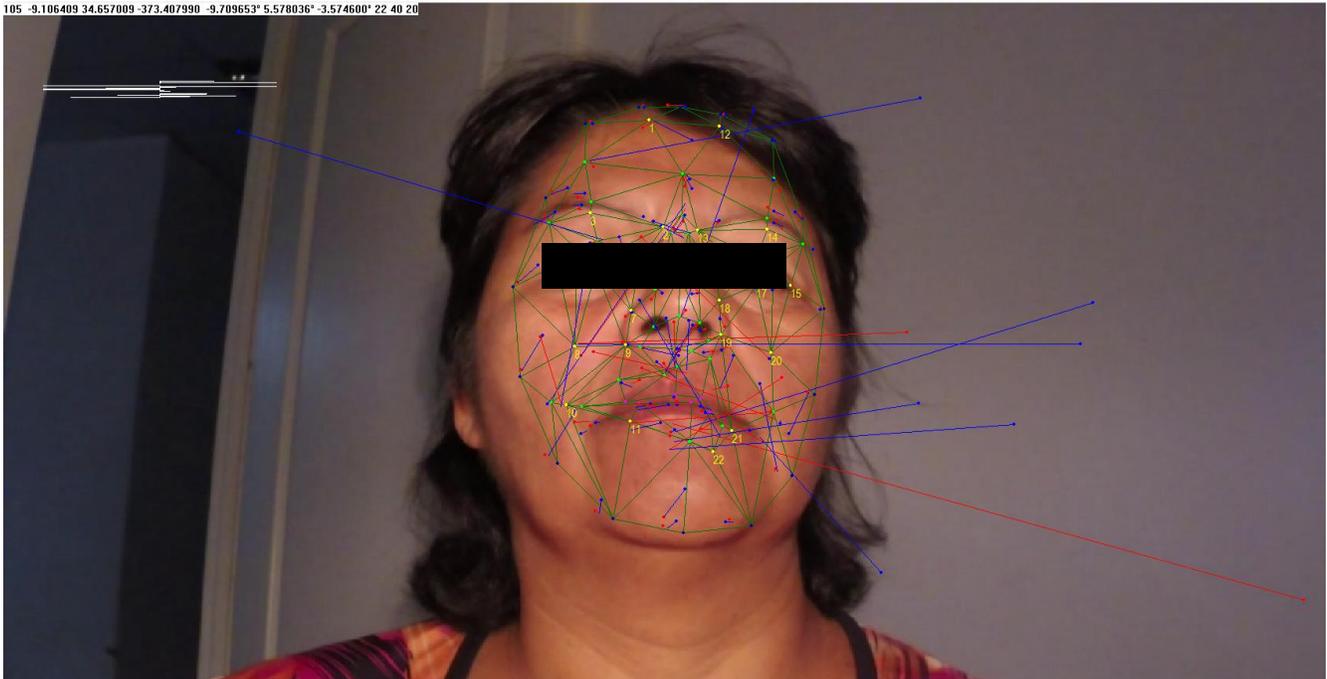


Figura B1: Errores de FO. Las líneas azules representan el FO frame a frame. Las líneas rojas, representan el FO respecto del key frame y el frame actual. Ambos flujos muestran varios puntos fuera del rostro.

Anexo C: Procedimiento de ajuste de marcadores pasivos y el método propuesto

El método propuesto entrega un log con la posición y la proyección de 22 puntos en un archivo de texto. Este log se editó para ser cargado en VisArtico. Se ajustó el tiempo de ambos marcadores usando la vista temporal (Figura C1), se igualó el sistema de coordenadas (Figura C2) y se comparó el desplazamiento de uno de los puntos de la boca (el punto 22 de acuerdo al método propuesto queda registrado como el punto 52 en VisArtico) muy cercano al marcador pasivo 19 (Figura C3). Finalmente, aprovechando los marcadores pasivos 25, 26, 27, 28 y 29 que corresponden a un sombrero que viste el sujeto, se eliminó la rotación de la cabeza y se comparó la apertura de la boca únicamente (Figura 25).

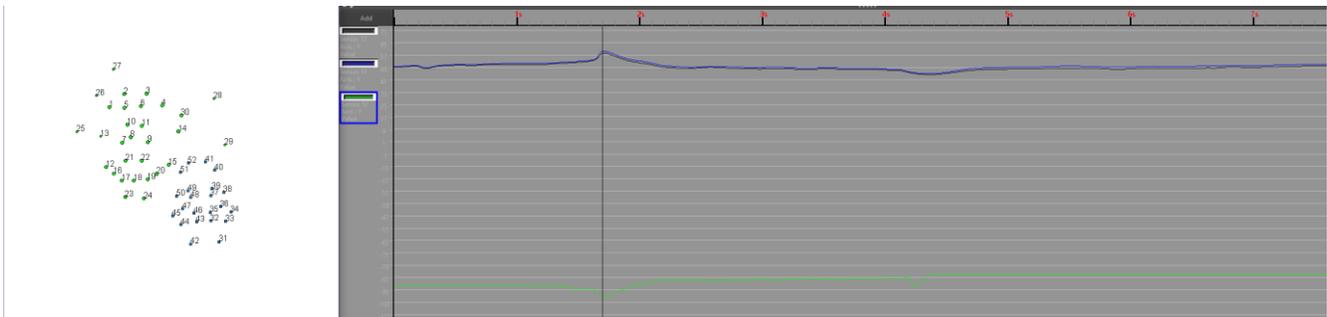


Figura C1: Ajuste de las grabaciones usando la vista temporal de VisArtico.

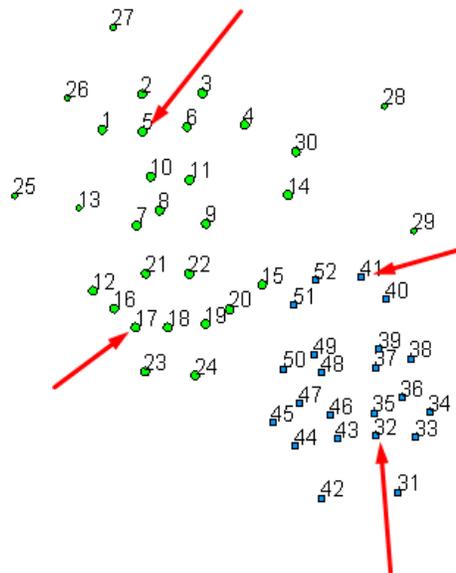


Figura C2: Puntos elegidos para efectuar la unificación del sistema de coordenadas. Notar que la información del método propuesto está invertida respecto del sistema de coordenadas del sistema de seguimiento cinemático, por lo que se igualan el punto 17 con el 32 y el 5 con el 41 (el punto 5 y 17 corresponden a marcadores pasivos, mientras que el punto 32 y 41 corresponden a puntos reportados por el método propuesto).



Figura C3: Punto de máxima apertura de la boca. Comparación del punto 19 (marcador pasivo) con el punto 22 del método propuesto (representado como punto 52 en VisArtico). El punto 17 tiene un movimiento similar al punto al 19.