

INSTITUTO TECNOLÓGICO DE APIZACO

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

**“ANÁLISIS DE LA EXPRESIÓN FACIAL CON KINECT PARA EL DIAGNÓSTICO
DE PARÁLISIS USANDO EL SISTEMA DE CLASIFICACIÓN NOTTINGHAM”**

TESIS

PARA OBTENER EL GRADO DE:
MAESTRO EN SISTEMAS COMPUTACIONALES

PRESENTA:
RAÚL CID CARRO

ASESORES:
DR. EDMUNDO BONILLA HUERTA
M.I.A. CARLOS PÉREZ CORONA

LÍNEA DE INVESTIGACIÓN:
INVESTIGACIÓN, DESARROLLO Y APLICACIÓN DE TECNOLOGÍAS INTELIGENTES

APIZACO, TLAXCALA; MÉXICO

NOVIEMBRE 2015

"2015, Año del Generalísimo José María Morelos y Pavón"

Apizaco, Tlax., 27 de octubre de 2015

No. de Oficio: DEPI/378/15

ASUNTO: **Se Autoriza Impresión de Tesis de Grado.**

ING. RAUL CID CARRO
CANDIDATO AL GRADO DE MAESTRO
EN SISTEMAS COMPUTACIONALES
No. de Control: **M13370003**
PRESENTE.

Por este medio me permito informar a usted, que por aprobación de la Comisión Revisora asignada para valorar el trabajo, mediante la Opción: **I Tesis de Grado por Proyecto de Investigación**, de la **Maestría en Sistemas Computacionales**, que presenta con el tema: "**ANÁLISIS DE LA EXPRESIÓN FACIAL CON KINECT PARA EL DIAGNÓSTICO DE PARALISIS USANDO EL SISTEMA DE CLASIFICACIÓN NOTTINGHAM**" y conforme a lo establecido en el Procedimiento para la Obtención del Grado de Maestría en el Instituto Tecnológico, la División de Estudios de Posgrado e Investigación a mi cargo le emite la:

AUTORIZACIÓN DE IMPRESIÓN

Debiendo entregar un ejemplar del mismo debidamente encuadernado y seis copias en CD en formato PDF, para presentar su Acto de Recepción Profesional a la brevedad.

Sin otro particular por el momento, le envío un cordial saludo.

ATENTAMENTE
PENSAR PARA SERVIR, SERVIR PARA TRIUNFAR®


M.A.D. MA. A. ACELA DAVILA JIMENEZ
JEFA DE LA DIVISIÓN DE ESTUDIOS DE
POSGRADO E INVESTIGACIÓN.



Secretaría de Educación Pública
Instituto Tecnológico de Apizaco
División de Estudios de Posgrado
e Investigación

MAADJ/mebr.

C.p. Expediente.



"2015, Año del Generalísimo José María Morelos y Pavón"

Apizaco, Tlax., 26 de octubre de 2015

ASUNTO: Aprobación del trabajo de Tesis de Maestría.

M.A.D. MA. A. ACELA DÁVILA JIMÉNEZ
JEFA DE LA DIVISIÓN DE ESTUDIOS DE
POSGRADO E INVESTIGACIÓN.
PRESENTE.

Por este medio se le informa a usted, que los integrantes de la **Comisión Revisora** para el trabajo de tesis de maestría que presenta el **ING. RAÚL CID CARRO**, con número de control **M13370003** candidato al grado de **Maestro en Sistemas Computacionales** y egresado del **Instituto Tecnológico de Apizaco**, cuyo tema es **"ANÁLISIS DE LA EXPRESIÓN FACIAL CON KINECT PARA EL DIAGNÓSTICO DE PARÁLISIS USANDO EL SISTEMA DE CLASIFICACIÓN NOTTINGHAM"**, fue:

A P R O B A D O

Lo anterior, al valorar el trabajo profesional presentado por el candidato y constatar que las observaciones que con anterioridad se le marcaron así como correcciones sugeridas para su mejora ya han sido realizadas.

Por lo que se avala se continúe con los trámites pertinentes para su titulación.

Sin otro particular por el momento, le envió un cordial saludo.

LA COMISIÓN REVISORA


DR. EDMUNDO BONILLA HUERTA


M.I.A. CARLOS PÉREZ CORONA


DR. JOSÉ FEDERICO RAMÍREZ CRUZ


DR. ROBERTO MORALES CAPORAL

C. p.- Interesado.

“Imagination is more important than knowledge. For knowledge is limited to all we know and understand, while imagination embraces the entire world, and all there ever will be to know and understand.”

Albert Einstein

DEDICATORIA

A toda mi familia, porque son una fuente de inspiración y motivación que me permite lograr mis objetivos y conquistar mis aspiraciones. En especial a mi padre, quien me obsequió el mejor regalo que alguien pudo darme, creer en mí.

Quizás has desaparecido de mi vida, pero nunca de mi corazón.

AGRADECIMIENTOS

A mis maestros de la Maestría en Sistemas Computacionales,
por sus enseñanzas y dedicación.

Al Dr. Edmundo Bonilla Huerta por haber dirigido mi trabajo de
tesis.

Al Instituto Tecnológico de Apizaco, por permitirme realizar mis
estudios de maestría.

Al CONACYT, por solventar mis estudios de maestría y darme
la oportunidad de alcanzar este objetivo personal de carácter
profesional.

RESUMEN

El presente trabajo de tesis muestra la implementación de un sistema para clasificar el nivel de parálisis facial que puede presentar una persona; a fin de lograr tal objetivo, se propone al *Kinect para Windows*® como una herramienta de desarrollo que facilite la extracción de los parámetros faciales necesarios para ser utilizados por el sistema de clasificación Nottingham.

El objetivo principal es brindar un resultado objetivo y lo más preciso posible que le permita al médico corroborar su diagnóstico y elimine la subjetividad durante el procedimiento de valoración diagnóstica. El sistema permite guardar imágenes del rostro del paciente junto con archivos de texto asociados que contienen información sobre la simetría facial. Estos archivos digitales generados, permitirán tener un expediente médico más completo; de esta manera es posible observar, comparar y comprobar el progreso del paciente.

La metodología para el desarrollo del sistema está compuesta de las siguientes etapas:

La primera etapa consiste en la identificación automática de los ocho puntos característicos en el rostro del paciente con ayuda del *Kinect*®. Estos puntos son necesarios para emplear el sistema de clasificación Nottingham.

En la segunda etapa se captura la imagen del rostro del paciente en reposo y al momento de realizar cada uno de los tres diferentes movimientos faciales (levantando las cejas, cerrando estrechamente los ojos y sonriendo), tal como lo establece el sistema de clasificación Nottingham.

En la tercera y última etapa se calcula el grado de parálisis al comparar las diferencias en las distancias entre el lado sano y el lado afectado del rostro de cada una de las imágenes obtenidas.

ABSTRACT

This thesis document presents the implementation of a system for classifying the level of facial paralysis that a person can be had in order to achieve this objective, it has been used a *Kinect for Windows*® as a development tool that facilitates the extraction of the facial parameters required to be used by the Nottingham grading system.

The main objective is to provide an appropriate and accurate result that allows to the doctor to confirm the diagnosis and eliminate the subjectivity during the diagnostic evaluation procedure. The system allows to save images of the patient's face, along with associated text files that contain information about facial symmetry. The obtained digital files would permit a more complete medical file; thus, it is possible to observe, compare and follow the health progress of the patient.

The methodology for the development of the system consists of the following stages: The first stage involves the automatic identification of the eight characteristic points in the face using the *Kinect*®. These points are necessary in order to use the Nottingham grading system.

In the second stage the system save the image of the patient's face at rest and at the time of doing each of the three different facial movements (raising eyebrows, tightly closing his eyes and smiling), as established in the Nottingham grading system.

In the third and final stage, the system calculate the degree of paralysis by comparing the differences in the distances between the healthy and the affected side of the face of each of the acquired images.

TABLA DE CONTENIDO

Resumen.....	I
Abstract	II
Tabla de contenido.....	III
Índice.....	IV
Índice de figuras	VII
Índice de tablas	IX

ÍNDICE

CAPÍTULO 1: INTRODUCCIÓN.....	1
1.1 Introducción	2
1.2 Planteamiento del problema	4
1.3 Justificación	5
1.4 Hipótesis	6
1.5 Objetivo general.....	6
1.5.1 Objetivos específicos	6
1.6 Pregunta de investigación.....	7
1.7 Organización de la tesis.....	7
1.8 Estado del Arte.....	8
CAPÍTULO 2: MARCO TEÓRICO.....	15
2.1 Introducción	16
2.2 Topografía del Sistema Nervioso.....	17
2.2.1 Sistema Nervioso Periférico (SNP).....	17
2.3 Parálisis Facial.....	22
2.3.1 Signos y síntomas.....	24
2.3.2 Epidemiología	25
2.3.3 Causas.....	26
2.3.4 Diagnóstico	26
2.4 <i>Kinect</i> ®.....	33
2.4.1 Especificaciones técnicas.....	34
2.4.2 Interfaz de Seguimiento Facial de <i>Kinect</i> ®.....	38
2.5 C Sharp (C#).....	47
2.5.1 Ventajas frente a otros lenguajes.....	51

2.5.2 Desventajas frente a otros lenguajes.....	52
2.6 XAML	52
2.6.1 Código XAML de la herramienta de adquisición de información.....	53
2.6.2 Ventajas frente a otros lenguajes.....	57
CAPÍTULO 3: METODOLOGÍA.....	58
3.1 Introducción	59
3.2 Visión global del sistema.....	59
3.3 Adquisición de los datos	60
3.3.1 Establecimiento de la comunicación con el <i>Kinect</i> ®.....	60
3.3.2 Detección y seguimiento del rostro de una persona	62
3.3.3 Discriminación de la información facial y visual de la persona	67
3.4 Almacenamiento de los datos	70
3.5 Evaluación de los datos	72
3.6 Descripción de la interfaz de usuario	76
3.6.1 Lienzo principal.....	76
3.6.2 Selección de la carpeta de almacenamiento	77
3.6.3 Selección de la expresión facial.....	77
3.6.4 Captura de imagen	77
3.6.5 Diagnóstico	77
CAPÍTULO 4: PRUEBAS Y CONCLUSIONES	79
4.1 Introducción	80
4.2 Análisis de pruebas.....	80
4.2.1 Consistencia en el diagnóstico	80
4.2.2 Aplicación del sistema en pacientes con parálisis facial	81
4.3 Conclusiones.....	87

4.4 Trabajo a futuro.....	88
REFERENCIAS BIBLIOGRÁFICAS	90
ANEXOS I: PUBLICACIONES	94
ANEXOS II: VINCULACIÓN	116

ÍNDICE DE FIGURAS

Fig. 1.1 Lesión del nervio facial que produce falta de movilidad y sensibilidad en el lado de la cara afectado.	3
Fig. 1.2 Captura de video del sistema de análisis de video tridimensional.....	13
Fig. 2.1 Anatomía topográfica del Sistema Nervioso.....	16
Fig. 2.2 Vista inferior de un cerebro humano y tronco cerebral mostrando los pares craneales.....	20
Fig. 2.3 Ilustración del origen y trayecto del nervio facial derecho e izquierdo y su recorrido.	21
Fig. 2.4 Clasificación de la parálisis facial.	23
Fig. 2.5 Puntos de referencia facial designados en el método Nottingham.....	32
Fig. 2.6 Estructura del Kinect® y sus componentes principales.....	34
Fig. 2.7 Rango válido del sensor de profundidad.	35
Fig. 2.8 Comparativa entre resoluciones.....	37
Fig. 2.9 Espacio de la cámara.....	38
Fig. 2.10 Puntos supuestamente reconocibles por Kinect®.....	44
Fig. 2.11 Mapeado de los puntos reconocibles y el modelo CANDIDE-3.....	45
Fig. 2.12 Ángulos de la postura de la cabeza.....	46
Fig. 2.13 Imagen RGB captada por Kinect®.....	49
Fig. 2.14 Superposición de puntos faciales.....	51
Fig. 2.15 Interfaz de usuario en XAML.....	56
Fig. 3.1 Estructura del sistema.....	60
Fig. 3.2 Diagrama de flujo del algoritmo de seguimiento facial usado por Kinect®65	
Fig. 3.3 Confirmación de detección del rostro.....	67
Fig. 3.4 Visualización de los índices de los puntos faciales reconocibles.....	68
Fig. 3.5 Visualización de sólo los ocho puntos de interés.....	69
Fig. 3.6 Distancias entre los puntos de interés.....	70
Fig. 3.7 Formato del nombre de los archivos.....	71
Fig. 3.8 Expresiones faciales que debe realizar el paciente.....	73
Fig. 3.9 Guía para el diagnóstico de parálisis facial mediante el sistema de clasificación Nottingham.....	73

Fig. 4.1 Resultados de la prueba de consistencia	81
Fig. 4.2 Fotografías de las cuatro expresiones faciales obtenidas por el sistema .	83
Fig. 4.3 Recuperación esperada por el médico especialista	86
Fig. 4.4 Comparación entre cámaras del modelo original (v1) y el modelo reciente (v2)	89

ÍNDICE DE TABLAS

Tabla 1.1 Lista de los sistemas objetivos de análisis de diagnóstico en las últimas décadas.....	9
Tabla 2.1 Escala de Botman y Jongkees.	28
Tabla 2.2 Escala de May.	28
Tabla 2.3 Escala de Smith.....	29
Tabla 2.4 Puntuación de Parálisis Facial de Stennert.	29
Tabla 2.5 Puntuación de Parálisis Facial (Defecto Secundario) de Stennert.	30
Tabla 2.6 Sistema de Clasificación de Parálisis Facial Sunnybrook (Ross et al). .	32
Tabla 2.7 Formatos soportados por DepthImageStream.	36
Tabla 2.8 Formatos soportados por ColorImageStream.	37
Tabla 2.9 Principales interfaces de <i>Kinect</i> ®.....	39
Tabla 2.10 Estructuras de datos utilizadas por la API.	39
Tabla 2.11 Explicación de los ángulos medidos por <i>Kinect</i> ®.	46
Tabla 3.1 Índices de expresiones faciales.....	71
Tabla 4 1 Obtención de distancias entre los puntos característicos.....	84
Tabla 4 2 Cálculo del valor de movilidad para ambos lados del rostro.....	85

CAPÍTULO 1: INTRODUCCIÓN

1.1 Introducción

Las ciencias de la salud, y la medicina en particular, son uno de los campos del saber más evolucionados y beneficiados por el uso de tecnología, al tiempo que registran un crecimiento exponencial tanto en el número de usuarios, como en el de instituciones que se han incorporado a la búsqueda de diferentes medios que permitan tener un mejor nivel de vida. Según David Banta, una tecnología (médica) es cualquier técnica o herramienta, producto o proceso, método o aparato, que permita ampliar las capacidades humanas [1]. Una forma común de clasificar a las tecnologías médicas es la siguiente:

- a) Tecnologías preventivas: Su propósito es proteger al individuo, impidiendo la presentación y/o el avance de la enfermedad.
- b) Tecnologías de diagnóstico: Sirven para determinar los momentos y/o procesos patológicos por los que pasa un paciente.
- c) Tecnologías terapéuticas o de rehabilitación: Su función es eliminar a la enfermedad o minimizar sus efectos (aquí cabe una subdivisión, para especificar que existen tecnologías con un sentido realmente curativo y otras de carácter paliativo, llamadas por Lewis Thomas *half-way technologies* [2]).
- d) Tecnologías de administración y organización: Son estratégicas para el otorgamiento correcto y oportuno de los servicios de salud [3].

En México la necesidad por los diferentes tipos de tecnología médica va en aumento. Según el censo realizado por el Instituto Nacional de Estadística y Geografía (INEGI) en el año 2010, existen en México 5 millones 739 mil 270 personas con discapacidad, lo que representa el 5.1% de la población total del país, en donde la parálisis facial ocupa uno de los diez primeros lugares de atención en diversas unidades de rehabilitación del país; esto debido en gran parte a que el nervio facial es el que más se lesiona de todos los pares craneales. La parálisis se produce cuando dicho nervio está inflamado o comprimido, sin embargo, se desconoce qué es lo que causa el daño. La mayoría de los investigadores coinciden que una infección viral causa el trastorno. Se cree que el nervio facial se inflama como reacción a la infección, causando presión dentro del acueducto de Falopio y

llevando a un infarto (muerte de las células nerviosas debido al suministro insuficiente de sangre y oxígeno). Esta lesión impide el movimiento normal de diferentes regiones del rostro como la frente, párpados y boca; así como la expresión de las emociones debido a la asimetría del rostro (ver Fig. 1.1 [4]). Dicho padecimiento exige una evaluación exhaustiva y precisa por parte del médico, debido a la discapacidad funcional motora que ocasiona, tiene un efecto negativo sobre la calidad de vida del paciente.

Existen dos tipos de parálisis facial: central y periférica. La parálisis facial central es contralateral a la lesión afectando a la mitad del rostro inferior mientras que la parálisis facial periférica es ipsilateral afectando a toda la mitad del rostro, este último tipo de parálisis tiene una incidencia epidémica, lo que podría sugerir un mecanismo etiopatogénico de tipo infeccioso o inflamatorio. Su incidencia es de alrededor de 20 a 30 casos por cada 100 mil personas al año. La incidencia de casos aumenta cuando se presenta hipertensión y altos niveles de colesterol. Esto resulta preocupante si se tiene en cuenta que en México, 80% de las personas con niveles inadecuados de colesterol lo desconocen y 52% de las y los hipertensos no saben que tienen presión alta. Pese a la alta incidencia en México, no se cuenta con información accesible destinada a los pacientes sobre este padecimiento y su tratamiento.



Fig. 1.1 Lesión del nervio facial que produce falta de movilidad y sensibilidad en el lado de la cara afectado.

Por lo tanto, la evaluación del grado de parálisis es muy importante, ya que contar con un método capaz de clasificar el nivel de movilidad facial de manera precisa, asegura tener un buen diagnóstico y tratamiento adecuado para el paciente. Uno de los retos inherentes en la clasificación de la parálisis del nervio facial, es alcanzar un consenso sobre precisamente lo que un sistema de clasificación de la función del nervio facial debe lograr. Un sistema de clasificación ideal sería un instrumento bien calibrado que sea conveniente de usar, sería a la vez objetivo, preciso y fiable, además de fácil aplicación ya sea en la clínica o en el laboratorio de investigación.

Este trabajo propone como una posibilidad usar el *Kinect para Windows*® como una herramienta de desarrollo que ayude en el diagnóstico de pacientes con algún tipo de parálisis facial, asimismo plantea el uso del sistema de clasificación Nottingham como evaluador de la función del nervio facial. El *Kinect*® es un dispositivo periférico inicialmente pensado como un simple controlador de juego, que gracias a los componentes que lo integran: sensor de profundidad, cámara RGB, arreglo de micrófonos y sensor de infrarrojos (emisor y receptor), es capaz de detectar cualquier movimiento corporal o sonido y posteriormente interpretarlo en acciones dentro de la interfaz. Gracias a toda la información que captura este dispositivo, es posible hacer uso de él para programar toda una serie de aplicaciones para diferentes áreas y usos diferentes a los lúdicos.

1.2 Planteamiento del problema

Generalmente las personas que registran síntomas como debilidad o entumecimiento en el rostro se dirigen con un médico especialista, el cual debe llevar a cabo una cuidadosa evaluación clínica y prueba física en la que se le pide al paciente que realice diferentes movimientos faciales, con el fin de determinar el rango de movilidad en los diferentes lados de su cara. Una vez terminada la prueba el médico evalúa los datos obtenidos y realiza un diagnóstico, el cual resulta determinante al momento de elegir el tratamiento correcto. Sin embargo, el procedimiento descrito previamente carece de una evaluación precisa, debido a que los resultados clínicos se ven influenciados por la subjetividad de cada médico, además la reproducibilidad del diagnóstico no es suficiente.

Ante ello se considera necesario la elección de un método preciso para la clasificación de la función del nervio facial como un prerrequisito para el diagnóstico efectivo y tratamiento de pacientes con parálisis. La anatomía del nervio facial es compleja y su fisiología única por lo que representa un reto en la valoración del daño en la movilidad facial. El enfoque tradicional utiliza valoraciones subjetivas por lo que el desarrollo de una herramienta de automatización en el proceso de diagnóstico médico, resulta de gran beneficio para el médico especialista, ya que le permitiría contar con un diagnóstico objetivo que utilizaría para corroborar el propio.

La información visible que es valorada mediante el juicio subjetivo del médico, ahora también pueden ser evaluada por un sistema inteligente capaz de reconocer y obtener la posición de 8 puntos anatómicos faciales claves a través del *Kinect*® en tiempo real, para que; utilizando el sistema de clasificación Nottingham, se compare la posición de cada uno de las puntos conseguidos en ambos lados del rostro con el fin de determinar una diferencia en la amplitud de movimiento entre los dos lados de la cara y de esta manera, poder asistir al médico en el diagnóstico al brindarle un resultado objetivo que le facilite elegir la terapia adecuada de recuperación.

1.3 Justificación

En la actualidad, se utilizan diferentes sistemas tradicionales de clasificación de la función del nervio facial que la mayoría de médicos emplean para determinar el grado de la enfermedad. Sin embargo, existen problemas en términos de objetividad y reproducibilidad en estos métodos que a su vez tienen la limitante de no poder ofrecer una evaluación precisa y resulta difícil entender el estado de recuperación de los pacientes. Aunque se han realizado varios intentos para resolver estos problemas, estos métodos no han sido ampliamente aplicados para el diagnóstico real tampoco.

Por tal motivo el uso innovador de herramientas tecnológicas como el *Kinect*®, aplicado en la resolución de problemas en el área médica, beneficia tanto al paciente como al médico, ya que el propósito de automatizar ciertos procesos, no es deshumanizar la práctica médica, sino eliminar lo más posible el error humano presente en ciertos procesos ambiguos como el diagnóstico.

Con el trabajo propuesto, más que resolver un problema, se busca mejorar el proceso de diagnóstico; al ofrecer una herramienta útil y fácil de usar que le permite al médico contar con datos confiables sobre la fisonomía de un paciente. De igual manera, el sistema ofrece la posibilidad de llevar un registro temporal fotográfico del paciente, permitiéndoles tanto al médico como al paciente observar los avances obtenidos de la terapia y/o cualquier otro tratamiento.

1.4 Hipótesis

A pesar que el *Kinect*® fue creado como una herramienta lúdica, este trabajo considera su potencial en el área médica, proponiéndolo como un recurso de apoyo en la evaluación de pacientes que han sufrido algún tipo de parálisis facial, al realizar un análisis alternativo, pero sobre todo objetivo, que permite al médico tener una confirmación en su diagnóstico, evitando con esto la subjetividad en el procedimiento. De igual forma, brinda un registro visual del progreso de cada paciente. Este sistema brinda una ayuda accesible e innovadora para cualquier clínica.

1.5 Objetivo general

Desarrollar un sistema que utilice al *Kinect*® como una herramienta de extracción de puntos faciales característicos en tiempo real, que permitan realizar de manera objetiva un diagnóstico alterno de parálisis facial mediante el sistema de clasificación Nottingham. El sistema debe crear un registro de datos e imágenes que permita verificar el progreso en la movilidad facial de cada paciente.

1.5.1 Objetivos específicos

- ✓ Elaborar y diseñar un sistema capaz de realizar la identificación y seguimiento del rostro de la persona en tiempo real usando la cámara del *Kinect*®.
- ✓ Obtener el modelo genérico en 3D (CANDIDE-3) del rostro detectado y guardar la ubicación de los vértices de los polígonos que lo conforman.
- ✓ Identificar los 8 vértices que mejor representen los puntos de interés en el sistema de clasificación Nottingham y sólo mostrar éstos sobre el rostro.

- ✓ Guardar en un archivo de texto las coordenadas de los 8 puntos de interés, de igual forma, guardar una imagen en formato JPG del rostro del paciente en reposo y al momento de realizar los 3 ejercicios de movilidad facial (levantando las cejas, cerrando estrechamente los ojos y sonriendo) que requiere el sistema de clasificación Nottingham.
- ✓ Calcular la distancia entre los puntos de interés obtenidos para obtener las diferencias en amplitud de movimiento existentes entre el lado sano y el lado afectado del rostro para cada imagen guardada.
- ✓ Determinar el nivel de parálisis facial evaluando las diferencias en amplitud de movimiento utilizando el sistema de clasificación Nottingham.
- ✓ Someter el sistema a pruebas con distintos pacientes para comprobar su efectividad.

1.6 Pregunta de investigación

¿Es posible la automatización del sistema de clasificación del nivel de parálisis facial Nottingham mediante el desarrollo de un sistema que utilice al *Kinect*® como una herramienta de adquisición de datos faciales, para que éste sea usado como una herramienta de ayuda en el diagnóstico de dicha enfermedad?

1.7 Organización de la tesis

El presente trabajo de tesis está dividido en 4 capítulos, a continuación se presenta una breve descripción de cada uno de ellos.

Capítulo I. En este capítulo se presenta la introducción al documento de tesis, la descripción del problema, la justificación, la hipótesis, los objetivos generales y específicos que se pretenden cumplir, la pregunta de investigación y el análisis del estado de arte.

Capítulo II. En este capítulo se incluye el marco teórico, describiendo los componentes involucrados en el nervio facial, posteriormente se explica la parálisis facial y se termina con una descripción del *Kinect*®.

Capítulo III. En este capítulo se presenta la metodología y las herramientas utilizadas para realizar el desarrollo del sistema y poder realizar la clasificación de la función del nervio facial correctamente.

Capítulo IV. En este capítulo se describen las conclusiones obtenidas durante todo el desarrollo del presente trabajo de tesis y se sugieren los posibles trabajos futuros.

1.8 Estado del Arte

La necesidad de métodos objetivos y consistentes para la evaluación de la parálisis facial dio como resultado proponer la aplicación de la tecnología en la evaluación de la función facial. El objetivo de estos recientes enfoques es describir los movimientos faciales cuantitativamente con mayor precisión y reproducibilidad de lo que sería con los métodos tradicionales; ya que, si bien estos métodos son simples, no consumen mucho tiempo y no requieren ningún equipo especial, sus resultados no son del todo exactos. Estas deficiencias son inherentes en todos los métodos de evaluación visual en donde se involucra el criterio de una persona, por tal motivo, se han creado varios sistemas de clasificación computarizados que buscan compensar estos problemas.

En la literatura dichos sistemas de clasificación pueden dividirse según su enfoque en dos tipos. El primer tipo se refiere a los sistemas en los que se automatiza un método tradicional médico en donde se utilizan principalmente técnicas de procesamiento de imágenes y reconocimiento de patrones. Con el fin de evaluar el grado de movimiento en las diferentes regiones de la cara, se les pide primero a los pacientes que realicen un cierto número de expresiones faciales por separado. La idea básica es comparar las diferencias entre los puntos clave o regiones en los dos lados de una cara. Las técnicas que intervienen en este tipo de métodos incluyen la detección del rostro, la detección de punto o región saliente, el agrupamiento y esquemas de aprendizaje automático. Este tipo de métodos pueden requerir la interacción del usuario con el software del sistema.

El segundo tipo de métodos tiene como objetivo evaluar automáticamente la parálisis facial y proporcionar los resultados de la evaluación sin que intervenga el usuario. Al igual que el primer tipo de métodos, se les pide a los pacientes que

realicen movimientos faciales instruidos por médicos. Los métodos de detección de regiones salientes se utilizan para medir la asimetría entre los dos lados de la cara y compara las variaciones de expresión entre el lado anormal y el lado normal. Se emplean máquinas de soporte vectorial para producir el grado de parálisis.

Cabe señalar que los dos tipos de métodos descritos anteriormente tienen el mismo problema clave, el cual puede ser descrito usando variabilidad intrínseca y extrínseca. La variabilidad intrínseca en los individuos puede ser debido a múltiples factores, incluyendo lesiones y cambios relacionado con la edad. La variabilidad extrínseca por otro lado, está relacionada con las condiciones ambientales y puede ser causada por la orientación de la vista, la iluminación, las sombras y los reflejos. Las variaciones de estos factores pueden dar como resultado una disminución significativa en la tasa de reconocimiento facial, y por lo tanto hacer que sea difícil detectar los puntos característicos [5].

El diagnóstico objetivo de la función facial es necesaria para planear y evaluar la intervención terapéutica en pacientes con parálisis facial. La Tabla 1.1 muestra algunos sistemas propuestos para analizar de manera objetiva el movimiento facial, todos ellos utilizan diversas técnicas para visualizar movimientos faciales complejos mediante la medición y análisis de puntos de referencia en el rostro, píxeles, variación en los reflejos de luz y las líneas de contorno de la superficie a través de fotografías o videograbaciones de los movimientos faciales [6].

Tabla 1.1 Lista de los sistemas objetivos de análisis de diagnóstico en las últimas décadas.

Autor	Método	Datos adquiridos
Burres 1985	Registro electromiográfico de la superficie y medición de puntos de referencia faciales especificados en reposo y durante expresiones faciales definidas con pinzas de mano	Desplazamiento de los marcadores durante el movimiento; sin análisis multiregional
Neely et al, 1992	Reflectancia de la luz facial	Cambios del contorno facial; ninguna medición de la dirección de la magnitud del movimiento

Johnson et al, 1994	Seguimiento de marcadores faciales en fotografías	Medición de la dirección del desplazamiento de los marcadores
Yuen et al, 1994	Teoría de la interferencia de la luz, en la que la iluminación a través de una rejilla crea líneas de contorno sobre un objeto	Evaluación cuantitativa de los movimientos faciales y la gravedad de la parálisis facial
Isono et al, 1996	Seguimiento de marcadores faciales en fotografías	Medición de la dirección del desplazamiento de los marcadores
Meier-Gallati et al, 1998	Reflectancia de la luz facial	Cambios del contorno facial; ninguna medición de la dirección de la magnitud del movimiento
Sargent et al, 1998	Fotografías digitales, seguimiento de los marcadores faciales, resta de píxeles	Análisis en 2D del movimiento facial
Frey et al, 1998	Video análisis en 3D, seguimiento de los marcadores faciales estáticos y dinámicos	Cuantificación en 3D de desplazamiento, dirección y patrón de tiempo de los movimientos faciales, análisis de la simetría global de la cara
Coulson et al, 1999	Video análisis en 3D, seguimiento de los marcadores faciales	Mediciones en 3D de los desplazamientos de los marcadores faciales
Somia et al, 2000	Seguimiento de los marcadores faciales alrededor de los ojos, casco especial para la cámara	Medición del desplazamiento, velocidad y aceleración del movimiento del párpado
Heckman y Schön-Hupka 2001	Fotografías digitales en serie, técnica de superposición digital	Medición de la movilidad de la frente: distancias, áreas y ángulos
Watchman et al, 2001	Video análisis en 2D, puntos de referencia marcados en la imagen digital	Velocidad, aceleración, dirección del movimiento y desplazamiento
Linstrom 2002	Fuente de luz infrarroja, sensibilidad de la luz infrarroja, marcadores faciales reflectantes de luz de la cámara	Cuantificación del desplazamiento de los marcadores, patrón de tiempo del movimiento
Nakamura et al, 2003	Videograbación de los movimientos faciales, medición en imágenes fijas	Medición de la anchura de la apertura del ojo durante diferentes movimientos de la boca

Tomat y Manktelow 2005	Videograbación, seguimiento de marcadores faciales alrededor de la boca en imágenes fijas	Análisis en 2D de los movimientos de la sonrisa, desplazamiento y ángulos de movimiento
Mishima et al, 2006	Sistema de captura de movimiento en video en 3D, sin marcadores faciales, cámaras infrarrojas	Análisis en 3D de los movimientos de la boca
Manktelow et al, 2008	Medición de puntos de referencia faciales especificados en reposo y en expresión facial máxima con regla de mano	Desplazamiento de marcadores durante movimiento; sin análisis multiregional y medición de dirección
Sawyer et al, 2008	Estereofotogrametría en 3D, puntos de referencia faciales superficiales	Análisis en 3D de la posición en descanso y sonrisa; medición de vectores, distancias y ángulos
Shu He et al, 2008	Medición de la simetría de los movimientos faciales en cada región facial.	Análisis de movimiento basado en patrones binarios locales de los dos lados de la cara
Junyu Dong et al, 2011	Proyección de integración para la detección de una región del rostro y sus bordes	Medición de la distancia de puntos localizados en los bordes de los ojos, cejas, nariz y boca
Anguraj y Padma 2012	Detección de bordes para obtener puntos de referencia que utiliza el algoritmo IECM.	Medición de distancias entre puntos del rostro
Limbeck et al, 2012	Seguimiento de marcadores faciales colocados en diferentes partes del rostro	Análisis en 2D de la posición de los marcadores y su tamaño
Azoulay et al, 2014	Detección de puntos de referencia faciales utilizando una cámara de celular	Medición de áreas triangulares y distancias lineales extraídas para la obtención de características
Katsumi et al, 2015	Medición de movimiento facial tridimensional	Análisis cinco movimientos faciales independientes y su simetría facial regional

Las herramientas objetivas de diagnóstico han sido desarrolladas para enfocarse en el movimiento de áreas específicas del rostro, como la frente, los ojos y los labios. Estas herramientas objetivas son más precisas que las herramientas subjetivas. Wu

et al, mostraron que los movimientos involuntarios que acompañan a un movimiento voluntario del rostro son diagnosticados significativamente más frecuente con herramientas de diagnóstico objetivas que con subjetivas [7].

El análisis tridimensional permite la evaluación más exacta de la función facial completa. Gross et al, encontraron que el análisis en 2D subestima amplitudes tridimensionales en casi un 43% [8]. Por lo tanto, el análisis tridimensional de la función facial es crucial en la evaluación pre-operatoria y post-operatoria de los movimientos faciales. Los sistemas de diagnóstico usando análisis de video tridimensional, como el sistema desarrollado por Frey, provee datos tanto cuantitativos como cualitativos en la función facial [9].

Este sistema consiste en una estructura con espejos, una rejilla de calibración, una videocámara digital y una estación de trabajo (*Workstation*). El procedimiento a grandes rasgos consiste en colocar 18 puntos de referencia en el rostro del paciente, posteriormente los sujetos son videograbados mientras realizan una serie de movimientos faciales. Las grabaciones resultantes son transferidas a una computadora, donde posteriormente se utilizan los softwares *Ulead®* y *Facialis®* (Laboratorio de Biomecánica del Instituto Federal de Suiza, Zúrich, Suiza) para seleccionar la secuencia de video más adecuada. Por último, se utiliza el software *FaciShow®* (Laboratorio de Biomecánica del Instituto Federal de Suiza, Zúrich, Suiza) que permite la visualización de los puntos de referencia en ambas perspectivas, (bidimensional o tridimensional) para cada movimiento individual, así como también la visualización de las trayectorias de los puntos de referencia en cada fase. Adicionalmente, todo el videoclip para cada momento puede ser revisado o la imagen de la cara del paciente, tanto en reposo como en el punto final de cada movimiento puede ser visto (ver Fig. 1.2 [6]).

El sistema fue completado e instalado por primera vez en el Laboratorio para Movimiento y Análisis de Imagen en la División de Cirugía Plástica y Reconstructiva, Departamento de Cirugía, Escuela de Medicina de Viena en enero de 1998. Desde entonces la evaluación de todos los pacientes tratados para parálisis facial se hace en este instituto [10].



Fig. 1.2 Captura de video del sistema de análisis de video tridimensional.

Aunque el proceso de análisis de estos sistemas es demasiado lento, se ha validado y probado su fiabilidad en la evaluación de la terapia para parálisis facial. Los objetivos a futuro del análisis de video tridimensional es distribuir este sistema para la evaluación de los movimientos faciales complejos, a través del cual se pretende crear un sistema internacional estandarizado de clasificación facial que se aplique fácilmente, sea eficiente y asequible en la rutina clínica diaria.

Si bien existe una diversidad de métodos y maneras de adquisición de datos, todos los sistemas comparten el uso de por lo menos una cámara profesional para la captura de imágenes que puedan ser posteriormente procesadas. Con el avance de la tecnología móvil, se han realizado recientemente aplicaciones que pretenden dar portabilidad a los sistemas ya existentes, sin embargo se ven gravemente afectadas por los problemas inherentes de los dispositivos móviles. En el presente trabajo, se optó por utilizar el *Kinect*®, ya que además de contar con una cámara RGB que obtiene imágenes a color, tiene un sensor de profundidad que funciona como un proyector de rayos infrarrojos; que combinado con un sensor CMOS monocromo, permite percibir el lugar en 3D en cualquier condición de luz ambiental. Esta característica adicional que posee el *Kinect*® en comparación a las cámaras fotográficas ordinarias usadas en otros sistemas, permite tener datos

complementarios del rostro que resultan útiles al momento de elaborar el sistema, además de brindar una solución ante el problema que representan las variabilidades.

CAPÍTULO 2: MARCO TEÓRICO

2.1 Introducción

El presente trabajo busca desarrollar sistema que brinde una valoración objetiva en el proceso de diagnóstico de pacientes con parálisis facial. En este sentido, es preciso aclarar algunos conceptos básicos en este capítulo. En primer lugar se definen los elementos que componen la topografía del Sistema Nervioso hasta llegar al nervio facial (ver Fig. 2.1). Posteriormente se describe la parálisis facial, comenzando por sus signos y síntomas, después su epidemiología, causas y terminando con su diagnóstico en donde se mencionan los métodos de clasificación más usados en los cuales se describe a grandes rasgos el método de Nottingham.

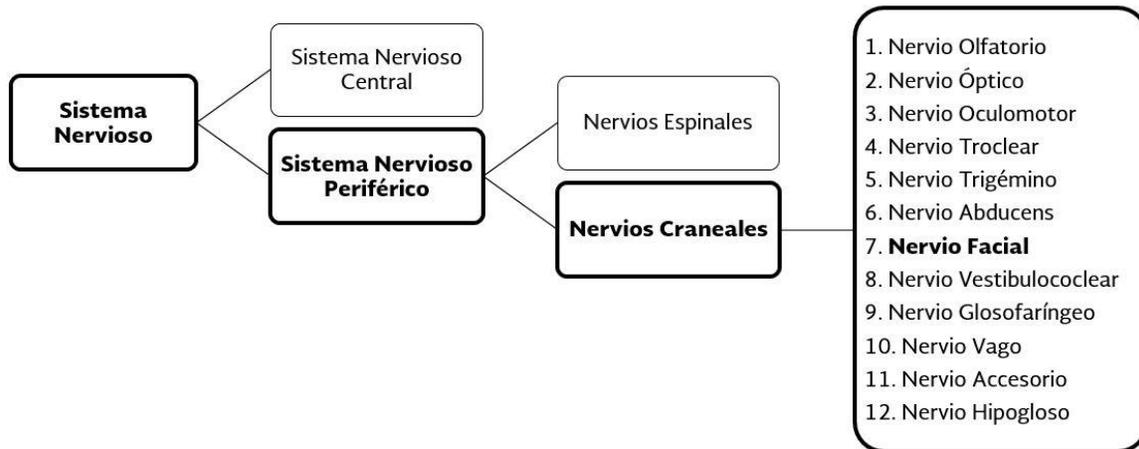


Fig. 2.1 Anatomía topográfica del Sistema Nervioso.

Una vez descritas todas las partes anatómicas involucradas y los aspectos clínicos de la parálisis facial, se ofrece una visión detallada de las principales tecnologías empleadas en la elaboración del trabajo, comenzando por el *Kinect*®. Al inicio se mencionan sus especificaciones técnicas y la utilidad que tienen en el trabajo, después se detalla la interfaz de seguimiento, en donde se describe el uso de cada uno de los componentes y variables que se utilizan. Para finalizar este capítulo, se indican los lenguajes de programación que se usaron para desarrollar el sistema y se enlistan sus ventajas y desventajas que presentan frente a otras posibles alternativas en el caso de haberlas.

2.2 Topografía del Sistema Nervioso

El Sistema Nervioso, es el más completo, complejo y desconocido de todos los sistemas que conforman el cuerpo humano, asegura junto con el Sistema Endócrino, las funciones del control del organismo. Es capaz de recibir e integrar innumerables datos procedentes de los distintos órganos sensoriales para lograr una respuesta del cuerpo, además en él se encuentran los centros de la consciencia, del pensamiento, de las emociones y de todas las cualidades humanas consideradas como superiores.

El Sistema Nervioso tiene tres funciones básicas: *la sensitiva, la integradora y la motora*. La función sensitiva le permite reaccionar ante estímulos provenientes tanto desde el interior del organismo como desde el medio exterior. Luego, la información sensitiva se analiza, se almacenan algunos aspectos de ésta y toma decisiones con respecto a la conducta a seguir; esta es la función integradora. Por último, puede responder a los estímulos iniciando contracciones musculares o secreciones glandulares; esta es la función motora.

Su constitución anatómica es muy compleja, y las células que lo componen, a diferencia de las del resto del organismo, carecen de capacidad regenerativa; y para su estudio puede dividirse en: *Sistema Nervioso Central (SNC)* y *Sistema Nervioso Periférico (SNP)*. Sin embargo, las dos partes del Sistema Nervioso trabajan en conjunto y son esenciales para la vida. Sin el SNP, el SNC no tendría ninguna entrada sensorial para procesar, por lo que le sería imposible reaccionar al medio ambiente. Del mismo modo, el SNP se basa en el SNC para coordinar la información procedente de diferentes partes del cuerpo y tomar decisiones acerca de cómo se debe reaccionar en una situación dada. Estas dos partes del Sistema Nervioso trabajan juntas y en sincronía para crear una experiencia consciente de nuestra vida cotidiana [11].

2.2.1 Sistema Nervioso Periférico (SNP)

El SNP está formado por nervios y neuronas que residen o se extienden fuera del SNC, hacia los miembros y órganos del cuerpo. A diferencia del SNC, el SNP no cuenta con la protección de estructuras óseas. Su función principal es conectar los

estímulos externos, internos y propioceptivos (sentir la disposición de los diferentes órganos) con el SNC, y éste a su vez con la periferia, para controlar órganos, aparatos o sistemas. Los nervios que componen el SNP, en definitiva, permiten la comunicación entre el SNC y el entorno interno o externo. De acuerdo al tipo de fibras, se puede hablar de nervios motores, nervios sensitivos y nervios mixtos. Respecto a la zona del SNC de la que proceden, por otra parte, los nervios pueden distinguirse entre *nervios craneales* y *nervios raquídeos*. Los craneales son aquellos que parten o llegan al encéfalo, mientras que los raquídeos, también denominados espinales, salen o llegan a la médula espinal. A su vez, aquellos que llevan información al resto del organismo se denominan eferentes (motores) y aquellos que traen información desde la periferia hasta el sistema nervioso central se denominan aferentes (sensitivos) [12].

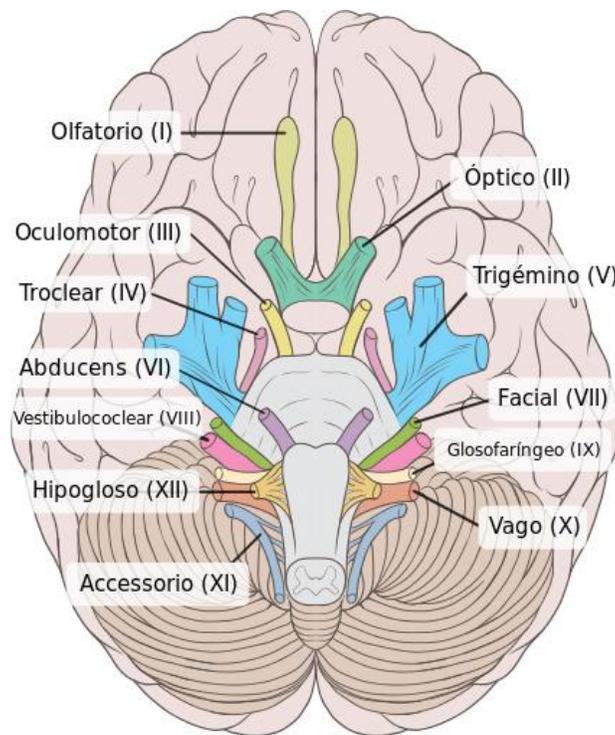
Nervios craneales

Los nervios craneales, también llamados pares craneales, son doce pares de nervios que emergen del cerebro o a nivel del troncoencefálico (ver Fig. 2.2). Estos nervios abandonan el cráneo a través de fisuras y forámenes para distribuirse en la cabeza y cuello principalmente (a excepción del décimo nervio craneal que inerva estructuras torácicas y abdominales).

1. Nervio Olfatorio (I par craneal): se origina en la mucosa olfatoria, cruza los agujeros de la lámina cribosa del etmoides y termina en el bulbo olfatorio. Es un nervio puramente sensorial y su función es la olfacción.
2. Nervio Óptico (II par craneal): se origina en las fibras que provienen de la retina, cruza el agujero óptico de la órbita y termina en el quiasma óptico. Es un nervio sensorial y su función en la visión.
3. Nervio Oculomotor (III par craneal): es un nervio mixto aunque principalmente motor. La función motora somática permite el movimiento del párpado y determinados movimientos del globo ocular. La actividad motora parasimpática condiciona la acomodación del cristalino y la constricción de la pupila o miosis.

4. Nervio Troclear (IV par craneal): es un nervio mixto aunque principalmente motor, cuya función motora permite el movimiento del globo ocular.
5. Nervio Trigémino (V par craneal): es un nervio mixto. La porción sensitiva transmite las sensaciones de tacto, dolor, temperatura y propiocepción de la cara. La porción motora inerva los músculos de la masticación.
6. Nervio Abducens (VI par craneal): es un nervio mixto aunque principalmente motor, cuya función motora permite movimientos del globo acular.
7. Nervio Facial (VII par craneal): es un nervio mixto. La porción sensitiva transporta la sensibilidad gustativa de los dos tercios anteriores de la lengua. La porción motora somática inerva la musculatura de la mímica facial. La porción motora parasimpática inerva las glándulas salivales y lagrimales.
8. Nervio Vestibulococlear (VIII par craneal): es un nervio mixto, principalmente sensorial. La función principal es transportar los impulsos sensoriales del equilibrio y la audición.
9. Nervio Glossofaríngeo (IX par craneal): es un nervio mixto. La porción sensorial transporta la sensibilidad gustativa del tercio posterior de la lengua. La porción motora somática inerva la musculatura que permita la elevación de la faringe durante la deglución. La porción motora parasimpática inerva la glándula parótida.
10. Nervio Vago (X par craneal): es un nervio mixto. La función sensorial transporta la sensibilidad de la epiglotis, faringe, así como estímulos que permiten el control de la presión arterial y la función respiratoria. La porción motora somática inerva los músculos de la garganta y cuello permitiendo la deglución, tos y la fonación. La porción motora parasimpática inerva la musculatura lisa de los órganos digestivos, el miocardio y las glándulas del tubo digestivo.
11. Nervio Accesorio (XI par craneal): es un nervio mixto principalmente motor que inerva músculos deglutorios, el músculo trapecio y el músculo esternocleidomastoideo.
12. Nervio Hipogloso (XII par craneal): inerva la musculatura lingual.

Para algunos autores ni el nervio olfatorio ni el óptico deberían incluirse en el SNP, ya que serían extensiones del cerebro, mientras que el resto emerge del troncoencefálico. Algunos pares craneales sólo contienen fibras aferentes, otros sólo eferentes y algunos ambos tipos de fibras. La Nomenclatura Anatómica Internacional incluye al nervio terminal como nervio craneal, a pesar de ser atrófico en los humanos y estar estrechamente relacionado con el nervio olfatorio. Los pares craneales tienen un origen aparente que es el lugar donde el mismo sale o entra en el encéfalo. El origen real es distinto de acuerdo a la función que cumplan [13].



© (2013) PATRICK J.

Fig. 2.2 Vista inferior de un cerebro humano y tronco cerebral mostrando los pares craneales.

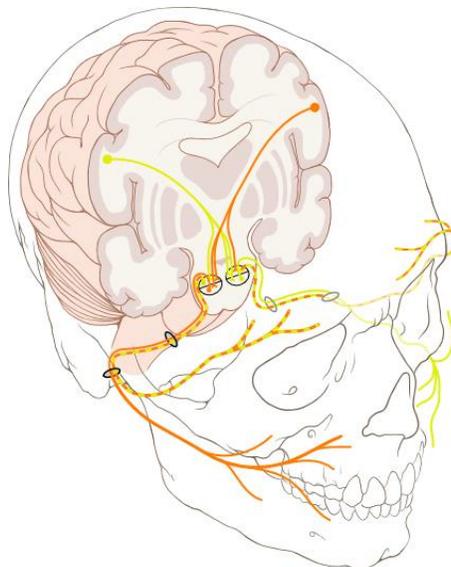
Se pueden distribuir los pares craneales en 4 grupos:

- Cerebro medio: I y II pares.
- Mesencéfalo: III y IV pares.
- Protuberancia: V, VI, VII y VIII pares.
- Bulbo: IX, X, XI, XII pares

Esta disposición permite localizar topográficamente una posible lesión en el sistema nervioso, según la semiología derivada de la afectación de un determinado par craneal.

Nervio Facial

El nervio facial es un nervio craneal mixto, es decir, contiene fibras sensitivas, sensoriales y motoras. Por ser un par craneal, consiste en dos fibras nerviosas, el nervio facial propiamente dicho y el nervio intermedio o intermediario de Wrisberg (ver Fig. 2.3). Una corre por el lado derecho de la cara y el contralateral por la izquierda. Parte del tallo cerebral, justo entre el puente troncoencefálico y el bulbo raquídeo; sus fibras rodean el núcleo del VI par craneal y salen del neuroeje a nivel del surco bulboprotuberencial, en el llamado ángulo pontocerebeloso, junto con el nervio vestibulococlear (VIII par craneal) también llamado nervio auditivo. Penetra el hueso temporal por el conducto auditivo interno y sigue por el acueducto de Falopio para salir por el agujero estilomastoideo; atraviesa la glándula parótida donde se divide en dos ramas terminales: la temporofacial y la cervicofacial, que se distribuyen en todos los músculos de la cara y el cutáneo del cuello, a excepción del elevador del párpado superior que es inervado por el III par craneal. La rama temporofacial inerva los músculos frontal y orbicular [14].



© (2006) PATRICK J.

Fig. 2.3 Ilustración del origen y trayecto del nervio facial derecho e izquierdo y su recorrido.

El nervio facial está compuesto por fibras motoras y sensitivas que se originan en tres grupos nucleares: núcleo motor, núcleos parasimpáticos (salival superior y lacrimal) y núcleo sensitivo.

1. El núcleo motor proporciona la inervación a los músculos de la expresión facial, los músculos auriculares, el músculo estapedio, el vientre posterior del digástrico, y el músculo estilogloideo.
2. El núcleo salival superior proporciona la inervación a las glándulas salivales submandibular y sublingual, y a las glándulas palatinas y nasales.
3. El núcleo lacrimal proporciona la inervación para las glándulas lacrimales.
4. El núcleo sensitivo recibe aferencias gustativas de los dos tercios anteriores de la lengua.

Para explorar la parte motora se pide al paciente que realice movimientos faciales (arrugar la frente, cerrar los ojos, fruncir los labios, sonreír o soplar). La musculatura facial inferior puede observarse en los pacientes comatosos observando el gesto de dolor al presionar firmemente en la apófisis estiloides. El gusto puede explorarse utilizando sustancias con sabor dulce, amargo, ácido y salado aplicados en cada mitad de los dos tercios anteriores de la lengua. La disfunción sensitiva se localizará en la zona postero-superior del conducto auditivo externo y en la mitad superior del pabellón auricular [15].

2.3 Parálisis Facial

La parálisis facial es la debilidad o parálisis de la musculatura inervada por el nervio facial (VII par craneal) debida a una lesión del VII par craneal en cualquier lugar de su recorrido desde el núcleo de origen a las estructuras que inerva. La lesión impide el movimiento normal de diferentes músculos como: frente, párpados y boca; así como, la expresión de las emociones. La parálisis de los músculos faciales causa asimetría, ocasionando problemas psicológicos y emocionales.

Existen dos tipos de parálisis facial: *Parálisis Facial Periférica* (PFP) y *Parálisis Facial Central* (PFC) (ver Fig. 2.4). La PFC es causada por afección de la neurona motora superior, y se acompaña con frecuencia de hemiplejía del mismo lado de la lesión. Afectando al movimiento voluntario de la parte inferior de la parte inferior de

la cara, pero no a los músculos frontal y orbicular de los párpados. Además, suelen ser normales los movimientos faciales reactivos a emociones. La PFP, también es conocida como parálisis de Bell, de neurona motora inferior o parálisis idiopática; es causada por la afección aguda del nervio facial a nivel periférico; origina pérdida del movimiento voluntario en todos los músculos faciales del lado afectado y es la más común de las parálisis. La PFP puede ser unilateral o bilateral, completa si abarca todos los músculos del lado afectado, o parcial si sólo afecta a un grupo de ellos. La forma de parálisis más común es unilateral, afectando la hemicara del nervio lesionado [16].

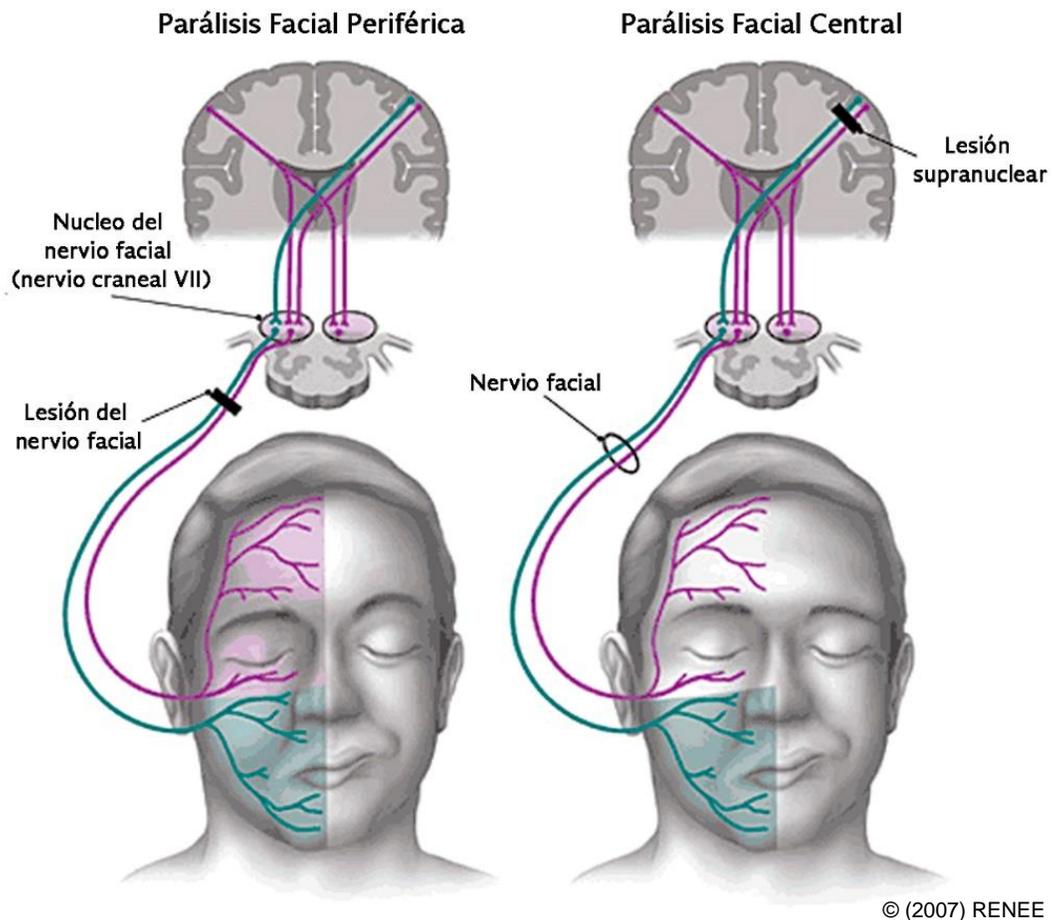


Fig. 2.4 Clasificación de la parálisis facial.

2.3.1 Signos y síntomas

Los signos y síntomas de la parálisis facial dependen del lugar de la lesión en el nervio facial. A medida que la lesión es más periférica se puede observar pérdida de movimiento del lado afectado de la cara. Los pliegues y arrugas de la hemicara están ausentes o indefinidos. El párpado superior, inferior, mejillas y comisura de la boca se ven caídas. Se presenta un incremento en la sensibilidad a los cambios de temperatura y hay una mayor secreción de saliva y lágrimas. Hay problemas visuales a consecuencia del trastorno de movilidad o cierre palpebral, tales como visión borrosa o incompleta. También existe una exposición de la córnea a causa de la dificultad para cerrar los ojos; así como, pérdida del reflejo de parpadeo.

La asimetría ocasionada por la parálisis se evidencia más cuando el paciente realiza los siguientes movimientos: formar arrugas horizontales en la frente al elevar las cejas, fruncir el entrecejo, arrugar la nariz, apretar labios y abocinar la boca para soplar o silbar. Asimismo, se observará asimetría al sonreír y hablar; en ocasiones la voz adquiere un tono nasal. Los líquidos ingeridos se pueden derramar, los alimentos sólidos se quedan acumulados entre la encía y el maxilar, y pueden morderse las paredes de las mejillas al masticar por falta de sensibilidad y tono muscular.

Una alteración que se presenta en la audición es la hiperacusia, la cual se define como hipersensibilidad a los sonidos cotidianos, percibidos como insoportables, fuertes o dolorosos. Aunque la hipersensibilidad se puede presentar como un daño a nivel central, en el caso de la parálisis facial se presenta si la lesión es proximal al nervio del estapedio, tras la disminución o ausencia del reflejo estapedial. Dicho reflejo se presenta con frecuencia, como un mecanismo de protección contra sonidos fuertes, mediante la contracción bilateral del músculo del estapedio en el oído medio, lo cual ocurre alrededor de los 85 decibeles.

Otras características de la parálisis facial periférica son la presencia de otitis, dolor cercano al mastoide (50% de los casos), formación excesiva de lágrimas, pérdida sensorial en la piel y músculo, y pérdida del sabor en la mitad de la lengua (33% de los casos). Cuando se presenta comorbilidad entre la parálisis facial periférica y

diabetes *mellitus*, los pacientes presentan con mayor frecuencia sintomatología motora, mientras que las neuropatías sensoriales son menos frecuentes [16].

2.3.2 Epidemiología

La parálisis facial periférica tiene una incidencia epidémica, lo que podría sugerir un mecanismo etiopatogénico de tipo infeccioso o inflamatorio. Su incidencia es de 20 a 30 casos por 100 mil personas al año. Se reporta que la frecuencia de padecer diabetes *mellitus* y parálisis facial es de un 0.4 a un 2.5% aunque no se ha reportado una relación causa-efecto entre estos dos padecimientos [17].

En general, los datos estadísticos indican que el trastorno afecta de igual forma a hombres y mujeres, en cualquiera de los dos lados de la cara. Ha sido descrita en todas las edades, aunque en un estudio se reportan picos de incidencia entre los 20 a 29 años y entre los 50 a 59 años, mientras que otra investigación indica que el pico de incidencia es a los 40 años [18]. Existe la posibilidad de que haya más de un pico de edad para la incidencia de la parálisis facial debido a que en general, se trata de un padecimiento transitorio y puede estar determinado por distintas causas de acuerdo a la edad. La prevalencia en adultos mayores aumenta debido a que la recuperación es más lenta y puede llegar a ser crónica.

La mujer en edad reproductiva es afectada de dos a cuatro veces más que el hombre de la misma edad, y la mujer embarazada 3.3 veces más que la que no lo está. Algunos estudios reportan casos donde se presenta una parálisis facial en el embarazo, principalmente durante el tercer trimestre de gestación. Hay pocos casos en donde se presenta durante el primer trimestre. Asimismo, en el 22% de los casos de mujeres que presentaron dicho trastorno también presentaron preeclampsia, por lo que se sospecha que la parálisis durante el embarazo puede ser predictor de preeclampsia [19].

El riesgo de padecer parálisis de Bell parece ser estacional, siendo más común en invierno. Esto puede deberse a los cambios de temperatura que afectan al sistema inmune, haciéndolo vulnerable ante padecimientos víricos o infecciosos.

2.3.3 Causas

Cuando la parálisis facial periférica no tiene una causa aparente se le conoce como parálisis facial primaria o idiopática (parálisis de Bell), mientras que al tener una causa detectable, se le conoce como parálisis secundaria. Se ha descrito que la causa más común de la parálisis facial periférica es una infección viral sistemática del virus herpes simple.

Dentro de las principales causas de parálisis facial secundaria se encuentran: traumatismos, problemas vasculares como aneurisma de la arteria basilar, lesiones vasculares pontinas o hipertensión; así como tumores, procesos inflamatorios óticos y reacciones a drogas o medicamentos como la isoniazida. También se deben a cuadros neurológicos como polirradiculoneuropatías periféricas tipo Guillain-Barré, multineuritis, sífilis, lepra, cuadros meníngeo-infecciosos o encefalitis; asociadas a padecimientos como síndrome de Ramsay-Hunt, de Heerfordt, de Melkersson-Rosenthal o de Stevens-Johnson; enfermedades óseas; por ejemplo la enfermedad de Paget; enfermedades granulomatosas y del tejido conectivo como sarcoidosis; malformaciones congénitas de oído, problemas endócrinos como hipotiroidismo, esclerosis múltiple y en menores casos VIH [16].

La parálisis facial es un síntoma común que acompaña la enfermedad de Lyme (borreliosis) en los niños. En México su incidencia es poco frecuente; ya que no es una zona endémica de la enfermedad. Se ha encontrado que la parálisis facial puede aparecer después de la extracción dental, siendo muy baja su incidencia. Se atribuye esta afección a la anestesia, la cual inflama el nervio facial causándole degeneración, aunque también puede deberse a la infección por parte de virus contenido en la saliva [20].

2.3.4 Diagnóstico

La orientación diagnóstica incluye una cuidadosa evaluación clínica y exploración física. Partiendo de estas evaluaciones se realizan una serie de estudios que tienen como objetivo identificar la causa de la parálisis, los cuales son: análisis de sangre del fluido cerebro-espinal, rayos X, estudios de neuroimagen de conducción del nervio, electromiografía y electroneurografía. Estos estudios se solicitan de acuerdo

a sospecha clínica posterior a la evaluación clínica de signos y síntomas empleando algún método de diagnóstico.

Métodos de diagnóstico

Los métodos usados actualmente para clasificar el nivel de la función del nervio facial se dividen en dos enfoques: *tradicionales* o *informáticos*. El término “enfoque tradicional” es usado para referirse a los métodos que no requieren ningún equipo de cómputo especializado. Una gran parte de estos enfoques se basan en la utilización de las evaluaciones subjetivas del observador. Por otra parte en los métodos con un “enfoque informático” se implica el uso de equipo especializado para medir y cuantificar los datos digitales de manera objetiva. El mayor inconveniente con este enfoque es el costo y la operatividad.

Enfoques tradicionales

Los enfoques tradicionales para la evaluación de la función facial han evolucionado considerablemente durante el último medio siglo. Los informes iniciales en la literatura eran de carácter descriptivo. En 1970, May preparó un análisis detallado sobre el diagnóstico de parálisis facial y desarrollo un sistema subjetivo basado en 10 criterios relevantes para el tono facial y la expresión [21].

En 1983, John House revisó ocho sistemas de graduación y evaluó los méritos de cada uno de ellos. Su sistema de clasificación sigue siendo un marco conceptual útil. House clasificó las escalas del nervio facial como *brutas*, *regionales* o *específicas*. Las escalas brutas hacen una evaluación general de la función motora facial.

Algunos de los primeros ejemplos de escala bruta se muestran en la Tabla 2.1 y en la Tabla 2.2 [22]. Las escalas brutas de la función facial evalúan la función general y asignan un grado que refleje la severidad de la parálisis y los efectos secundarios simultáneamente. Estas escalas brutas son descriptivas y como resultado no pueden ser manipuladas matemáticamente [23].

Tabla 2.1 Escala de Botman y Jongkees.

Clase	Significado
0	Actividad facial normal
I	Parálisis ligera: normal en reposo, normal al hablar, los ojos se pueden cerrar, alguna asimetría al reír y silbar
II	Parálisis moderada: normal en reposo, asimetría al hablar y reír, los ojos no se pueden cerrar.
III	Parálisis severa: asimetría en reposo, disfunción en movimientos
IV	Parálisis total: sin tono, pérdida total de la función. Contractura de los músculos puede resultar en la mejora aparente, y la degeneración atrófica puede causar un defecto más grave

Tabla 2.2 Escala de May.

Grado de la lesión	Grado	Etiología	Valoración
Normal	I	Lesiones incompletas ¹	Excelente recuperación; movimiento individual adecuado, cierre de los ojos, sonrisa.
Leve (apenas perceptible)	II	Lesiones incompletas ¹	Movimiento apropiado, pero el movimiento de masas presente en los ojos o la boca.
Moderado (diferencia obvia)	III	Nervio seccionado (después de la reparación) ²	Algunos movimientos y simetría razonable pero incapacidad para cerrar completamente los ojos, incompetencia oral.
Grave (debilidad incapacitante, espasmo, sincinesia)	IV	Nervio seccionado (después de la reparación) ²	Simetría facial pobre.
Sin movimiento	V	Nervio seccionado	Cara flácida.

¹Parálisis de Bell, herpes zoster ótico, otitis media, fractura del hueso temporal.

²Tumor, trauma, etc.

Las escalas regionales en contraste, atribuyen puntajes independientes a diferentes áreas de la función facial, a veces con ponderación para reflejar la mayor o menor importancia de determinadas áreas de la cara, como el movimiento de cierre de los ojos o la boca. En los sistemas regionales, las puntuaciones de la frente pueden tener un inferior impacto debido a un menor impacto en la estética, la incapacidad funcional relativamente menos asociada, o el pronóstico más favorable para la recuperación de la función. Un ejemplo de la escala regional se muestra en la Tabla 2.3 [22].

Tabla 2.3 Escala de Smith.

Escala	0 ninguna función I – 0% a 25% de función II – 25% a 50% de función III – 50% a 75% de función IV – 75% a 100% de función
Áreas a graduar	0 a IV reposo 0 a IV frente y arrugas 0 a IV ojo 0 a IV boca

Cada área es graduada como un porcentaje. El porcentaje es después convertido en un número entre 0 (ninguna función) y 4 (normal). La puntuación para cada área es sumada y después dividida por 4, dando una puntuación máxima de 4.

Las escalas específicas piden al observador responder “sí” o “no” a preguntas sobre áreas específicas de la cara. La escala Stennert mostrada en las Tabla 2.4 y Tabla 2.5 [24] es un ejemplo de escala específica.

Tabla 2.4 Puntuación de Parálisis Facial de Stennert.

Parámetro que se evalúa	Valoración
Tono en reposo	
Diferencia entre fisuras palpebrales	<input type="checkbox"/> 3mm y más
Perdida del surco nasolabial (si está presente del lado normal)	<input type="checkbox"/> Sí
Caída del ángulo de la boca	<input type="checkbox"/> Sí
Motilidad	
Ceño fruncido (al menos el 50% del lado normal)	<input type="checkbox"/> No es posible
Cierre palpebral incompleto	
Inervación Leve (como en el sueño)	<input type="checkbox"/> Sí
Inervación Máxima	<input type="checkbox"/> Sí
Exposición de dientes	
Dientes caninos superiores e inferiores	<input type="checkbox"/> No visible
Incisivo segundo superior (todo el ancho)	<input type="checkbox"/> No visible
Silbando (disminución de la distancia entre el philtrum y el ángulo de la boca en comparación con el lado normal)	<input type="checkbox"/> Menos del 50%

El número de casillas marcadas por 10 es igual a la puntuación de parálisis para una puntuación total posible de 100%. La escala es de doble ponderación (La cara en reposo se cuenta con 40% y en movimiento con 60%)

Tabla 2.5 Puntuación de Parálisis Facial (Defecto Secundario) de Stennert.

Parámetro que se evalúa	Valoración
Hiperacusia	No anotada en esta escala
Gustación dañada	No anotada en esta escala
Sincinesia en la frente, ojos, surco nasolabial, esquina de la boca o la barbilla	
Presente en al menos un área	<input type="checkbox"/> Sí
Presente en tres o más áreas	<input type="checkbox"/> Sí
Espasmo	
	<input type="checkbox"/> Presente
	<input type="checkbox"/> Fuertemente presente
	<input type="checkbox"/> Inconveniente
Lagrimeo	
	<input type="checkbox"/> Menos del 30%
	<input type="checkbox"/> Menos del 30% + cierre de la tapa incompleta
	<input type="checkbox"/> 0%
Contracturas	<input type="checkbox"/> Presente
Lágrimas de cocodrilo	<input type="checkbox"/> Presente

El número de casillas marcadas por 10 es igual a la puntuación de parálisis para una puntuación total posible de 100%.

La escala es de doble ponderación (La cara en reposo se cuenta con 40% y en movimiento con 60%)

La consideración de las ventajas y desventajas de cada uno de estos primeros sistemas, proporcionaron la base para una escala bruta que House introdujo en 1983. Esta escala bruta pronto fue modificada por Brackmann y Barrs para dar lugar a los actuales seis niveles de la escala House-Brackmann [25]. Esta escala, la cual fue formalmente adoptada como el estándar universal de la Academia Americana de Otorrinolaringología y Cirugía de Cabeza y Cuello en recomendación del Comité de Trastornos del Nervio Facial, procuraba agrupar pacientes con similares grados de discapacidad. La escala de House-Brackmann es técnicamente un sistema “bruto” en el que se requiere al observador para dar un nivel que refleje el grado de parálisis y el déficit secundario de una vez. Esta escala, sin embargo, incorpora algún nivel de ponderación dependiendo de la región de un déficit dado.

La escala de House-Brackmann, aunque resulta útil en la normalización de los informes de la parálisis del nervio facial, es imperfecta. En primer lugar, debido a la amplia gama de movimientos faciales comprendida en diferentes grados, no puede

distinguir los cambios sutiles en la función facial. Como resultado, la diferenciación entre los mejores y peores resultados de procedimientos de injerto de nervio puede ser imperfecta para discernir. En segundo lugar, los pacientes pueden tener disfunción en una región facial que encaja en un grado, sin embargo, ajustarse en un grado diferente en otra región. Este aspecto puede dar lugar a inconsistencia en la puntuación [26]. Informes anteriores han señalado que la escala House-Brackmann tiene limitaciones significativas de concordancia entre observadores. Estas preocupaciones han llevado a la propuesta de sistemas adicionales.

El primero de ellos fue el sistema Burres-Fisch. El sistema Burres-Fisch es un sistema objetivo sobre la base de mediciones lineales del desplazamiento de puntos en la cara de referencia. Las distancias se midieron con calibradores de mano, así como con electrodos de superficie que producían registros electromiográficos. Este enfoque proporciona un índice de medición lineal que es una escala graduada continua. El sistema Burres-Fisch fue diseñado para minimizar el prejuicio del observador y la subjetividad. Entre las limitaciones de este enfoque están la incapacidad para hacer simultáneamente registros de diferentes regiones faciales, la falta de evaluación de defectos secundarios y la más importante, la excepcionalmente ardua y lenta tarea de medir y realizar los cálculos requeridos [27].

Dos enfoques adicionales son el sistema de clasificación Nottingham y la escala Sunnybrook. El sistema Nottingham mantuvo el concepto de Burres-Fisch de medición objetiva pero permitió una evaluación más rápida y agregó criterios para defectos secundarios. Las diferencias en las mediciones de los puntos de referencia mostrados en la Fig. 2.5 en un estado de reposo y tras el máximo esfuerzo son sumadas y utilizadas para generar un porcentaje cuando se compara contra el lado contralateral. La presencia de defectos secundarios está marcada por separado. Esta escala ha sido identificada por varios investigadores como posiblemente ser la más fácil y reproducible de todas, por esta razón fue elegida como el sistema de clasificación del nivel de parálisis facial para este trabajo. Las limitaciones de este

método incluyen la incapacidad para evaluar la disfunción del nervio facial bilateral y el uso de sólo modificadores descriptivos para describir defectos secundarios [28].

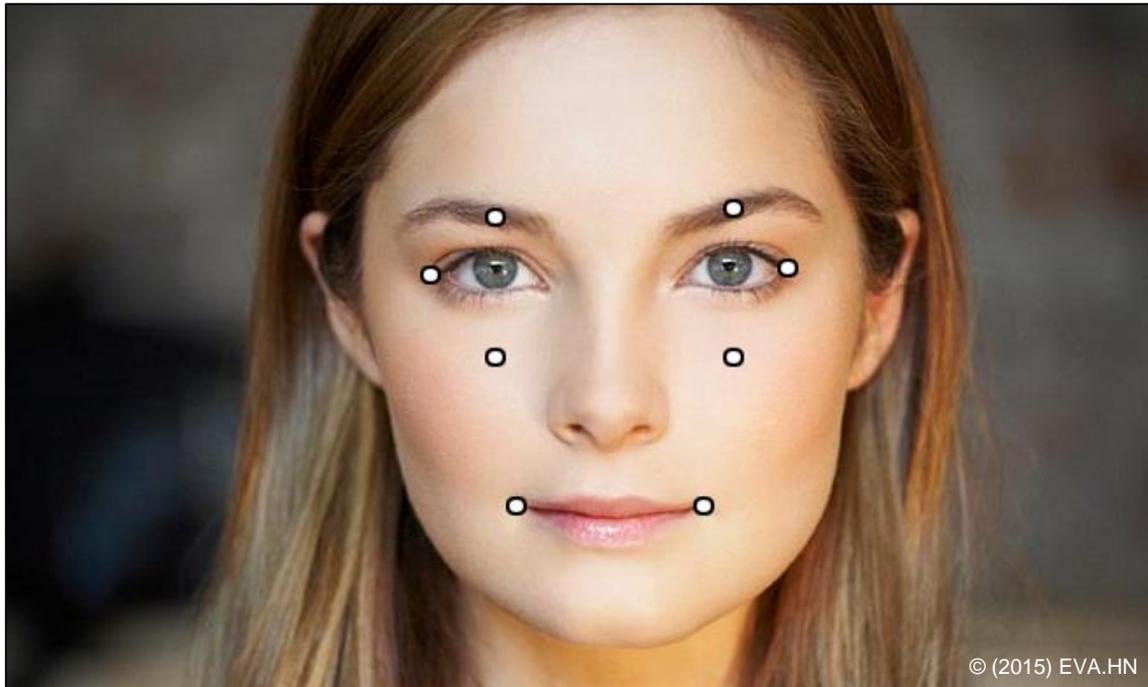


Fig. 2.5 Puntos de referencia facial designados en el método Nottingham.

La escala de Sunnybrook (ver Tabla 2.6), propuesto por Ross, es una escala subjetiva ponderada que incorpora defectos primarios y secundarios en el cálculo de una única puntuación. Los autores informaron que este enfoque era capaz de distinguir entre los niveles de disfunción del nervio facial no discernibles con la escala House-Brackmann. Sin embargo, este enfoque sigue siendo en gran medida una escala subjetiva y entre los defectos secundarios considera sólo la sincinesia [29].

Tabla 2.6 Sistema de Clasificación de Parálisis Facial Sunnybrook (Ross et al).

Parámetro	Descubrimiento	Valor del Punto
Puntuación de simetría al descansar ¹		
Ojo	Normal o anormal	0 o 1 punto

Mejilla (pliegue nasolabial)	Normal, alterado o ausente	0 o 1 o 2 puntos
Boca (caída o tirada)	Normal o anormal	0 o 1 punto
Puntuación de movimiento voluntario ²		
Arrugas de la frente	Sin movimiento a normal	0 a 5 puntos
Cierre de los ojos suave	Sin movimiento a normal	0 a 5 puntos
Sonrisa con boca abierta	Sin movimiento a normal	0 a 5 puntos
Gruñido	Sin movimiento a normal	0 a 5 puntos
Fruncido de labios	Sin movimiento a normal	0 a 5 puntos
Puntuación de sincinesia ³		
Arrugas de la frente	Ninguna, media, moderada o severa	0 a 4 puntos
Cierre de los ojos suave	Ninguna, media, moderada o severa	0 a 4 puntos
Sonrisa con boca abierta	Ninguna, media, moderada o severa	0 a 4 puntos
Gruñido	Ninguna, media, moderada o severa	0 a 4 puntos
Fruncido de labios	Ninguna, media, moderada o severa	0 a 4 puntos

Cálculo: (Suma de los puntos de simetría en reposo x 5) – (suma de los puntos de movimiento voluntario x 4) – (suma de los puntos de sincinesia x 1) = puntuación final (máximo de 100).

¹Puntuación de simetría al descansar compara el lado afectado con el lado normal.

²Puntuación de movimiento voluntario evalúa el nivel de excursión muscular comparado con el lado normal.

³Puntuación de sincinesia evalúa las contracciones musculares asociadas con cada expresión.

2.4 Kinect®

El *Kinect*® es un dispositivo físico creado por Alex Kipman y desarrollado por *Microsoft*® para la consola *Xbox 360*, fue lanzado en noviembre del 2010 y desde febrero del 2012 para PC a través de *Windows 7* y *Windows 8*. *Kinect*®. Permite a los usuarios controlar e interactuar con la consola sin necesidad de tener contacto físico con un control de videojuegos tradicional, mediante una interfaz natural de usuario que reconoce gestos, comandos de voz, objetos e imágenes.

Los componentes principales del *Kinect*® son indicados en la Fig. 2.6 [30], ellos son:

1. Micrófono multidireccional: Es un arreglo de cuatro micrófonos que pueden aislar la voz del usuario del ruido en la habitación. Mediante la comparación de la demora en cada micrófono, la fuente de la voz puede ser localizada.
2. Emisor de láser infrarrojo: Emite activamente un espectro infrarrojo cercano, el cual puede ser distorsionado por una superficie irregular y luego formado aleatoriamente como manchas reflejadas. Las manchas pueden ser recibidas por el sensor de profundidad (Número 3 en la figura).

3. Sensor de profundidad (cámara infrarroja): Captura la señal de infrarrojos que se puede convertir en un mapa de profundidad.
4. Inclinador motorizado: El motor puede ser programado con el fin de lograr el mejor ángulo de visión.
5. Cable USB: Transmite el flujo de video, de profundidad y de audio. La fuente de alimentación adicional debe conectarse para obtener todas las funciones del *Kinect*®. (La fuente de poder del *Kinect*® es de 12W mientras que la fuente de poder de un cable USB normal es de 2.5W).
6. Cámara RGB: Captura el flujo de video de color.

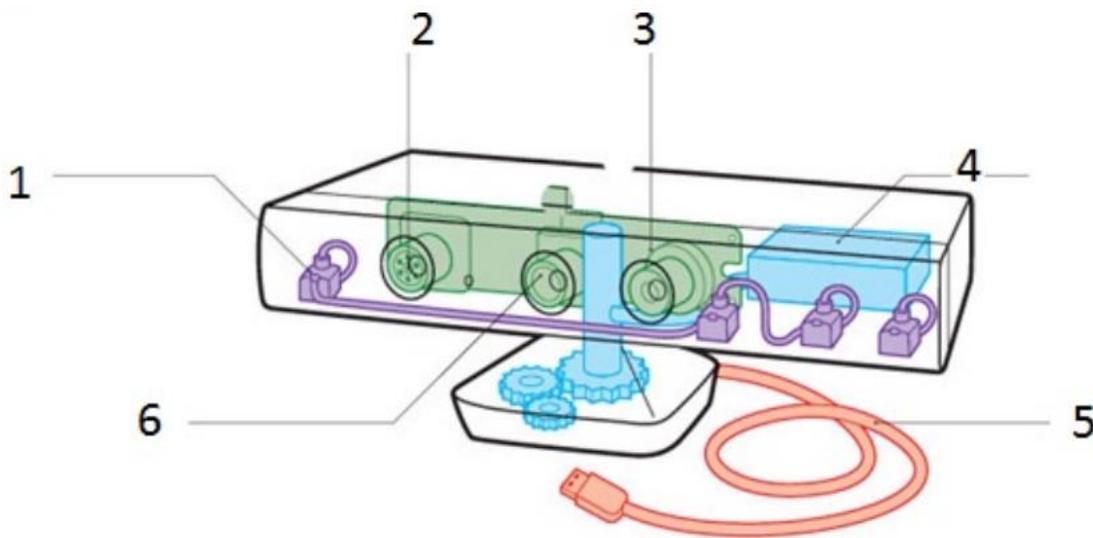


Fig. 2.6 Estructura del Kinect® y sus componentes principales.

2.4.1 Especificaciones técnicas

Sensor de profundidad

El sensor de profundidad permite tener una visión 3D del entorno de *Kinect*®. El emisor genera rayos infrarrojos y el sensor detecta los rayos que rebotan en los objetos o personas de la habitación. Los rayos reflejados son convertidos en información de profundidad para poder medir la distancia entre un objeto y el sensor, lo que permite generar la imagen de profundidad. Esto es especialmente relevante para el trabajo, al hacer posible conocer con precisión la distancia de los diferentes

puntos del rostro a la cámara. El sensor de profundidad cuenta con un rango válido que va desde 0.8 hasta 4 metros de distancia, tal como se muestra en la Fig. 2.7.

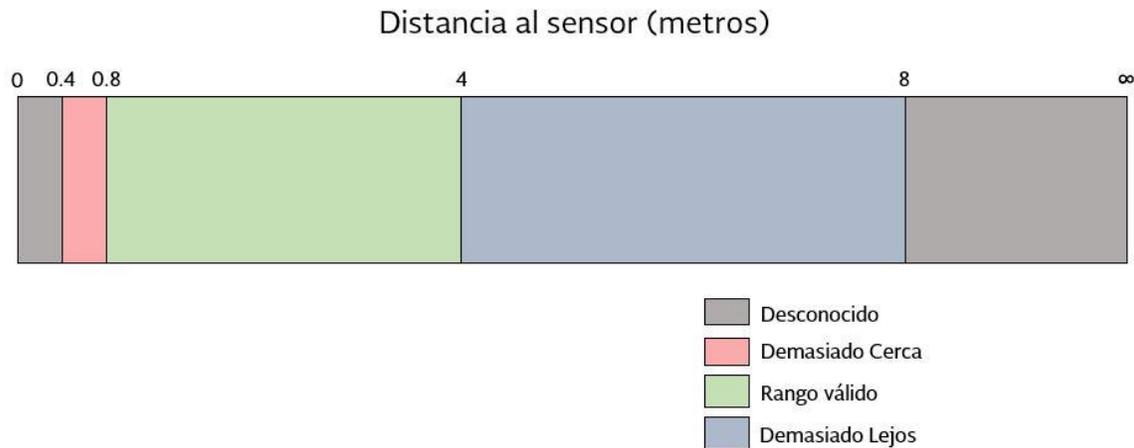


Fig. 2.7 Rango válido del sensor de profundidad.

En una imagen de profundidad, el color de cada píxel nos dirá qué tan lejos está esa parte de la imagen de la cámara. Puesto que estos valores corresponden directamente al lugar donde se encuentran los objetos en el espacio, estos son mucho más útiles en la determinación para saber en dónde empieza un objeto, en donde acaba otro y si hay alguna persona alrededor.

También, debido a la forma en que el *Kinect*® crea la imagen profunda, esta no es sensible a las condiciones de luz en la habitación en el momento que fue capturada. El *Kinect*® capturará la misma imagen de profundidad en una habitación iluminada o en una oscura.

Una imagen de profundidad además contiene información tridimensional precisa acerca de lo que se encuentra frente al *Kinect*®. A diferencia de una cámara convencional, que captura cómo se ven las cosas, una cámara de profundidad captura donde están estas. La información tridimensional del entorno del sensor es enviada desde el *Kinect*® hasta la computadora a través de un flujo de datos llamado *DepthImageStream*, que se puede configurar para enviar la información con diferentes grados de precisión y frecuencias de actualización. La Tabla 2.7 muestra

los diferentes formatos que se pueden utilizar para capturar y transmitir las imágenes de profundidad.

Tabla 2.7 Formatos soportados por *DepthImageStream*.

Nombre de la Propiedad	Descripción
Resolution320x240Fps30	Imágenes de 320x240 píxeles a 30 imágenes por segundo
Resolution640x480Fps30	Imágenes de 640x480 píxeles a 30 imágenes por segundo
Resolution80x60Fps30	Imágenes de 80x60 píxeles a 30 imágenes por segundo
Undefined	Formato sin definir

Dicha información es almacenada en un buffer denominado *DepthImageFrame*, que contiene la información de profundidad de un fotograma, por lo que es actualizado cada vez que el sensor captura un nuevo fotograma. Esto implica que, si los cálculos a realizar con dicho fotograma necesitan más tiempo para ejecutarse que el tiempo entre fotogramas, se deberá almacenar la información en otro buffer aparte; y probablemente no se podrá analizar la información de los fotogramas producidos durante dichos cálculos en tiempo real.

Cámara RGB

Kinect® cuenta con un sensor óptico que permite obtener imágenes a color en tiempo real. Dicho sensor cuenta con un campo de visión de 43° en vertical por 57° en horizontal y es capaz de capturar fotogramas a razón de 30 imágenes por segundo. Al igual que el sensor de profundidad, crea un flujo de datos de la cámara al ordenador llamado *ColorImageStream*, cuyos fotogramas son almacenados en el buffer *ColorImageFrame*.

También admite diferentes formatos en los que capturar y transmitir las imágenes (ver Tabla 2.8). El formato elegido dependerá de las necesidades de velocidad de captura y de precisión de la información recibida; y se puede elegir visualizar la imagen infrarroja para poder observar la escena en casos de luminosidad muy baja.

Tabla 2.8 Formatos soportados por ColorImageStream.

Nombre de la Propiedad	Descripción
RawYuvResolution640x480Fps15	Imágenes Raw YUV de 640x480 píxeles a 15 imágenes por segundo
RgbResolution1280x960Fps12	Imágenes RGB de 1280x960 píxeles a 12 imágenes por segundo
RgbResolution640x480Fps30	Imágenes RGB de 640x480 píxeles a 30 imágenes por segundo
YuvResolution640x480Fps15	Imágenes YUV de 640x480 píxeles a 15 imágenes por segundo
Undefined	Formato sin definir

Para la elaboración de este trabajo, los dos formatos que se consideraron más adecuados elegir fueron *RgbResolution1280x960Fps12* y *RgbResolution640x480Fps30*. El primero debido a la gran precisión que proporciona, y el segundo debido a que permite una tasa de captura de imágenes más elevada. En la Fig. 2.8 se muestra una comparativa entre ambos formatos de captura.



RgbResolution1280x960Fps12



RgbResolution640x480Fps30

Fig. 2.8 Comparativa entre resoluciones.

Finalmente, debido a la gran diferencia de calidad entre las imágenes se decidió optar por el formato *RgbResolution1280x920Fps12*. A pesar de la menor tasa de fotogramas por segundo que permite, resulta más importante obtener imágenes de

gran calidad de los rostros de los sujetos. La principal causa de esta decisión es la relevancia que tiene dentro del trabajo poder realizar un análisis del movimiento facial de forma precisa.

2.4.2 Interfaz de Seguimiento Facial de *Kinect*®

El Kit de Desarrollo de Software de Seguimiento Facial de Windows (SDK de Seguimiento Facial), junto con el Kit de Desarrollo de Software de *Kinect* para Windows (SDK de *Kinect para Windows*®), contienen una serie de librerías e interfaces que permiten seguir rostros humanos en tiempo real. El motor del seguimiento facial del SDK de Seguimiento Facial analiza los datos de entrada capturados por la cámara *Kinect*®, calcula la posición de la cara y expresiones faciales, y hace esa información disponible para cualquier aplicación en tiempo real.

Sistema de Coordenadas

El SDK de Seguimiento Facial usa el sistema de coordenadas del *Kinect*® para obtener sus resultados de seguimiento en 3D. El origen de coordenadas se localiza en el centro óptico de la cámara (ver Fig. 2.9 [30]); el eje Z apunta hacia el usuario, el eje Y apunta hacia arriba y el eje X hacia la izquierda del sensor. Las medidas se obtienen en metros para las distancias y en grados para los ángulos de rotación.

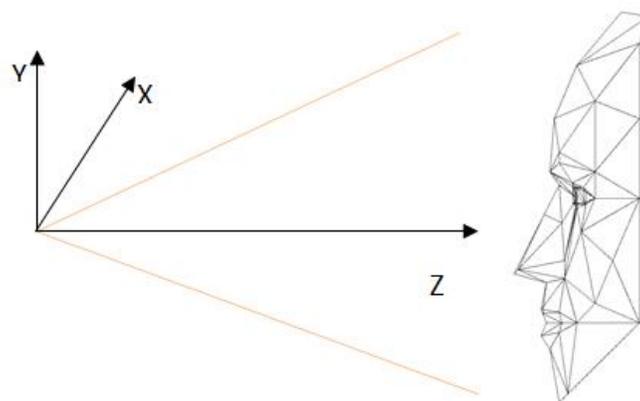


Fig. 2.9 Espacio de la cámara.

Imágenes de entrada

El SDK de Seguimiento Facial acepta como entrada imágenes de color y profundidad del *Kinect*®. La calidad de seguimiento puede verse afectada por la calidad de esas imágenes de entrada, ya que las imágenes oscuras o borrosas dificultan la tarea del seguimiento y reconocimiento facial. Por otro lado, las caras más grandes o cercanas son más fáciles de seguir.

Descripción de la API

La Interfaz de Programación de Aplicaciones (API, por sus siglas en inglés) de Seguimiento Facial, fue diseñada para funcionar en tiempo real durante el procesamiento de las imágenes de entrada. Para poder funcionar correctamente se sirve de una serie de interfaces (ver Tabla 2.9), que son los encargados de extraer, actualizar y procesar la información fotograma a fotograma.

Tabla 2.9 Principales interfaces de Kinect®.

Interfaz	Función
IFTFaceTracker	Principal interfaz de seguimiento facial
IFTResult	Resultado de una operación de seguimiento
IFTImage	Una interfaz de poyo que agrupa varios buffers de imágenes
IFTModel	Un modelo 3D ajustado al rostro seguido

También utiliza las estructuras de datos presentados presentadas en la Tabla 2.10.

Tabla 2.10 Estructuras de datos utilizadas por la API.

Estructura	Descripción
FT_SENSOR_DATA	Contiene todos los datos de entrada para una operación de seguimiento facial
FT_CAMERA_CONFIG	Contiene la configuración del sensor de video o profundidad
FT_VECTOR2D	Contiene los puntos de un vector 2D
FT_VECTOR3D	Contiene los puntos de un vector 3D
FT_TRIANGLE	Contiene un modelo facial triangular 3D
FT_WEIGHTED_RECT	Contiene un rectángulo ponderado devuelto por el SDK de Seguimiento Facial

Interfaz *IFTFaceTracker*

El interfaz principal es *IFTFaceTracker*, y se puede instanciar mediante el método *FTCreateFaceTracker*. Tras su inicialización, permite seguir un rostro de forma síncrona enviando imágenes de color y profundidad mediante el uso del objeto *IFTImage*. Estos resultados son devueltos por medio de una instancia de *IFTResult*. *IFTFaceTracker* proporciona el método *CreateFTResult* para crear una instancia de *IFTResult* en la que se almacenen los resultados del seguimiento facial específicos para el modelo que está empleando la instancia de *IFTFaceTracker*. Una aplicación necesitará crear una instancia de *IFTResult* para poder comenzar el seguimiento facial, y puede usar el método de *IFTFaceTracker* para obtener un arreglo de áreas que potencialmente contengan caras a partir de los datos de imagen (FT_SENSOR_DATA) proporcionados por la aplicación. Depende del programador interpretar los resultados y decidir qué caras seguir.

Para comenzar el seguimiento de una cara se necesita llamar al método *StartTracking*. Éste es un método caro que busca en la imagen de entrada una cara, determina su orientación y comienza a seguirla. Se pueden proporcionar instrucciones indicando en qué región de la imagen buscar una cara (pROI) o enviar NULL para buscar en toda la imagen, la primera cara que se encuentre será a la que se siga. Otro parámetro obtenible es la orientación de la cabeza. Dicho parámetro se puede obtener a partir de los datos del esqueleto proporcionados por *Kinect*®. Sin este parámetro se podrá realizar un seguimiento facial, pero los resultados iniciales podrían no ser óptimos.

Una vez que *StartTracking* ha comenzado a seguir un rostro de la forma indicada por *IFTResult*, la aplicación deberá continuar el seguimiento facial mediante llamadas sucesivas a *ContinueTracking*, que usa información de las llamadas anteriores a *StarTracking* o *ContinueTracking* para mantener un seguimiento facial preciso. Se debe seguir llamando a *ContinueTracking* hasta que se desee terminar el seguimiento facial o éste haya fallado, por ejemplo debido a que la persona cuya cara se estaba siguiendo salga fuera de la imagen de la cámara. Un fallo en el seguimiento facial es indicado mediante el estado *IFTResult*. Para comenzar de

nuevo el seguimiento facial, la aplicación llama a *StartTracking* de nuevo y después a *ContinueTracking* de forma secuencial.

ContinueTracking es una función relativamente rápida que usa la información existente del estado de una cara seguida, por ello es mucho más barato que *StartTracking*, que debe generar la información facial por completo. En algunas ocasiones es necesario llamar continuamente a *StartTracking*, especialmente si la aplicación tiene una tasa de imágenes baja o si la cara se mueve demasiado rápido entre imágenes, por lo que no es posible hacer un seguimiento continuo.

Interfaz *IFTResult*

La interfaz *IFTResult* proporciona acceso a los datos obtenidos como resultado de las llamadas de seguimiento facial realizadas mediante los métodos *IFTFaceTracker.StartTracking* e *IFTFaceTracker.ContinueTracking*. *IFTResult* se crea llamando al método *IFTFaceTracker.CreateFTResult*. *IFTFaceTracker* proporciona el método *CreateFTResult* ya que los resultados están relacionados con el modelo subyacente con el que *IFTFaceTracker* ha sido inicializado, y por lo tanto su creación depende en gran medida de éste.

Para determinar si el seguimiento facial ha tenido éxito (en cuyo caso se devuelve S_OK) se debe llamar al método *GetStatus*. Sobre una llamada exitosa de seguimiento facial, *IFTResult* proporciona acceso a la siguiente información:

- *GetFaceRect*: Un rectángulo expresado en función de las coordenadas de la imagen devuelta por la cámara que rodea la cara que está siguiendo.
- *Get2DShapePoints*: Coordenadas 2D de los puntos clave sobre la cara alineada en coordenadas de imagen de video. Incluye los 87 puntos 2D indicados.

Interfaz *IFTImage*

Los datos de imagen (video y profundidad) para el seguimiento facial se proporcionan mediante FT_SENSOR_DATA, una estructura de datos que contiene punteros a una imagen de video y a una imagen de profundidad. Se debe usar el

método *FTCreateImage* para crear instancias de *IFTImage*. *IFTImage* encapsula los datos de una imagen usada en seguimiento facial y especifica los formatos de imagen soportados mediante el atributo *FTIMAGEFORMAT*. Además proporciona buffer para guardar la imagen (*Allocate*) y también permite realizar un almacenamiento externo para ésta en disco (*Attach*). En el primer caso (*Allocate*) *IFTImage* libera la memoria asignada cuando se llama al método *Reset*. En el caso de (*Attach*) el programador es responsable de gestionar la memoria para indicar cómo guardar la imagen.

IFTImage proporciona varios métodos de acceso a la información sobre la imagen: el formato, la altura, anchura, tamaño de imagen y bytes por píxel. Además, *IFTImage* proporciona acceso al buffer. A esta interfaz también pertenecen métodos para la copia rápida de imágenes y la visualización de líneas de depuración, lo que la convierte en una interfaz sumamente útil desde el punto de vista del desarrollador.

Interfaz *IFTModel*

La interfaz *IFTModel* proporciona un sistema para convertir los puntos de seguimiento facial en una malla de vértices 3D en el espacio de cámara, que se superpone sobre el rostro de la persona que está siendo seguida. Su instancia es devuelta por el método *IFTFaceTracker.GetFaceModel*. Esta interfaz proporciona múltiples métodos para obtener varias propiedades del modelo:

- *GetSUCount*, *GetAUCount*: Devuelve el número de *Shape Units (SUs)* o *Animation Units (AUs)* usadas en el modelo lineal 3D.
- *GetTriangles*: Devuelve los triángulos expresados mediante sus vértices en coordenadas 3D de la malla del modelo facial. Cada triángulo tiene el índice de sus tres vértices listando en sentido horario.
- *GetVertexCount*: Devuelve el número de vértices existentes en la malla del modelo 3D.

Además, *IFTModel* proporciona dos métodos para obtener un modelo de cara 3D en el espacio de video de la cámara o proyectando en el plano de video de la cámara. Estos métodos son:

- *Get3DShape*: Devuelve los vértices del modelo de cara 3D transformados mediante las *Sus*, *AUs*, la escala de dimensionado, rotación y translación enviados.
- *GetProjectedShape*: Devuelve los vértices del modelo de cara 3D transformados mediante las *Sus*, *AUs*, la escala de dimensionado, rotación y translación enviados, y los proyecta sobre la imagen de video. Es especialmente útil a la hora de superponer la malla 3D o los puntos faciales detectados sobre el rostro de la persona seguida y poder cotejar los datos.

Puede usarse el método *FTCreateFaceTracker* para crear una instancia de *IFTFaceTracker* y el método *FTCreateImage* sirve para crear una instancia de *IFTImage*. *IFTResult* e *IFTModel* son creados a partir de *IFTFaceTracker*.

Salidas del Seguimiento Facial

Cada vez que los métodos *StartTracking* o *ContinueTracking* sean llamados, *FTResult* se actualizará. *FTResult* contiene la siguiente información de un usuario seguido:

1. Estado del seguimiento
2. Puntos 2D
3. Postura 3D de la cara
4. *AUs*, aunque su precisión es insuficiente para poder ser empleadas en el reconocimiento de expresiones faciales

Puntos faciales reconocibles por *Kinect*®

En este punto del trabajo, se pudo comprobar que la información proporcionada por la documentación de *Microsoft*® acerca de los puntos faciales que es capaz de detectar *Kinect*® era errónea. Supuestamente, la API de Seguimiento Facial rastrea los 87 puntos 3D indicados en la Fig. 2.10 [30], además de otros 13 que no aparecen en la imagen y se explican a continuación.

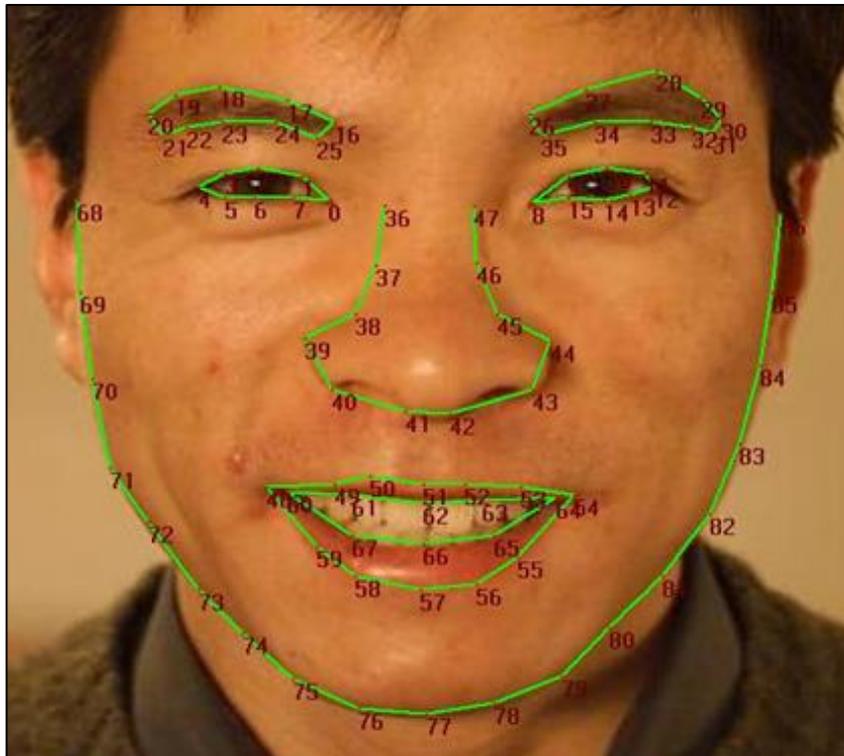


Fig. 2.10 Puntos supuestamente reconocibles por *Kinect*®.

Los 13 puntos restantes no mostrados en la Fig. 2.10 corresponden con:

87: Centro del ojo izquierdo

88: Centro del ojo derecho

89: Centro de la nariz

90-94: Interior de la ceja izquierda

95-99: Interior de la ceja derecha

No obstante, las pruebas realizadas a la hora de captar los puntos faciales empleando la cámara *Kinect*® demostraron que la información proporcionada por *Microsoft*® no era correcta. En lugar de los 100 puntos que se esperaba obtener, se pudo comprobar que la cámara realizaba un seguimiento de 121 puntos faciales, estos puntos están basados en el modelo facial CANDIDE-3 que utiliza el SDK de Seguimiento Facial (ver Fig. 2.11 [30]).

Este cambio puso de manifiesto uno de los puntos débiles de la cámara *Kinect*® como mecanismo de reconocimiento automático de expresiones faciales, y este es que el *Kinect*® es incapaz de detectar el grado de apertura de los ojos de un rostro. Esto se debe a que no todos los puntos faciales son actualizados por *Kinect*® fotograma por fotograma, si no que existen regiones como el contorno de los ojos y la nariz que sólo son empleados en el momento de reconocer y seguir la cara [31].

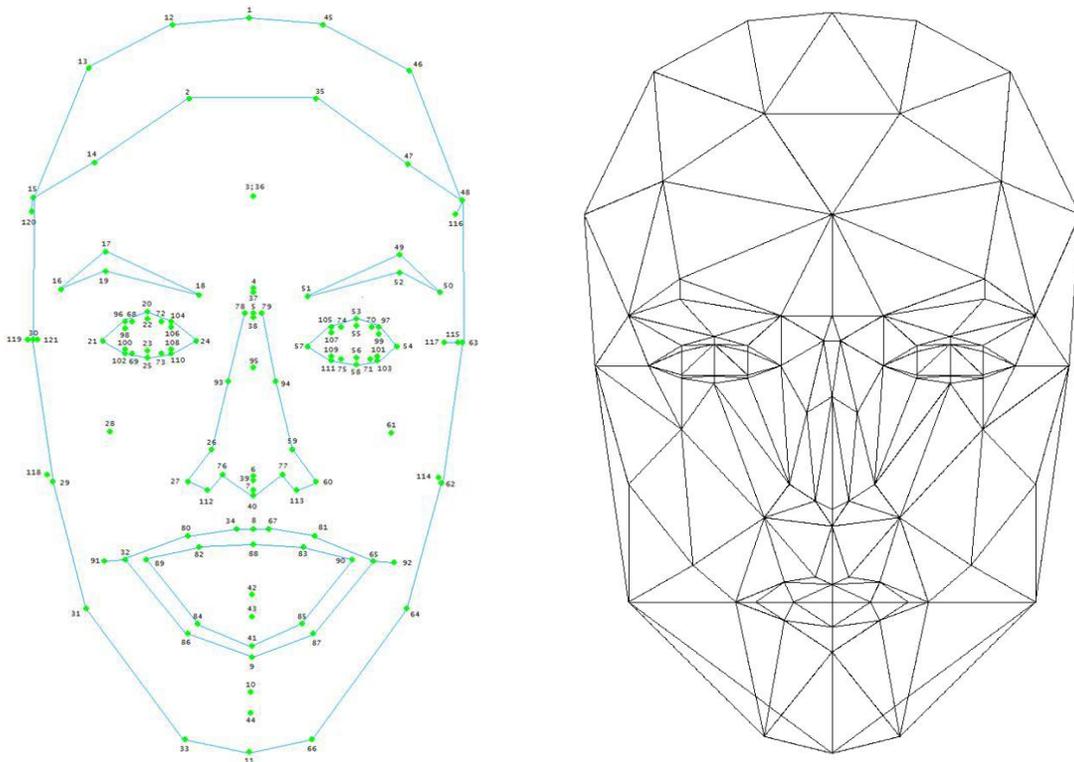


Fig. 2.11 Mapeado de los puntos reconocibles y el modelo CANDIDE-3.

Postura 3D de la cara

La posición (X,Y,Z) de la cabeza del usuario se reporta basándose en un sistema de coordenadas diestro, es decir, establecido con el origen en el sensor, el eje Z apuntando hacia el usuario, el eje Y apuntando hacia arriba y el eje X apuntando hacia la izquierda del sensor. Las distancias se miden en metros. La postura de la cabeza del usuario queda definida mediante tres ángulos (ver Fig. 2.12 [30]).

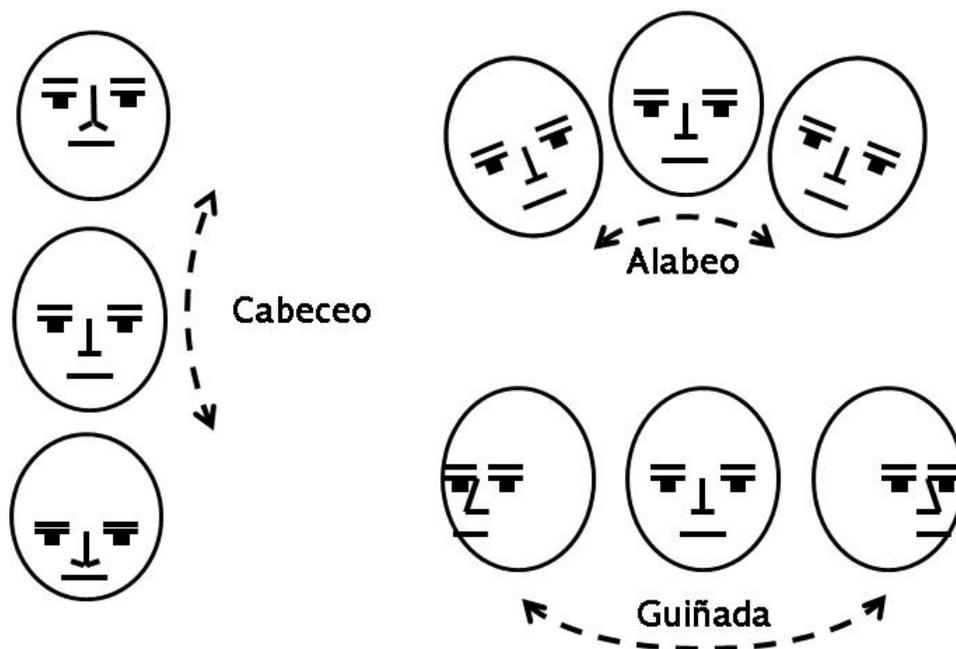


Fig. 2.12 Ángulos de la postura de la cabeza.

Los ángulos se expresan en grados, con valores que varían entre -180° y $+180^\circ$ (ver Tabla 2.11).

Tabla 2.11 Explicación de los ángulos medidos por Kinect®.

Ángulo	Valor y Posición de la Cabeza
Ángulo de Cabeceo o Pitch 0° = neutral	-90° = Mirando al suelo
	+90° = Mirando al techo
Ángulo de Alabeo o Roll 0° = neutral	El Seguimiento Facial capta cuando la cabeza del usuario tiene un ángulo de Cabeceo inferior a 20°, pero funciona de forma óptima cuando éste es inferior a 10°.
	-90° = Horizontal y paralela al hombro derecho del sujeto +90° = Horizontal y paralela al hombro izquierdo del sujeto
Ángulo de Guiñada o Yaw 0° = neutral	El Seguimiento Facial capta cuando la cabeza del usuario tiene un ángulo de Alabeo inferior a 90°, pero funciona de forma óptima cuando éste es inferior a 45°.
	-90° = Girada completamente hacia el hombro derecho del sujeto +90° = Girada completamente hacia el hombro izquierdo del sujeto
	El Seguimiento Facial capta cuando la cabeza del usuario tiene un ángulo de Guiñada inferior a 45°, pero funciona de forma óptima cuando éste es inferior a 30°.

2.5 C Sharp (C#)

El lenguaje de programación C Sharp, también llamado C# es un lenguaje de programación orientado a objetos desarrollado por *Microsoft*® e incluido en la plataforma .NET. Emplea una sintaxis derivada de C y C++, aunque C Sharp fue diseñado como una mejora de ambos lenguajes. Este lenguaje a supuesto uno de los pilares fundamentales en la elaboración del proyecto, ya que su potencia visual y su gran sinergia con el SDK de *Kinect*® lo convierten en el lenguaje más adecuado para la elaboración de las herramientas de adquisición de la información y de análisis de resultados. El lenguaje C Sharp, junto con el entorno de desarrollo *Microsoft Visual Studio*® permite el desarrollo de aplicaciones visualmente atractivas de una forma rápida, sencilla e intuitiva [30].

A continuación se mostrarán y explicarán algunos fragmentos del código del sistema para comprender mejor el funcionamiento de éste y del propio lenguaje de programación.

```
public MainWindow()
{
    InitializeComponent();
    var faceTrackingViewerBinding = new System.Windows.Data.Binding("Kinect") {
        Source = sensorChooser };
    faceTrackingViewer.SetBinding(FaceTrackingViewer.KinectProperty,
        faceTrackingViewerBinding);
    sensorChooser.KinectChanged += SensorChooserOnKinectChanged;
    sensorChooser.Start();
}
```

La función mostrada se ejecuta al crear la ventana principal del programa, llamada *MainWindow*. Posteriormente, se prepara el programa para crear un vínculo con el sensor *Kinect*® y poder empezar a recibir información de él. Finalmente, con la instrucción *sensorChooser.Start* se inicia la búsqueda de sensores *Kinect*® conectados al equipo. En caso de que la búsqueda tenga éxito, la variable

sensorChooser será la que contenga la información del sensor a partir de ahora, y permitirá acceder a los parámetros de éste, como se podrá observar posteriormente.

```
private void KinectSensorOnAllFramesReady(object sender, AllFramesReadyEventArgs
allFramesReadyEventArgs)
{
    using (var colorImageFrame = allFramesReadyEventArgs.OpenColorImageFrame())
    {
        if (colorImageFrame == null)
        {
            return;
        }
        // Copia de la imagen RGB para mostrarla en pantalla
        var haveNewFormat = this.currentColorImageFormat !=
colorImageFrame.Format;
        if (haveNewFormat)
        {
            this.currentColorImageFormat = colorImageFrame.Format;
            this.colorImageData = new byte[colorImageFrame.PixelDataLength];
            this.colorImageWritableBitmap = new WriteableBitmap(
colorImageFrame.Width, colorImageFrame.Height, 96, 96,
PixelFormats.Bgr32, null);
            ColorImage.Source = this.colorImageWritableBitmap;
        }
        colorImageFrame.CopyPixelDataTo(this.colorImageData);
        this.colorImageWritableBitmap.WritePixels(
new Int32Rect(0, 0, colorImageFrame.Width, colorImageFrame.Height),
this.colorImageData, colorImageFrame.Width * Bgr32BytesPerPixel, 0);
    }
}
```

El método *KinectSensorOnAllFramesReady* es ejecutado cada vez que la cámara *Kinect*® envía un nuevo fotograma al sistema, por lo tanto puede emplearse para capturar la información de dicho fotograma con el fin de realizar los cálculos pertinentes sobre él o mostrar la información RGB por pantalla. En este caso, el fotograma capturado por la cámara RGB será mostrado por pantalla con el fin de

permitir al usuario del sistema apreciar la información que está siendo captada por la cámara *Kinect*® (ver Fig. 2.13). Para ello, en primer lugar es necesario comprobar que se ha recibido correctamente una imagen RGB mediante el método *allFramesReadyEventArgs.OpenColorImageFrame*, que devolverá NULL en caso negativo. Una vez asegurada la recepción de la imagen se comprueba que el formato sea correcto, si no lo fuera es necesario adaptar el formato de la imagen recibida a la superficie donde la imagen será mostrada. Finalmente, las dos últimas funciones sirven para guardar el fotograma en memoria y mostrarlo por pantalla.



Fig. 2.13 Imagen RGB captada por *Kinect*®.

A pesar de poder apreciar la información visual recibida por *Kinect*®, en la Fig. 2.13 no es posible notar si el sensor ha sido capaz de detectar y seguir el rostro de la persona, ni se puede visualizar dicha información facial. Para ello se deberá indicar que cada vez que se genere un nuevo fotograma, se superpongan los datos faciales recibidos con la imagen RGB. Este resultado se puede conseguir introduciendo una

llamada al método *DrawFaceModel* cada vez que se produzca una ejecución del método *KinectSensorOnAllFramesReady*.

```
public void DrawFaceModel(DrawingContext drawingContext)
{
    if (!this.lastFaceTrackSucceeded || this.skeletonTrackingState !=
        SkeletonTrackingState.Tracked)
    {
        facepointsReady = false;
        return;
    }
    facepointsReady = true;
    var faceModelPts = new List<Point>();
    for (int i = 0; i < this.facePoints.Count; i++)
    {
        // Ajusta para 1280 x 920
        faceModelPts.Add(new Point((this.facePoints[i].X)/2,
            (this.facePoints[i].Y )/2 ));
    }
    var faceModelGroup = new GeometryGroup();
    for (int i = 0; i < faceModelPts.Count; i++)
    {
        var facePoints = new GeometryGroup();
        var ellipseGeometry = new EllipseGeometry(faceModelPts[i], 1, 1);
        facePoints.Children.Add(ellipseGeometry);
        faceModelGroup.Children.Add(facePoints);
    }
    drawingContext.DrawGeometry(Brushes.Yellow, new Pen(Brushes.Yellow, 1.0),
        faceModelGroup);
}
```

El método *DrawFaceModel* comprueba que se haya detectado un rostro y actualmente esté siendo seguido. En caso negativo, la función finaliza al no ser posible dibujar los puntos faciales. En caso afirmativo, procede a almacenar todos los puntos faciales devueltos por *Kinect*® en la lista de puntos bidimensionales *pointsModel* y finalmente los dibuja en la superficie de dibujo *drawingContext*. De

ese modo, los puntos faciales detectados quedarán superpuestos con la imagen RGB captada por *Kinect*®; y será posible determinar la precisión del seguimiento facial. Esta característica añadida resulta especialmente útil ya que la cámara *Kinect*® necesita que se den unas condiciones de luz y proximidad del sujeto adecuadas, a pesar de los márgenes admitidos; y la superposición de los puntos faciales permite la detección de deficiencias en alguna de estas variables. En la Fig. 2.14 se puede apreciar el resultado final de esta operación.



Fig. 2.14 Superposición de puntos faciales.

2.5.1 Ventajas frente a otros lenguajes

- Permite el paso de parámetros a las funciones por valor por defecto, o por referencia mediante el uso de la palabra reservada *ref*. Asimismo, es posible enviar un número variable de parámetros a un mismo método mediante la sobrecarga de éste, es decir, mediante su definición sucesiva con diferente

número de parámetros. Esto permite una gran flexibilidad en el uso de funciones.

- En C# están disponibles una mayor cantidad de tipos de datos que en C++ o Java, lo que permite llevar a cabo tareas más completas y específicas.
- C# permite realizar declaraciones en el espacio de nombres, es decir, al iniciar la programación de una aplicación es posible declarar las clases del propio espacio de nombres.
- A diferencia de Java, C# permite la utilización de cualquier clase como un objeto, por lo que permite la definición de propiedades en cualquier clase.
- C# permite la ejecución de diferentes versiones de un mismo software de forma simultánea. Esto es posible gracias a la capacidad de mantener múltiples versiones de una misma clase, colocándolas en diferentes espacios de nombres.
- Permite la implementación de aplicaciones multihilo de forma muy simplificada.

2.5.2 Desventajas frente a otros lenguajes

- Es necesario disponer de una versión reciente de Visual Studio para poder programar eficientemente en C#.
- C# es un lenguaje complejo de aprender debido a su gran potencia y cantidad de librerías disponibles.
- Los programas desarrollados en C# sólo pueden ser ejecutados en entornos Windows. Para el desarrollo de este trabajo, no es un verdadero inconveniente, ya que las librerías de *Kinect®* han sido diseñadas por *Microsoft®* para ser empleadas exclusivamente en un entorno *Windows®*.

2.6 XAML

XAML ha sido el lenguaje de programación elegido para la estructuración de la interfaz de usuario de la herramienta de adquisición de imágenes e información facial debido a su facilidad de uso e interpretación. El origen de XAML se remonta a los años 70, cuando IBM creó un lenguaje llamado GML (Generalized Markup Language) con el objetivo de almacenar grandes cantidades de información. Debido

al éxito de GML, en 1986 la Organización Internacional de Normalización (ISO, por sus siglas en inglés) decidió normalizar el lenguaje, lo que dio lugar a la aparición de SGML. Con la aparición de la web y el lenguaje HTML en 1989, SGML aumentó considerablemente su popularidad. No obstante, las páginas web exclusivamente creadas en HTML resultaron ser caóticas y presentaban una gran cantidad de errores.

XML surgió entonces con el objetivo de mitigar los errores de sintaxis y facilitar la tarea al programador. XAML es una versión de XML optimizada para la creación sencilla de interfaces de usuario visualmente potentes. XAML proviene de las siglas en inglés de Lenguaje de Etiquetado de Aplicaciones Extensible (Extensible Application Markup Language). Es un lenguaje declarativo basado en el uso de etiquetas al igual que HTML, lo que permite que la programación esté muy bien estructurada y diferenciada. En la actualidad XAML ha experimentado un gran crecimiento, al haber sido el lenguaje de programación elegido para la estructuración de la interfaz de usuario de las aplicaciones para dispositivos Android en su entorno de desarrollo nativo.

2.6.1 Código XAML de la herramienta de adquisición de información

Todos los archivos XAML deben tener un elemento raíz, es decir, un objeto que será la base conceptual de toda la estructura de programación del documento. Este objeto típicamente será una ventana, un lienzo (canvas) o un gráfico.

```
<Window
    x:Class="FaceTrackingBasics.MainWindow"
    xmlns="http://schemas.microsoft.com/winfx/2006/xaml/presentation"
    xmlns:x="http://schemas.microsoft.com/winfx/2006/xaml"
    xmlns:local="clr-namespace:FaceTrackingBasics"
    Title="Seguimiento Facial Kinect"
    Closed="WindowClosed"
    Height="700" Width="770" ResizeMode="NoResize"
    >
    <Window.Resources>
        <SolidColorBrush x:Key="MediumGreyBrush" Color="#ff6e6e6e"/>
```

```

        <SolidColorBrush x:Key="KinectPurpleBrush" Color="#ff52318f"/>
        <SolidColorBrush x:Key="KinectBlueBrush" Color="#ff00BCF2"/>
        <Style TargetType="{x:Type Image}">
        <Setter Property="SnapsToDevicePixels" Value="True"/>
        </Style>
    </Window.Resources>

```

Inicialmente se establece como elemento raíz del documento una ventana con el título de “Seguimiento Facial Kinect” con unas dimensiones de 700 píxeles de alto por 770 píxeles de ancho; se configura la propiedad *ResizeMode* para que la ventana no pueda modificarse en tamaño por el usuario. Sobre esta ventana se definen una serie de recursos (*Resources*) que posteriormente serán empleados en la elaboración del estilo de la ventana. Finalmente se indica a la aplicación que se debe emplear la resolución nativa del dispositivo sobre el que está corriendo a la hora de renderizar la ventana.

```

<Grid Name="layoutGrid" Margin="10,0,10,5">
    <Grid.RowDefinitions>
        <RowDefinition Height="Auto"/>
        <RowDefinition Height="*/>
    </Grid.RowDefinitions>
    <DockPanel Grid.Row="0" Margin="0 0 0 0">
        <Image DockPanel.Dock="Left" Source="Images\Logo.png"
            Stretch="Fill" Height="32" Width="81" Margin="0 10 0 5"/>
        <TextBlock DockPanel.Dock="Right" Margin="0 0 -1 0"
            VerticalAlignment="Bottom" Foreground="{StaticResource
            MediumGreyBrush}" FontFamily="Segoe UI" FontSize="18">Nottingham
            System</TextBlock>
        <Image Grid.Column="1" Source="Images\Status.png" Stretch="None"
            HorizontalAlignment="Center" Margin="0 0 0 5"/>
    </DockPanel>

```

En este fragmento del código se crea una cuadrícula (*Grid*) que contiene y organiza los elementos dentro de la interfaz, y se establece que la altura de sus filas (*Rows*) se adapte automáticamente al tamaño de la ventana. Después se crea un panel que

ocupa la primera fila de la cuadrícula, y que contiene una imagen con el logo de *Kinect*® a la izquierda, el estado actual de la conexión con la cámara en el centro y el texto “Nottingham System” a la derecha.

```
<Viewbox Grid.Row="1" Stretch="Uniform" HorizontalAlignment="Center"
Margin="0,0,0,0">
    <Grid Name="MainGrid" Width="640" Height="480">
        <Image Name="ColorImage" Margin="0,-12,0,12"/>
        <local:FaceTrackingViewer x:Name="faceTrackingViewer"
Margin="0,-12,0,12" />
    </Grid>
</Viewbox>
<DockPanel Grid.Row="2" Margin="0,573,0,7">
    <Button x:Name="selectFolder" Content="Selecciona carpeta"
Height="25" Width="120" Click="selectDirectory_Click"
HorizontalAlignment="Left"/>
    <TextBox x:Name="path" Height="25" TextWrapping="Wrap" Text=""
Width="240" HorizontalAlignment="Left" Margin="4,0,0,0"
VerticalContentAlignment="Center" IsReadOnly="True"
HorizontalScrollBarVisibility="Auto"
VerticalScrollBarVisibility="Auto"/>
    <TextBlock Height="25" TextWrapping="Wrap" Text="Selecciona una
expresión facial" FontSize="10" Width="70" Margin="4,0,0,0"
HorizontalAlignment="Left" TextAlignment="Center"/>
    <ComboBox x:Name="facialExpression" Width="120" Height="25"
HorizontalAlignment="Left" Margin="4,0,0,0"
ItemsSource="{StaticResource List}"/>
    <Button x:Name="capture" Content="Capturar" Height="25"
Width="88" HorizontalAlignment="Left" Margin="4,0,0,0"
IsEnabled="False" Click="captureImage_Click"/>
    <Button x:Name="diagnostic" Content="Diagnóstico" Height="25"
HorizontalAlignment="Left" Margin="4,0,0,0" Width="75"
Click="diagnostic_Click"/>
</DockPanel>
</Grid>
</Window>
```

Después se crea una segunda fila de la cuadrícula en la que se ubica el lienzo en el que poder mostrar la información visual de *Kinect® (ColorImage)*. Finalmente se crea una tercera fila en la que se coloca un panel que contienen los elementos principales de control del sistema. Cada uno de estos elementos será descrito a detalle posteriormente en el documento.

A continuación se muestra en la Fig. 2.15 el resultado final de la interfaz de usuario de la aplicación de adquisición de información facial.

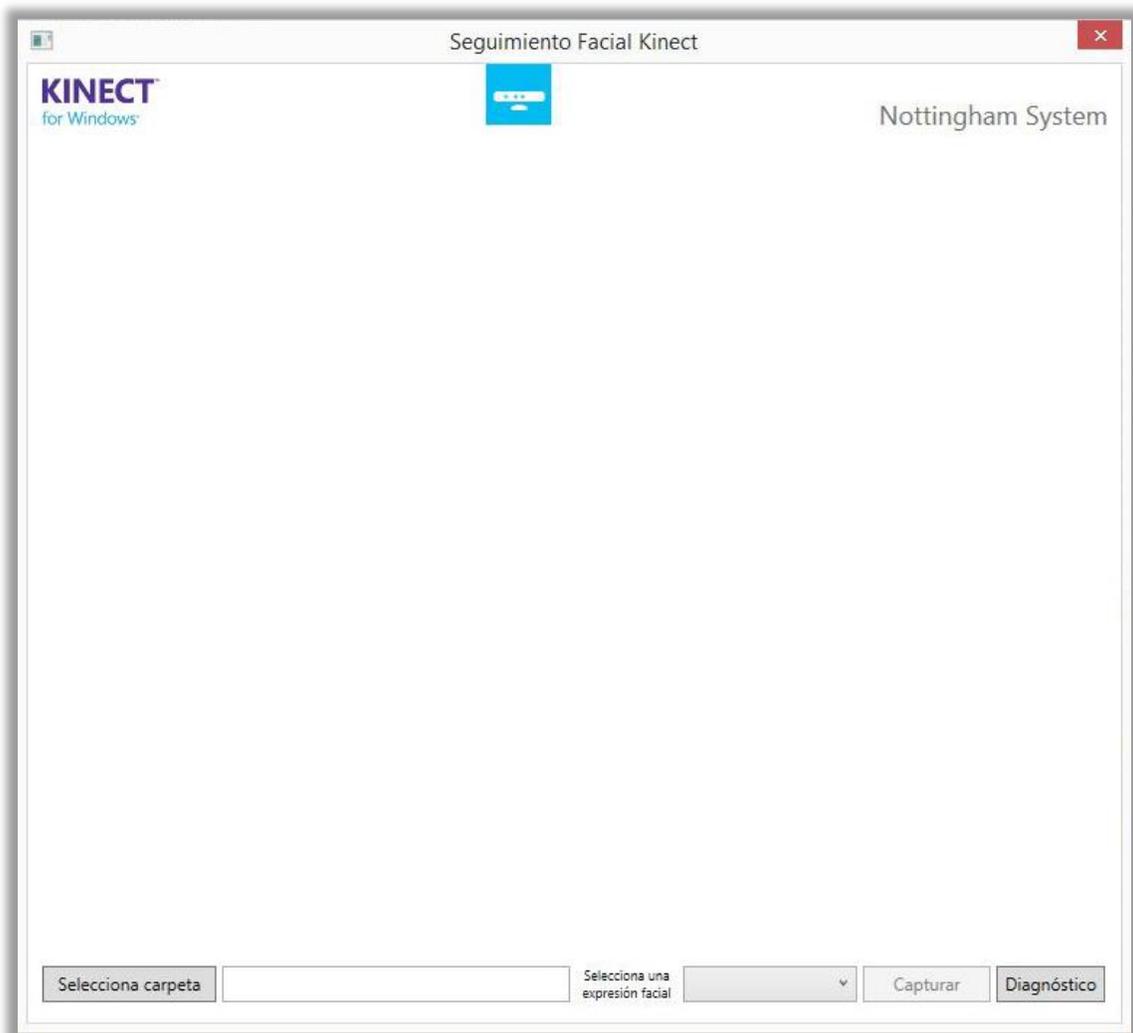


Fig. 2.15 Interfaz de usuario en XAML.

2.6.2 Ventajas frente a otros lenguajes

- Es un lenguaje capaz de ejecutar tareas típicas de lenguajes más complejos, como pueden ser la asignación de variables o la creación y asignación de listas y colecciones.
- Permite establecer una clara separación entre la programación de la interfaz de usuario y de la propia aplicación, lo que facilita en gran medida las labores de depuración del software.
- Debido a su diseño eminentemente enfocado a la creación de interfaces, con XAML es posible crear estructuras visuales muy complejas mediante el empleo de un número reducido de instrucciones sencillas, permitiendo enfocar los esfuerzos del desarrollo a la parte funcional de las aplicaciones.
- Permite la elaboración de aplicaciones que realicen un uso intensivo de gráficos complejos ya que, al permitir interoperabilidad con *DirectX*, la renderización de los gráficos será acelerada por hardware.

CAPÍTULO 3: METODOLOGÍA

3.1 Introducción

En este capítulo se ofrecerá una visión en profundidad de la estructura y el funcionamiento del sistema elaborado. Esto permitirá comprender todos los procesos que es necesario llevar a cabo por parte del sistema. Para ello, en primer lugar se mostrará la visión global del sistema, mostrando a grandes rasgos su estructura. Posteriormente se procederá a explicar con detalle la disposición interna de cada una de las tres fases que componen el sistema desarrollado. En este apartado, se expondrá el funcionamiento de cada una de ellas y se explicarán las funciones más relevantes.

3.2 Visión global del sistema

Tras realizar un profundo estudio del estado del arte acerca de las metodologías para el diagnóstico de la parálisis facial, se puso en manifiesto la necesidad de crear un sistema capaz de ofrecer una propuesta objetiva para asistir al médico en el proceso de diagnóstico de la parálisis facial. Con ese fin, se desarrolló el sistema que consta de tres fases principales: adquisición de los datos, almacenamiento de los datos y evaluación de los datos.

En la primera fase, la adquisición de los datos se realizó obteniendo la información visual y facial proporcionada por el *Kinect*® en tiempo real. Para lograr esta tarea, se crearon diferentes herramientas, las cuales resultan esenciales dentro del trabajo, ya que la exactitud y fiabilidad con que son recogidos los datos dependen todo su posterior análisis y los resultados obtenidos.

Para la fase de almacenamiento de los datos, se consideró guardar una imagen del paciente en formato JPEG al momento de realizar alguna de las expresiones faciales previamente seleccionadas. Al mismo tiempo, se creará un archivo de texto asociado a la imagen recién guardada que contiene las distancias entre los puntos de interés para el sistema de clasificación Nottingham.

La última fase toma como entrada los archivos de texto generados por la fase anterior y realiza sobre ellos los cálculos necesarios para obtener un resultado (diagnóstico). La Fig. 3.1 muestra las tres fases anteriormente descritas que forman parte del sistema.



Fig. 3.1 Estructura del sistema.

3.3 Adquisición de los datos

El cometido principal de esta fase es permitir la visualización en tiempo real de las imágenes RGB captadas por el *Kinect*® e identificar la información de los puntos faciales de interés para el sistema de clasificación Nottingham. De esta manera el usuario puede determinar la adecuación de las condiciones del procedimiento y modificar cualquiera de ellas en caso de ser necesario. Además, también permite apreciar si el *Kinect*® está detectando un rostro mediante la superposición en la interfaz de los puntos faciales de interés detectados con el rostro de la persona. Esta visión aporta información adicional acerca de la validez del procedimiento, ya que la no detección del rostro o su detección defectuosa podrían deberse a unas condiciones no óptimas en cuanto a luminosidad o distancia de la persona a la cámara del *Kinect*®.

La obtención de los datos constituye la fase inicial del ciclo de vida del sistema. Su funcionamiento se descompone en las siguientes etapas que serán descritas posteriormente:

1. Establecimiento de la comunicación con el sensor Kinect.
2. Detección y seguimiento del rostro de una persona.
3. Discriminación de la información facial y visual de la persona.

3.3.1 Establecimiento de la comunicación con el *Kinect*®

Para comenzar a recibir información del *Kinect*®, el primer paso a ejecutar es la búsqueda de sensores conectados a la computadora y detectar la validez de los mismos.

```

private void SensorChooserOnKinectChanged(object sender, KinectChangedEventArgs
kinectChangedEventArgs)
{
    KinectSensor oldSensor = kinectChangedEventArgs.OldSensor;
    KinectSensor newSensor = kinectChangedEventArgs.NewSensor;
    if (oldSensor != null)
    {
        oldSensor.AllFramesReady -= KinectSensorOnAllFramesReady;
        oldSensor.ColorStream.Disable();
        oldSensor.DepthStream.Disable();
        oldSensor.DepthStream.Range = DepthRange.Default;
        oldSensor.SkeletonStream.Disable();
        oldSensor.SkeletonStream.EnableTrackingInNearRange = false;
        oldSensor.SkeletonStream.TrackingMode = SkeletonTrackingMode.Default;
    }
}

```

En primer lugar se debe comprobar que no exista una conexión anteriormente establecida con un sensor *Kinect*®. En caso de que ya existiera, la conexión se deshace mediante la ruptura de los flujos de datos con la cámara.

```

    if (newSensor != null)
    {
        try
        {
            newSensor.ColorStream.Enable(ColorImageFormat.RgbResolution128
            0x960Fps12);
            newSensor.DepthStream.Enable(DepthImageFormat.Resolution640x48
            0Fps30);
            try
            {
                newSensor.DepthStream.Range = DepthRange.Near;
                newSensor.SkeletonStream.EnableTrackingInNearRange =
                true;
            }
            catch (InvalidOperationException)
            {
                newSensor.DepthStream.Range = DepthRange.Default;
                newSensor.SkeletonStream.EnableTrackingInNearRange =
                false;
            }
            newSensor.SkeletonStream.TrackingMode =
            SkeletonTrackingMode.Seated;
            newSensor.SkeletonStream.Enable();
            newSensor.AllFramesReady += KinectSensorOnAllFramesReady;
        }
        catch (InvalidOperationException){}
    }
}

```

Después de romper la posible conexión con el sensor antiguo, se comprueba que haya un sensor conectado actualmente y se establece la nueva conexión. Durante el establecimiento se fija el formato en que la imagen RGB (1280x960 píxeles a 12 fps) y la imagen tridimensional (640x480 a 30 fps) serán transmitidas por la cámara.

También se indica a la cámara que envíe la información del esqueleto de la persona detectada, lo cual proporciona una mayor fiabilidad a la hora de ubicar tridimensionalmente su rostro. El modo de captar el esqueleto se establece como *Seated*, por lo que no es necesario disponer de información sobre la parte inferior del cuerpo, y por tanto es más flexible que el modo estándar.

3.3.2 Detección y seguimiento del rostro de una persona

Una vez se ha establecido correctamente la conexión con la cámara *Kinect*®, el programa comienza a mostrar en pantalla las imágenes RGB enviadas por ésta. A continuación inicia la búsqueda de rostros dentro del espacio captado por el *Kinect*®, apoyándose para ello en la información visual y tridimensional.

```
private void OnAllFramesReady(object sender, AllFramesReadyEventArgs
allFramesReadyEventArgs)
{
    ColorImageFrame colorImageFrame = null;
    DepthImageFrame depthImageFrame = null;
    SkeletonFrame skeletonFrame = null;
    try
    {
        colorImageFrame = allFramesReadyEventArgs.OpenColorImageFrame();
        depthImageFrame = allFramesReadyEventArgs.OpenDepthImageFrame();
        skeletonFrame = allFramesReadyEventArgs.OpenSkeletonFrame();
        if (colorImageFrame == null || depthImageFrame == null || skeletonFrame
== null)
        {
            return;
        }

        // Comprueba si hay cambios de formato de imagen
        // FaceTracker no se ocupa de eso por lo que debemos resetear
        if (this.depthImageFormat != depthImageFrame.Format)
        {
            this.ResetFaceTracking();
            this.depthImage = null;
            this.depthImageFormat = depthImageFrame.Format;
        }
        if (this.colorImageFormat != colorImageFrame.Format)
        {
            this.ResetFaceTracking();
            this.colorImage = null;
        }
    }
}
```

```

        this.colorImageFormat = colorImageFrame.Format;
    }
    // Crea búfers para almacenar copias de los datos con los que trabajamos
    if (this.depthImage == null)
    {
        this.depthImage = new short[depthImageFrame.PixelDataLength];
    }
    if (this.colorImage == null)
    {
        this.colorImage = new byte[colorImageFrame.PixelDataLength];
    }
    // Obtiene la información del esqueleto
    if (this.skeletonData == null || this.skeletonData.Length !=
        skeletonFrame.SkeletonArrayLength)
    {
        this.skeletonData = new
            Skeleton[skeletonFrame.SkeletonArrayLength];
    }
    colorImageFrame.CopyPixelDataTo(this.colorImage);
    depthImageFrame.CopyPixelDataTo(this.depthImage);
    skeletonFrame.CopySkeletonDataTo(this.skeletonData);

```

Como ya se explicó en el capítulo anterior, la función *OnAllFramesReady* se ejecuta cada vez que el *Kinect*® captura un nuevo fotograma y lo envía al sistema a través de varios flujos de datos. Por ello, este es el método más adecuado en el que incluir las actualizaciones del seguimiento facial.

En primer lugar los datos de la imagen RGB, la imagen de profundidad y el esqueleto detectado deben ser abiertos y asignados a unas variables que permitirán su posterior tratamiento (*colorImageFrame*, *depthImageFrame* y *skeletonFrame*). Si alguno de los tipos de datos no existiera (*null*) la función se termina ya que no es posible realizar el seguimiento facial.

A continuación se guarda la imagen RGB en una nueva variable y se muestra por pantalla para que el usuario de la herramienta pueda visualizar la información captada por el *Kinect*®.

```

    // Actualiza la lista de los rastreadores y los rastreadores con la
    // información actual del frame
    foreach (Skeleton skeleton in this.skeletonData)
    {
        if (skeleton.TrackingState == SkeletonTrackingState.Tracked
            || skeleton.TrackingState == SkeletonTrackingState.PositionOnly)
        {
            // Queremos tener un registro de cualquier esqueleto,

```

```

        // rastreado o no rastreado
        If
        (!this.trackedSkeletons.ContainsKey(skeleton.TrackingId))
        {
            this.trackedSkeletons.Add(skeleton.TrackingId, new
                SkeletonFaceTracker());
        }
        // Da a cada rastreador el frame actualizado
        SkeletonFaceTracker skeletonFaceTracker;
        if (this.trackedSkeletons.TryGetValue(skeleton.TrackingId,
            out skeletonFaceTracker))
        {
            skeletonFaceTracker.OnFrameReady(this.Kinect,
                colorImageFormat, colorImage, depthImageFormat,
                depthImage, skeleton);
            skeletonFaceTracker.LastTrackedFrame =
                skeletonFrame.FrameNumber;
        }
    }
    this.RemoveOldTrackers(skeletonFrame.FrameNumber);
    this.InvalidateVisual();
}
finally
{
    if (colorImageFrame != null)
    {
        colorImageFrame.Dispose();
    }
    if (depthImageFrame != null)
    {
        depthImageFrame.Dispose();
    }
    if (skeletonFrame != null)
    {
        skeletonFrame.Dispose();
    }
}
}
}

```

En este último fragmento de la función, se ejecuta un bucle para el tratamiento de todos los esqueletos detectados en el fotograma. Para cada esqueleto se comprueba si ya estaba siendo seguido o se trata de un esqueleto nuevo, en el caso de ser nuevo se incluye a la lista de esqueletos activos. Si un esqueleto activo no es detectado durante un intervalo de 100 fotogramas consecutivos se borrará de la lista para ahorrar espacio en memoria.

A continuación se asigna al rastreador facial del esqueleto el nuevo valor para actualizar el estado del rostro. Esta operación es empleada por la API de seguimiento facial del *Kinect*® la cual hace uso de los algoritmo de Viola y Jones

[32] para la detección y del algoritmo presentado en la Fig. 3.2 para el seguimiento facial.

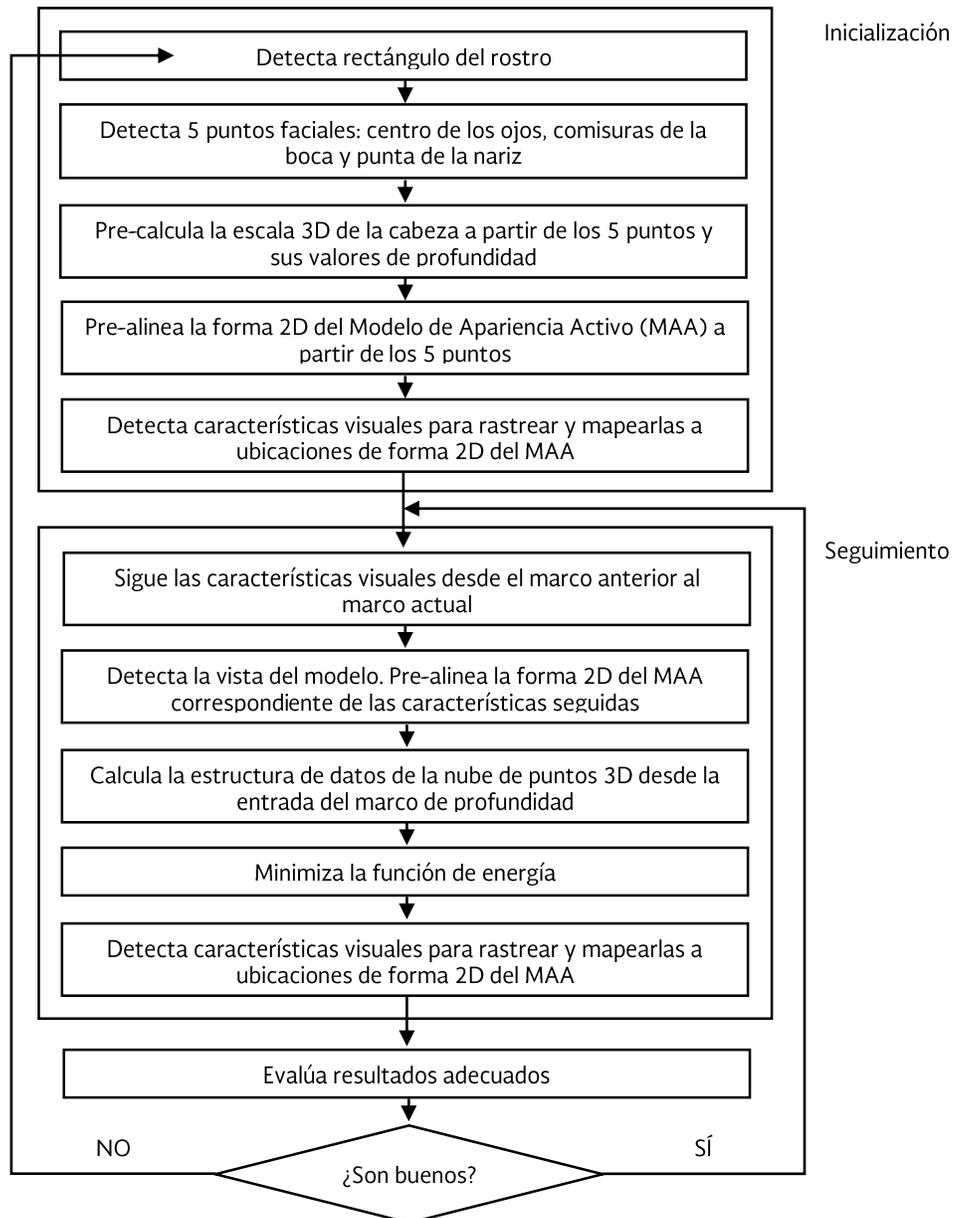


Fig. 3.2 Diagrama de flujo del algoritmo de seguimiento facial usado por Kinect®

En el diagrama de flujo anterior se comienza por encontrar un rectángulo del rostro en un cuadro de video usando un detector de rostros. Después utiliza una red

neuronal para encontrar los 5 puntos dentro del área de la cara (centro de los ojos, comisuras de la boca y punta de la nariz). Se pre-calcula la escala del rostro rastreado a partir de los 5 puntos sin proyección al espacio de la cámara 3D y se escala el modelo del rostro 3D apropiadamente.

Cuando los datos de profundidad no están disponibles (algunos cuadros de profundidad pueden no tener datos alrededor de las áreas del rostro), se asume la escala de la cabeza del usuario para que coincida con el tamaño promedio de la cara. También se alinea una forma 2D del MMA a estos cinco puntos característicos. Esto mejora la convergencia inicial cuando minimizamos la función de energía 3.1.

$$E = W_{2D}E_{2D} + W_{2D3D}E_{2D3D} + W_{depth}E_{depth} + W_{temp}E_{temp} + W_{fseg}E_{fseg} + W_{reg}E_{reg} \quad (3.1)$$

Donde E_{2D} , E_{2D3D} , E_{depth} , E_{temp} , E_{fseg} , E_{reg} son términos de regularización L2 y W_{2D} , W_{2D3D} , W_{depth} , W_{temp} , W_{fseg} , W_{reg} son los pesos escalares correspondientes.

También se siguió a Zhou et al. [34] (en términos de inicialización del rastreo); se inicializa el siguiente cuadro de forma facial 2D basado en la correspondencia encontrada por un emparejador robusto de características locales entre ese cuadro y el cuadro previo. Esto mejora la estabilidad en el rastreo de movimientos faciales rápidos y reduce el número de iteraciones de Gauss-Newton requeridas para minimizar la fórmula 3.1 [33].

Estos algoritmos de aprendizaje mejoran la precisión del reconocimiento facial con cada nuevo fotograma que envía la cámara. El punto óptimo de dicho aprendizaje se alcanza cuando se han recibido unos 300 fotogramas, lo que equivale a 26 segundos con la resolución escogida en el sistema.

Cuando existen esqueletos activos con información válida de sus rastreadores faciales (rostro visible correctamente), el sistema permite apreciar visualmente el seguimiento facial realizado superponiendo los 121 puntos reconocidos con la

imagen RGB captada por *Kinect*® (ver Fig. 3.3). Esto permite apreciar de forma sencilla si se está produciendo un seguimiento facial correcto, o si se deben ajustar las condiciones de adquisición de información.

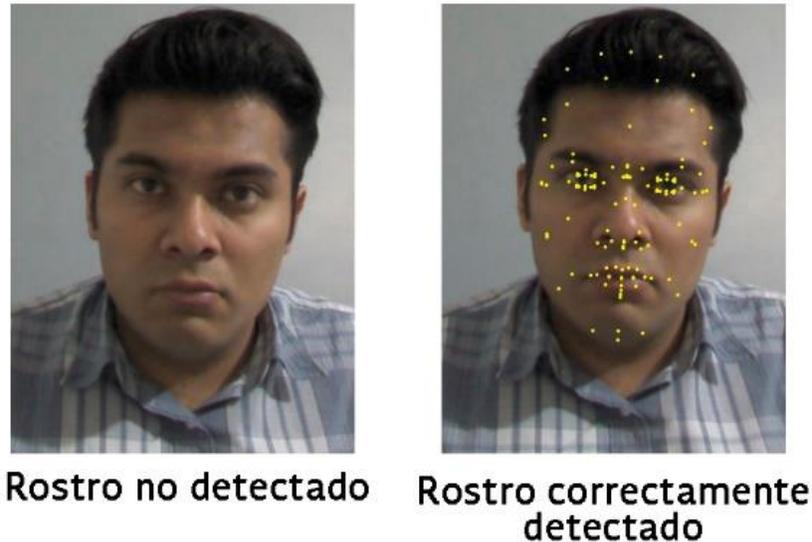


Fig. 3.3 Confirmación de detección del rostro

3.3.3 Discriminación de la información facial y visual de la persona

Una vez visualizados sobre la imagen RGB los 121 puntos, es necesario seleccionar sólo aquellos cuya ubicación coincida lo más cercanamente posible con los ocho puntos de interés del sistema de clasificación Nottingham (ver Fig. 2.5). Para realizar dicha tarea, primero fue necesario mostrar el índice de todos los puntos sobre la imagen RGB, esto con el fin de facilitar la identificación del índice de aquellos que se consideren tienen una ubicación similar o igual a los puntos utilizados en el sistema de clasificación. La Fig. 3.4 muestra los índices en color blanco de los 121 puntos reconocibles. La identificación correcta de los índices de los puntos seleccionados es importante pues se usarán posteriormente para obtener sus coordenadas.



Fig. 3.4 Visualización de los índices de los puntos faciales reconocibles

Los índices de los puntos que fueron considerados como candidatos al tener una ubicación similar a los puntos del sistema de clasificación son: 18, 51, 20, 53, 27, 60, 31 y 64. Una vez que se han discriminado los ocho puntos de interés, resulta innecesario mostrar el resto en la imagen RGB. Por esta razón, el código siguiente muestra la modificación al método *DrawFaceModel* visto en el capítulo anterior, para que sólo estos índices se almacenen en la lista de puntos *faceModelPts* y así sean sólo éstos los que se muestren sobre la imagen (ver Fig. 3.5).

```
public void DrawFaceModel(DrawingContext drawingContext)
{
    if (!this.lastFaceTrackSucceeded || this.skeletonTrackingState !=
        SkeletonTrackingState.Tracked)
    {
        facepointsReady = false;
        return;
    }
    facepointsReady = true;
}
```

```
var faceModelPts = new List<Point>();
// Puntos de interés para el sistema
var faceNottinghamPoints = new List<int> {18,51,20,53,27,60,31,64};
for (int i = 0; i < this.facePoints.Count; i++)
{
    // Agrega sólo los puntos de interés
    if (faceNottinghamPoints.Contains(i))
    {
        // Ajuste para 1280 x 920
        faceModelPts.Add(new Point( (this.facePoints[i].X )/2,
        (this.facePoints[i].Y )/2 ));
    }
}
var faceModelGroup = new GeometryGroup();
for (int i = 0; i < faceModelPts.Count; i++) //faceModel
{
    var facePoints = new GeometryGroup();
    var ellipseGeometry = new EllipseGeometry(faceModelPts[i], 1, 1);
    facePoints.Children.Add(ellipseGeometry);
    faceModelGroup.Children.Add(facePoints);
}
// Dibuja los puntos
drawingContext.DrawGeometry(Brushes.Yellow, new Pen(Brushes.Yellow, 1.0),
faceModelGroup);
}
```



Fig. 3.5 Visualización de sólo los ocho puntos de interés

3.4 Almacenamiento de los datos

En esta fase del sistema se guarda tanto la información facial como visual del paciente al momento de realizar alguna de las cuatro expresiones faciales establecidas. Para esto se utilizarán dos tipos de archivos; el primero, es un archivo de texto que contiene la información facial necesaria para realizar el diagnóstico objetivo por medio del sistema de clasificación Nottingham. Dentro del archivo se guardan las distancias específicas entre pares de puntos de interés. La Fig. 3.6 muestra la relación entre los puntos que forman las distancias que utiliza el sistema para evaluar el nivel de movilidad facial.

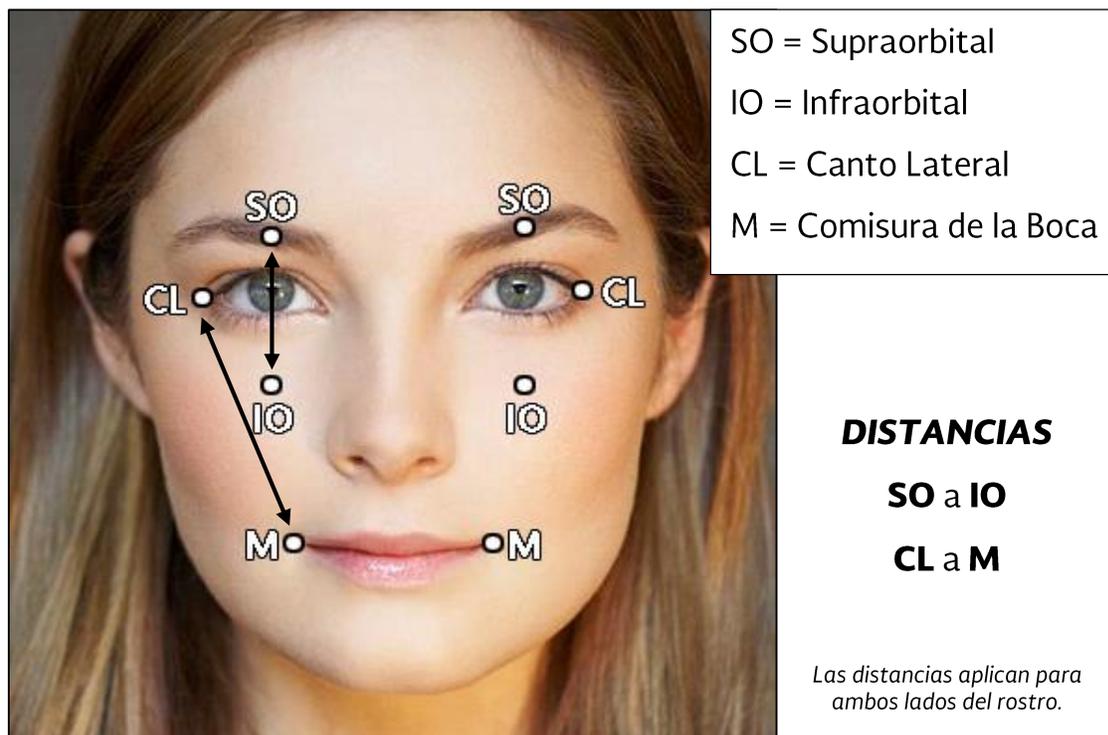


Fig. 3.6 Distancias entre los puntos de interés

El segundo archivo, es una fotografía del paciente, ésta sirve para complementar su expediente y llevar un registro visual del avance del paciente. La fotografía es guardada en un formato RGB con una resolución de 1280x920 píxeles.

Ambos archivos son creados simultáneamente en la misma carpeta y con el mismo nombre que es designado desde el sistema por defecto. En el sistema de clasificación Nottingham se requiere que el paciente realice cuatro expresiones faciales específicas, para diferenciar cada una de ellas se agregó un índice al final del nombre del archivo (ver Tabla 3.1).

Tabla 3.1 Índices de expresiones faciales

Expresión facial	Índice
En reposo	00
Levantando las cejas	01
Cerrando estrechamente los ojos	02
Sonriendo	03

El formato propuesto para nombrar a los archivos es presentado en la Fig. 3.7. Este formato fue definido de manera sencilla, pensando en incorporar sólo los elementos necesarios que permitan una identificación rápida por parte del médico

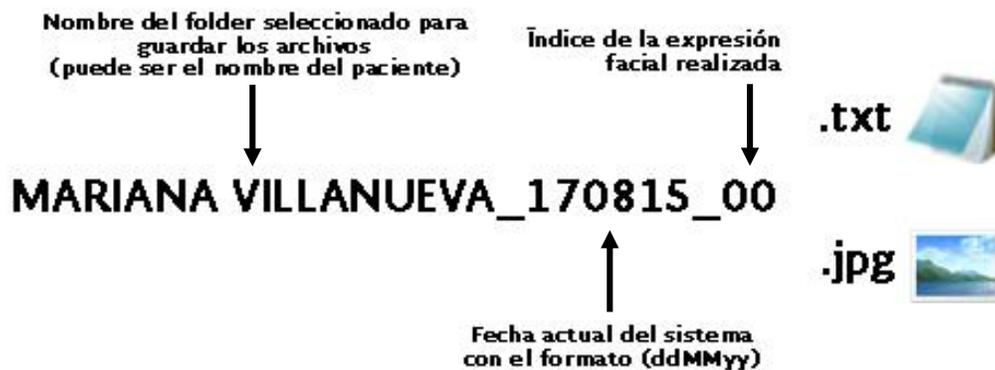


Fig. 3.7 Formato del nombre de los archivos

El código a continuación, describe la función *captureImage_Click*, esta función se encuentra asociada al evento *click* del botón “Captura” de la interfaz y es la responsable de crear los archivos cada vez que se invoca. En el código se observa la composición del formato del nombre de los archivos.

```

private void captureImage_Click(object sender, RoutedEventArgs e)
{
    if (facialExpression.SelectedIndex > -1)
    {
        switch (facialExpression.Text)
        {
            case "En reposo":
                FaceTrackingViewer.fen = 00;
                break;
            case "Levantando cejas":
                FaceTrackingViewer.fen = 01;
                break;
            case "Cerrando ojos":
                FaceTrackingViewer.fen = 02;
                break;
            case "Sonriendo":
                FaceTrackingViewer.fen = 03;
                break;
        }
        string nameFile = pathName + "/" + folderName + "_" +
            DateTime.Now.ToString("ddMMyy") + "_" + FaceTrackingViewer.fen + ".jpg";
        nameFiletxt = pathName + "/" + folderName + "_" +
            DateTime.Now.ToString("ddMMyy") + "_" + FaceTrackingViewer.fen;
        using (FileStream stream = new FileStream(nameFile, FileMode.Create))
        {
            JpegBitmapEncoder encoder = new JpegBitmapEncoder();
            encoder.Frames.Add(BitmapFrame.Create(this.colorImageWritableBitmap));
            encoder.Save(stream);
        }
    }
}

```

3.5 Evaluación de los datos

En esta última fase se calcula un porcentaje que representa el nivel de simetría facial que tiene una persona, donde el 100% o valores muy cercanos a él representan un rostro sano (sin parálisis). Este valor se obtiene mediante la evaluación de los datos faciales que se encuentran en los archivos de textos generados en la fase anterior. Por tal motivo, la interfaz del sistema le permite al médico seleccionar los cuatro archivos de texto referentes a las cuatro expresiones faciales que se le pidió al paciente realizar. La Fig. 3.8 [35] muestra ejemplos visuales de las expresiones faciales que el paciente debe realizar y resalta en rojo el área de los músculos mayormente involucrados en el movimiento que se esté realizando. El rostro en reposo es tomado como referencia en términos de movilidad; las distancias entre los puntos al realizar algún movimiento son comparadas con las distancia de esos puntos pero con el rostro en reposo.

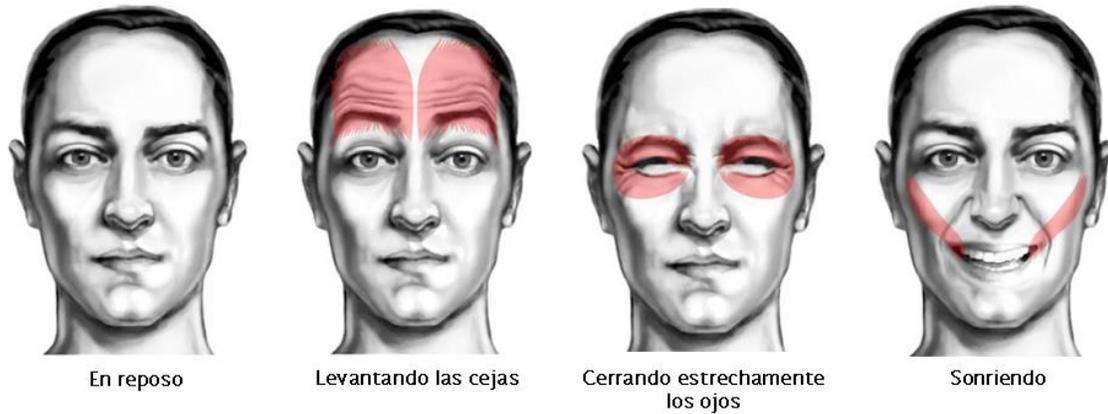


Fig. 3.8 Expresiones faciales que debe realizar el paciente en el sistema

Una vez seleccionados los archivos, el sistema extrae los valores de las distancias de cada uno de ellos y sigue el procedimiento clásico del sistema de clasificación Nottingham. La Fig. 3.9 [36] muestra la guía utilizada en las clínicas para diagnosticar la parálisis facial siguiendo el sistema de clasificación Nottingham. En la primera mitad de la guía, se muestra una imagen con la disposición de los ocho puntos de interés y las cuatro distancias importantes entre ellos, en la segunda mitad, se visualiza el procedimiento para calcular el porcentaje de simetría.

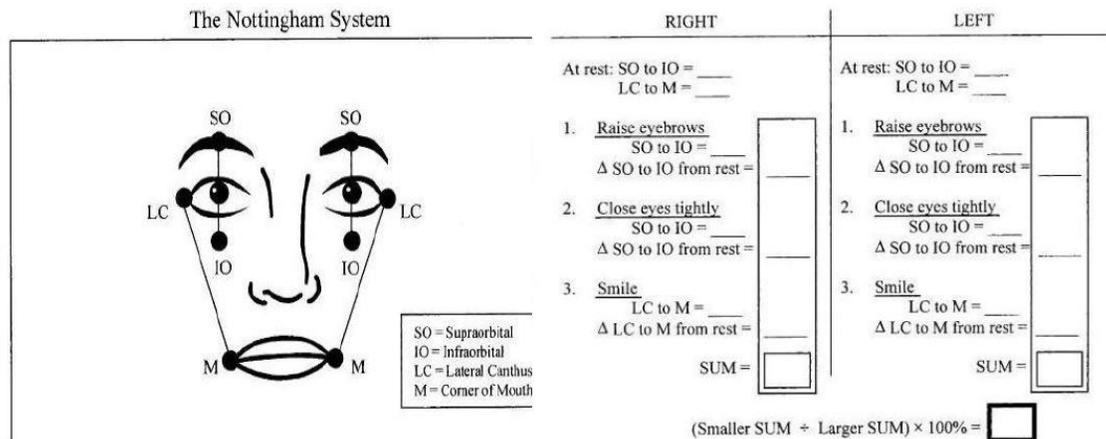


Fig. 3.9 Guía para el diagnóstico de parálisis facial mediante el sistema de clasificación Nottingham

El procedimiento descrito en la guía fue codificado en la función *diagnostic_Click*, esta función se encuentra asociada al evento *click* del botón “Diagnóstico” de la interfaz.

```
private void diagnostic_Click(object sender, RoutedEventArgs e)
{
    double rightSide = 0.0;
    double leftSide = 0.0;
    double diagnosticResult = 0.0;
    OpenFileDialog ofd = new OpenFileDialog();
    ofd.Multiselect = true;
    ofd.Filter = "Text|*.txt|All|*.*";
    ofd.ShowDialog();
    // Distancias de rostro en reposo
    double l_S02IOr = 0.0;
    double l_LC2Mr = 0.0;
    double r_S02IOr = 0.0;
    double r_LC2Mr = 0.0;
    // Distancias de rostro al realizar una expresión
    double l_S02IO, l_LC2M, r_S02IO, r_LC2M = 0.0;
    // Variable que confirma que exista un archivo del rostro en reposo
    Boolean rest_face = false;
    string[] result = ofd.FileNames;
    foreach (string y in result)
    {
        string index_file = y.Substring(ofd.FileName.Length-6,2);
        StreamReader sr = new StreamReader(y);
        if (index_file == "00") // Existe el archivo del rostro en reposo
        {
            l_S02IOr = Convert.ToInt32(sr.ReadLine());
            l_LC2Mr = Convert.ToInt32(sr.ReadLine());
            r_S02IOr = Convert.ToInt32(sr.ReadLine());
            r_LC2Mr = Convert.ToInt32(sr.ReadLine());
            rest_face = true;
        }
    }
    if (rest_face)
    {
        foreach (string y in result)
        {
            double l_diff, r_diff = 0.0;
            string index_file = y.Substring(ofd.FileName.Length-6,2);
            StreamReader sr = new StreamReader(y);
            if (index_file == "01") // Levantando cejas
            {
                l_S02IO = Convert.ToInt32(sr.ReadLine());
                r_S02IO = Convert.ToInt32(sr.ReadLine());
                l_diff = Math.Abs(l_S02IO - l_S02IOr);
                r_diff = Math.Abs(r_S02IO - r_S02IOr);
                leftSide += l_diff;
                rightSide += r_diff;
            }
            if (index_file == "02") // Cerrando ojos
            {
                l_S02IO = Convert.ToInt32(sr.ReadLine());
                r_S02IO = Convert.ToInt32(sr.ReadLine());
            }
        }
    }
}
```

```

        l_diff = Math.Abs(l_S02IO - l_S02IOr);
        r_diff = Math.Abs(r_S02IO - r_S02IOr);
        leftSide += l_diff;
        rightSide += r_diff;
    }
    if (index_file == "03") // Sonriendo
    {
        l_LC2M = Convert.ToInt32(sr.ReadLine());
        r_LC2M = Convert.ToInt32(sr.ReadLine());
        l_diff = Math.Abs(l_LC2M - l_LC2Mr);
        r_diff = Math.Abs(r_LC2M - r_LC2Mr);
        leftSide += l_diff;
        rightSide += r_diff;
    }
}
// Cálculo del diagnóstico
diagnosticResult = (Math.Min(rightSide, leftSide) / Math.Max(rightSide,
leftSide)) * 100;
}
rest_face = false;
}

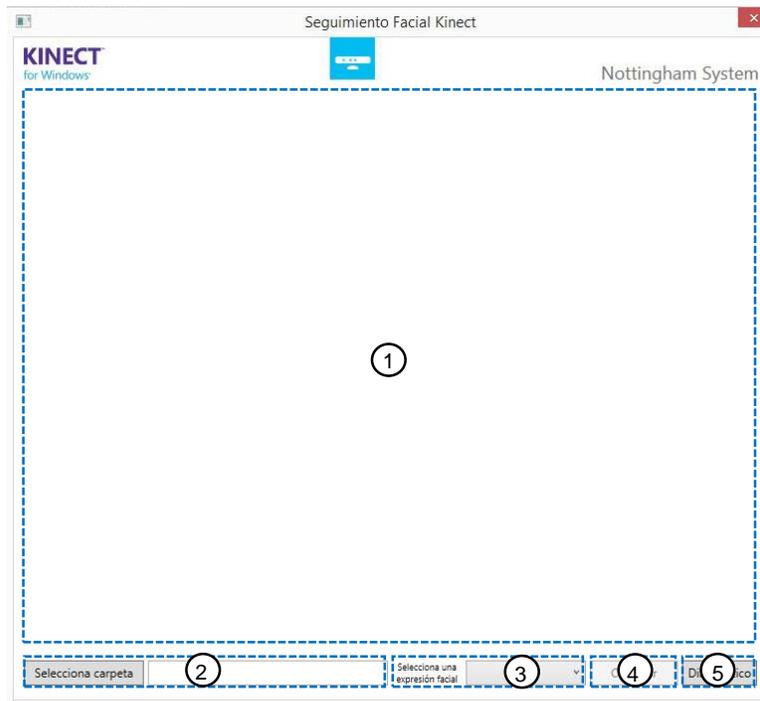
```

Como se puede observar en el código anterior, el sistema de clasificación Nottingham se basa fundamentalmente en la comparación de movilidad entre ambos lados del rostro. Esto quiere decir que en un rostro “normal” se espera que la movilidad facial se encuentre presente en igual proporción en ambos lados, mientras que en una persona con parálisis facial periférica existiría una diferencia de movilidad entre ambas partes del rostro, en donde el lado con mayor movilidad es considerado como el lado sano o el lado “normal”. Este lado es aquel que el sistema toma como referencia para evaluar el desempeño del lado afectado, ya que no todas las personas presentan una extensión en los movimientos de la cara uniforme, por lo que es necesario evaluarlas utilizando un parámetro individual.

En el caso de las personas sanas se esperan que los valores de movilidad en ambos lados sean iguales, sin embargo, es usual encontrar una diferencia muy pequeña, esto se debe a una asimetría facial leve la cual es absolutamente normal; es común tener un lado de la cara mayor que el otro, tanto a lo ancho como en lo alto, tener una ceja más alta, un párpado más abajo, la nariz o los labios algo desviados, etc. Estas asimetrías se deben a diferencias en el crecimiento de las estructuras óseas de cada lado y, en ocasiones, a la mayor o menor expresividad de cada lado de la cara.

3.6 Descripción de la interfaz de usuario

La interfaz de usuario se divide en cinco partes. A continuación se describen el funcionamiento secuencial que tienen cada una de estas partes (ver Fig. 3.10).



1. Lienzo principal
2. Selección de la carpeta de almacenamiento
3. Selección de la expresión facial
4. Captura de la imagen
5. Diagnóstico

Fig. 3.10 Partes que compone la interfaz

3.6.1 Lienzo principal

Se presenta en la parte central de la interfaz y muestra la imagen en tiempo real que capta la cámara del *Kinect*®, de no ser así es posible que exista algún problema de comunicación entre el *Kinect*® y la computadora. Una vez que se detecte la imagen, el paciente debe ser colocado siguiendo las especificaciones técnicas del dispositivo, vistas en el capítulo anterior. Para esto, se recomienda crear un escenario fijo, donde no se modifiquen elemento del entorno como la luz o la ubicación de la cámara o del paciente, esto con el fin de disminuir el número de variables que puedan alterar los resultados. En este escenario, se debe dar preferencia a la ubicación del rostro, tratando de colocarlo en el centro del lienzo para que sea lo más visiblemente posible.

3.6.2 Selección de la carpeta de almacenamiento

Consta del botón “Seleccionar carpeta” que al hacer click, despliega un cuadro de dialogo donde el médico debe seleccionar la carpeta donde desea guardar los archivos con la información facial y visual del paciente. Como sugerencia se propone que el nombre de la carpeta sea único y representativo de cada paciente. Una vez seleccionada la carpeta, la ubicación de ésta será mostrada en el cuadro de texto que se encuentra a lado derecho del botón “Seleccionar carpeta”.

3.6.3 Selección de la expresión facial

Para seleccionar la expresión facial que se va a capturar se cuenta con un cuadro de lista desplegable. Este cuadro contiene un listado con las cuatro expresiones faciales que se requiere que el paciente realice. Es importante prestar atención a esta selección para generar los archivos correctamente y no haya duplicidad.

3.6.4 Captura de imagen

El botón “Capturar” se encuentra por defecto deshabilitado, éste se habilita hasta que el médico haya seleccionado una carpeta de almacenamiento y el sistema haya reconocido un rostro. Una vez que se da click sobre el botón, se muestra un mensaje indicando que la captura fue realizada correctamente y por consiguiente, la creación de los archivos con la información facial y visual del paciente fue exitosa.

3.6.5 Diagnóstico

Para utilizar el botón “Diagnóstico” de la última parte de la interfaz, es necesario contar antes con los archivos generados en la parte de *Captura de imagen*. Al dar click sobre el botón, se despliega un cuadro de dialogo, donde el médico debe seleccionar sólo los 4 archivos relacionados a las expresiones faciales que hizo un paciente en un día determinado. Sí los archivos seleccionados son correctos, se mostrará un mensaje con el porcentaje de simetría facial que presenta la persona. El botón “Diagnóstico” permanece siempre habilitado, lo que permite al médico acceder a esta parte del sistema de manera automática, sin necesidad de capturar nuevamente las imágenes del rostro de un paciente y, de esta manera; realizar el cálculo del nivel de parálisis sobre los archivos ya existentes las veces que quiera.

La interfaz cuenta con todas las restricciones necesarias para que cualquier médico con poca experiencia en el uso de sistemas de cómputo pueda operarla sin dificultades.

CAPÍTULO 4: PRUEBAS Y CONCLUSIONES

4.1 Introducción

En este último capítulo se describen con detalle las pruebas que se realizaron, esto con el propósito de demostrar la efectividad y funcionalidad del sistema mediante el análisis de los resultados obtenidos después de realizar el procedimiento. Este análisis de pruebas permitió percibir los límites del sistema, pero sobretodo las mejoras que se pueden realizar en él, ya que el constante avance de la tecnología permite actualizar componentes del sistema como el *Kinect®*, que sin duda mejorarían significativamente su rendimiento. Para finalizar el capítulo se dan conclusiones sobre el trabajo en general.

4.2 Análisis de pruebas

Se realizaron dos pruebas diferentes con propósitos específicos. La primera prueba se hizo para corroborar la consistencia del resultado en el proceso de diagnóstico, la segunda prueba consistió en la aplicación del procedimiento de diagnóstico usando el sistema en personas que tienen cierto nivel de parálisis facial. Ambas pruebas contribuyen en la validación de los resultados obtenidos por el sistema. El proceso para el desarrollo de cada una de ellas es descrito a continuación.

4.2.1 Consistencia en el diagnóstico

Para comprobar la solidez de los resultados obtenidos durante el proceso de diagnóstico por parte del sistema, se ideó calcular el porcentaje de simetría facial de sólo una persona sana por un lapso de 10 días, con el fin de demostrar que todos los resultados obtenidos son iguales o muy similares y las diferencias que llegan a existir entre ellos es debido, en su mayoría, por variables como la iluminación o el posicionamiento del rostro. Gracias a esta prueba, se detectó que la presencia de barba en abundancia en el sujeto, pueden llegar a interferir en la ubicación correcta de algunos de los puntos de interés del sistema. Sin embargo, este factor no representa un obstáculo significativo en la aplicación del procedimiento, por lo que no es necesario afeitar por completo el rostro, pero se obtienen resultados más precisos. La Fig. 4.1 muestra una tabla con los resultados obtenidos de esta prueba.

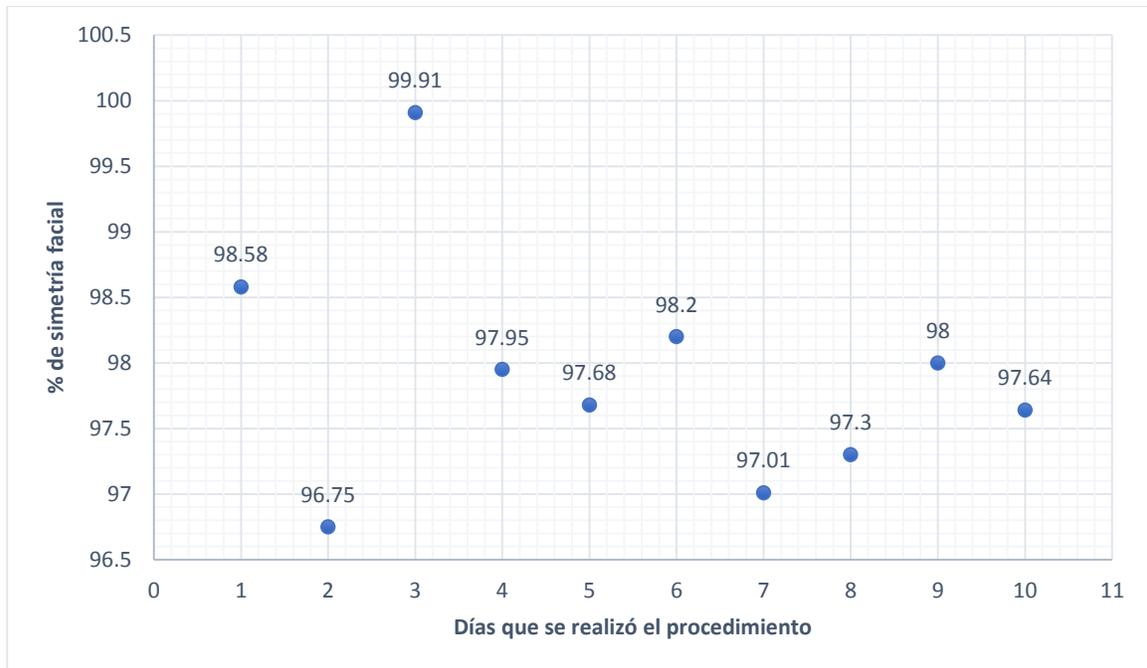


Fig. 4.1 Resultados de la prueba de consistencia

Como se puede observar en la figura anterior, todos los valores son similares entre sí; su media aritmética es de 97.60 y tienen una desviación estándar de sólo 1.05. Con los resultados obtenidos se puede asegurar que la prueba cumplió su objetivo al demostrar que los valores presentan sólo una pequeña variabilidad esperada debido a variables externas descritas con anterioridad. Los porcentajes altos obtenidos confirman la simetría natural de un rostro sano y del mismo modo evidencian la pequeña asimetría natural del rostro explicada anteriormente, esto confirma una gran solidez por parte del sistema.

4.2.2 Aplicación del sistema en pacientes con parálisis facial

Tal como se observó en la prueba anterior, el nivel de parálisis facial puede concebirse como el grado de simetría en los movimientos que se puede alcanzar en ambos lados del rostro. En este sentido, el sistema de clasificación Nottingham expresa este grado o nivel en forma de porcentaje, el cual se puede definir como el cociente del parámetro de movilidad del lado afectado entre el parámetro de movilidad del lado sano.

En otras palabras, se considera que el paciente se recupera de la parálisis cuando ambos lados del rostro (afectado y sano) adquieren un nivel de movilidad igual o muy similar entre sí. La determinación de dicho nivel de simetría facial es un reto, ya que es diferente para cada individuo, esto se debe a que no todas las personas con parálisis presentan un mismo grado de movilidad en su lado sano del rostro. En la práctica tradicional esto resulta muy complicado; ya que tal como lo muestra un estudio reciente, se determinó que por lo menos se debe tener 3 milímetros de asimetría facial en la comisura de la boca, en la supraorbital o en ambas partes del rostro para que una persona pueda llegar a percibir la asimetría [37].

Para la realización de esta prueba se acudió a la Unidad Básica de Rehabilitación (URB) ubicada en el municipio de Panotla. La UBR es una extensión en servicios del Sistema Municipal DIF la cual proporciona atención integral a personas con discapacidad temporal o permanente, por medio de:

- Educación para la Salud.
- Valoración médica.
- Rehabilitación a través de diferentes modalidades terapéuticas.
- Control y seguimiento de pacientes
- Capacitación y orientación a la familia.
- Referencia a otros niveles de atención en aquellos casos que por su complejidad no puedan ser atendidos en la UBR.

Uno de los principales tipos de discapacidad que se atienden en la Unidad son las secuelas de daño neurológico (parálisis facial), por lo que se solicitó realizar pruebas del sistema en pacientes con parálisis facial de la Dra. Graciela García Pérez quien funge como responsable de esta Unidad. Por motivos de confidencialidad sólo 2 personas dieron autorización para utilizar su información y datos obtenidos. La Fig. 4.2 muestra las evidencias del funcionamiento del sistema en dos de los pacientes; la señora Lucina Juárez de 75 años y del señor Roberto Ordoñez de 51 años, ambos presentan parálisis facial en el lado derecho del rostro.

Paciente: Lucina Juárez



En reposo



Levantando las cejas



Cerrando estrechamente los ojos



Sonriendo

Paciente: Roberto Ordoñez



En reposo



Levantando las cejas



Cerrando estrechamente los ojos



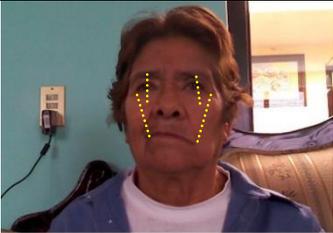
Sonriendo

Fig. 4.2 Fotografías de los pacientes realizando las cuatro expresiones faciales

A continuación se describen los pasos realizados para el cálculo del porcentaje de simetría facial siguiendo el sistema de clasificación Nottingham.

1. Una vez que se verifique el correcto posicionamiento de los puntos faciales a través de la interfaz, el médico procede a capturar las imágenes de las 4 expresiones faciales requeridas para la obtención de las distancias (Fig. 3.6) entre los puntos característicos (Tabla 4.1).

Tabla 4.1 Obtención de distancias entre los puntos característicos

Fotografías	Distancia	Lado izquierdo	Lado Derecho
 En reposo	SO a IO CL a M	42.64 61.34	47.96 72.56
 Levantando las cejas	SO a IO	56.57	47.71
 Cerrando estrechamente los ojos	SO a IO	44.01	45.48
 Sonriendo	CL a M	59.29	68.81

- Una vez obtenidas las distancias en cada expresión facial, se calculan las diferencias entre las distancias presentes del rostro en reposo contra las contenidas en las demás expresiones faciales. Esto debido a que el rostro en reposo sirve como referencia para determinar el nivel de movimiento presente. Al conseguir todos los valores, éstos se suman para cada lado; obteniendo así los valores de movilidad (Tabla 4.2).

Tabla 4.2 Cálculo del valor de movilidad para ambos lados del rostro

Lado izquierdo		Lado derecho	
SO a IO <i>En reposo</i>	SO a IO <i>Levantando las cejas</i>	SO a IO <i>En reposo</i>	SO a IO <i>Levantando las cejas</i>
42.64 – 56.57 = 13.93		47.96 – 47.71 = 0.25	
SO a IO <i>En reposo</i>	SO a IO <i>Cerrando estrechamente los ojos</i>	SO a IO <i>En reposo</i>	SO a IO <i>Cerrando estrechamente los ojos</i>
42.64 – 44.01 = 1.37		47.96 – 45.48 = 2.48	
CL a M <i>En reposo</i>	CL a M <i>Sonriendo</i>	CL a M <i>En reposo</i>	CL a M <i>Sonriendo</i>
61.34 – 59.29 = 2.05		72.56 – 68.81 = 3.75	
13.93 + 1.37 + 2.05 = 17.35		0.25 + 2.48 + 3.75 = 6.48	

Tras realizar las operaciones necesarias se obtiene un valor de movilidad de 17.35 para el lado izquierdo y de 6.48 para el lado derecho, estos valores reflejan la diferencia que visiblemente es percibida.

- El último paso consiste en dividir el valor más pequeño entre el valor más grande, dando como resultado 0.37 que multiplicado por 100 nos da el porcentaje de asimetría facial.

Todo el procedimiento detallado anteriormente es realizado internamente por el sistema, el cual como se observó en el cálculo desglosado dio un valor de 37.46% de simetría facial en la Sra. Lucina y de 78.41% en el Sr. Roberto, este resultado fue corroborado por la Dra. Graciela, la cual aprobó el resultando validando de esta

manera el correcto funcionamiento del sistema, al ser un valor que representa con certeza el nivel de parálisis.

La extensión del daño nervioso determina el alcance de la recuperación, la mejoría es gradual y los tiempos de recuperación varían. Con o sin tratamiento, la mayoría de los individuos comienza a mejorar dentro de las 2 semanas del inicio de los síntomas y la mayoría se recupera completamente, regresando a su función normal dentro de los 3 a 6 meses. Para algunos, sin embargo, los síntomas pueden durar más tiempo, dependiendo de diferentes variables. La Fig. 4.3 muestra una gráfica que representa el pronóstico de la Dra. Graciela en la recuperación de sus pacientes. Como se puede observar, se espera una mejora en ambos casos, sin embargo el factor de la edad influye en el tiempo y grado de recuperación.

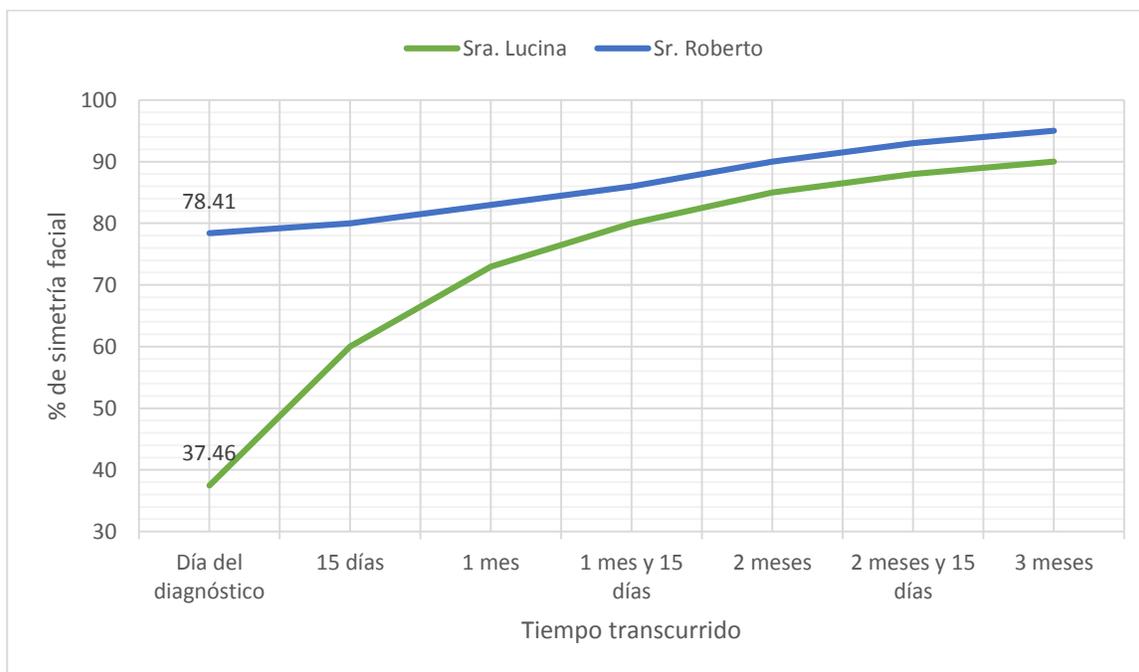


Fig. 4.3 Recuperación esperada por el médico

Gracias al sistema, el médico ahora cuenta con una herramienta que le permite evaluar no sólo el progreso de sus pacientes; sino sus propios pronósticos de recuperación, dando como resultado una mejora en su práctica médica.

4.3 Conclusiones

Del desarrollo del presente trabajo de tesis y las pruebas realizadas se pueden concluir lo siguiente:

Actualmente la tecnología es un factor vital para el desarrollo y evolución de una sociedad con anhelo de crecimiento. La tecnología se ha convertido en un componente esencial en todas las áreas, pero se hace imprescindible en el área de la salud. Hoy por hoy, se llegan a utilizar los avances tecnológicos no sólo para curar sino también para prevenir enfermedades; y posteriormente para todo tipo de investigación médica, la cual gracias a la tecnología ha realizado importantes descubrimientos. Con el trabajo realizado se comprobó la multifuncionalidad del *Kinect*®; la variedad en sus características permitió modelar una solución al problema de manera objetiva y sencilla, demostrando ser una herramienta multiusos en más de un área.

Los impactos sociales y científicos de este sistema son sólo una pequeña parte de lo que el sistema será capaz de contribuir a la comunidad médica científica. La comunicación interdisciplinaria entre los campos de la medicina será capaz de usar un lenguaje objetivo estandarizado común para hablar sobre parálisis facial. Este lenguaje proporcionará diagnósticos objetivos, lo que lleva a mejores resultados y una mejor calidad de vida de los pacientes con deterioro de la función facial.

Un sistema estandarizado, objetivo y automático de análisis de video no sólo beneficiaría la cirugía plástica y reconstructiva, sino también a la otorrinolaringología, oftalmología, neurocirugía y la cirugía oral y maxilofacial. No sólo los pacientes con parálisis facial, sino también pacientes con malformaciones congénitas faciales (por ejemplo, labio leporino y paladar), deformidades faciales postraumático, enfermedades degenerativas y autoinmunes, o deformidades presentadas después de la cirugía de tumor se beneficiarían de un mejor análisis de forma facial y función.

Si bien, los médicos no tienen dificultad al realizar el proceso de diagnóstico de parálisis facial de la manera tradicional, la mayor aportación del sistema es la certeza de contar con resultados objetivos, respaldados por datos cuantitativos.

Además de complementar el expediente clínico de un paciente mediante la adición de un registro temporal gráfico, todo lo anterior manejando una interfaz simple que ahorra significativamente el tiempo dedicado en el procedimiento de diagnóstico.

En general el sistema desarrollado resulta ser una herramienta útil, simple y accesible, capaz de brindar un resultado objetivo para cualquier médico que guste corroborar su diagnóstico y ampliar el expediente de sus pacientes. El uso del sistema de clasificación Nottingham como escala del nivel de parálisis facial demostró ofrecer un refinamiento más allá del disponible con la escala mayormente usada (House-Brackmann). Sin duda alguna, la automatización de este sistema clasificatorio resultó de manera natural; las diferentes características del *Kinect*® brindaron una respuesta a cada cuestionamiento que surgía durante la realización del trabajo, el cual fue satisfactorio realizar.

Sin duda alguna el sistema desarrollado muestra sólo una pequeña parte del potencial que tiene el *Kinect*® en el análisis facial. El sistema puede ser modificado para realizar tareas específicas en diferentes áreas como:

- Reconocimiento de expresiones faciales.
- Manejo de interfaces humano computadora inteligentes.
- Identificación de emociones.
- Clasificación de razas a partir de características faciales.
- Análisis de la fisonomía.

4.4 Trabajo a futuro

Si bien se cumplieron satisfactoriamente todos los objetivos propuestos en el trabajo, con la salida reciente de la nueva versión del *Kinect*®, se evidenciaron mejoras que se pueden implementar posteriormente para optimizar aún más los resultados obtenidos. El *Kinect*® v2 lleva la interacción humano computadora a un nivel superior, que ofrece una mejor precisión, capacidad de respuesta y capacidades intuitivas para acelerar el desarrollo de aplicaciones que responden al movimiento, gestos y la voz. La cámara de color de la nueva versión se ha mejorado con video a 1080 píxeles que se pueden mostrar en la misma resolución que la

pantalla de visualización. Además ahora se puede hacer seguimientos de hasta 6 personas y de 25 articulaciones del cuerpo por persona, las posiciones rastreadas son anatómicamente correctas y estables y el rango de seguimiento es más amplio. La mayor fidelidad en profundidad hace que sea mucho más fácil de ver objetos más pequeños, ver todos los objetos más claramente, y visualizar objetos en tres dimensiones (3D). La Fig.4.4 muestra estas mejoras significativas de fidelidad en profundidad comparando el sensor original (lado izquierdo) con el sensor de la versión 2 (lado derecho).



Fig. 4.4 Comparación entre cámaras del modelo original (v1) y el modelo reciente (v2)

En general, se puede afirmar que el sensor *Kinect*® v2 cuenta con importantes mejoras en comparación con el modelo anterior, pero considerando el trabajo que se desarrolló, lo que hace llamar más la atención es que esta nueva versión ofrece una resolución 20 veces mayor, permitiéndole a la aplicación crear una malla de más de 1000 puntos para una representación más exacta de la cara de una persona.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Banta H. D. & Behney C. J. (1981). Policy Formulation and technology assessment. *Milbank Mem Fund Quartely*, 59(3), 445-79.
- [2] Lewis T. (1995). *The lives of a cell: Notes of a biology watcher*. Penguin Books.
- [3] Retting R. A. & Hartman A. J. (1979). The development of medical technology: a policy perspective. *Medical Technology*. National Center for Health Services Research. Research Proceedings Series.
- [4] Colegio Oficial de Fisioterapeutas de Navarra. (2015). Información y atención al cuidado. Pamplona, España. Recuperado el 25 de junio del 2015, de <http://www.cofn.net>.
- [5] Dong J., Wang Q., Wang S. & Liu L. (2011). Evaluation of Facial Paralysis Degree, Face Analysis, Modeling and Recognition Systems. *Five*, 79-90.
- [6] Tzou C. J., Pana I., Placheta E., Hold A., et al. (2012). Evolution of 3D Surface Imaging System in Facial Plastic Surgery. *69(2)*, 173-185.
- [7] Wu Z., Silverman C., Linstrom C., et al. (2005). Objective computerized versus subjective analysis of facial synkinesis. *115*, 2118-2122.
- [8] Gross M., Trotman C. & Moffatt K. (1996). A comparison of three-dimensional and two-dimensional analyses of facial motion. *Angle Orthod.* 66, 189-194.
- [9] Frey M. (1999). Smile reconstruction using the gracilis muscle. *Oper Tech Plast Reconstr Surg.* 6, 180-189.
- [10] Frey M., Giovanoli P., Gerber M., et al. (1999). Three-dimensional video analysis of facial movements: A new method to assess the quantity and quality of the smile. *Plastic and Reconstructive Surgery.* 104(7), 2032-2039.
- [11] Deutsch S. & Deutsch A. (1993). *Understanding the Nervous System: An Engineering Perspective*. Wiley-IEEE Press. New York. 394.

- [12] Molist P., Pombal M. A. & Mergías M. (2013). Atlas de histología vegetal y animal. Órganos animales. Sistema nervioso. Depto. De Biología Funcional y Ciencias de la Salud. Facultad de Biología. Universidad de Vigo. España, 26-29.
- [13] Calle M. L. & Casado I. (2011). Exploración de los pares craneales. Recordatorio de semiología. Unidad de Ictus. Sección de Neurología. Hospital San Pedro de Alcántara. Cáceres. España. 65-71.
- [14] Rivas I. (2007). Gastroenterología – Medicina Interna. Instituto de Corazón y Vasos. Mérida, Venezuela. Recuperado el 3 de enero del 2015, de <http://www.gastromerida.com>.
- [15] Cacho J., Cacabelos P & Sevillano M. D. (2011). Patología de los pares craneales. Servicio de Neurología. Hospital Universitario de Salamanca. Salamanca. España. 4799.
- [16] Rodríguez M. D., Mangas S., Ortiz M. G., Rosete H. S., Vales O. & Hinojosa R. (2011). Parálisis facial periférica. Tratamientos y consideraciones. *Arch Neurocién (Mex) INNN*. 16(3). 148-155.
- [17] Kiziltan M., Akalin M., Rahsan S. & Uluduz D. (2007) Peripheral neuropathy in patients with diabetes *mellitus* presenting as Bell's palsy. *Neurosci Lett*. 427,138-141.
- [18] Tiemstra J. & Khatkhate N. (2007). Bell's palsy: diagnosis and management. *Am Fam Physician*. 76(7), 997-1002.
- [19] Shapiro J., Yudin M. & Ray J. (1999). Bell's palsy and tinnitus during pregnancy: predictors of preeclampsia? Three cases and a detailed review of the literature. *Acta Otolaryngol*. 119. 647-651.
- [20] Tazi M., Soichot P. & Perrin D. (2003). Facial palsy following dental extraction: Report of two cases. *J Oral Maxillofac Surg*. 61, 840-844.

- [21] May M. (1970). Facial paralysis, peripheral type: a proposed method of reporting. (Emphasis on diagnosis and prognosis, as well as electrical and chorda tympani nerve testing). *Laryngoscope*. 80, 331-390.
- [22] Brenner M. J. & Gail J. (2004). Approaches to Grading Facial Nerve Function. *Seminars in Plastic Surgery*. 18(1), 13-21.
- [23] House J. W. (1983) Facial nerve grading systems. *Laryngoscope*. 93, 1056-1069.
- [24] Stennert E. (1977). Facial nerve paralysis scoring system. In: *Facial Nerve Surgery. Proceedings of the Third International Symposium on Facial Nerve Surgery; Zurich; 1976*. 543-547.
- [25] House J. W. & Brackmann D. E. (1985). Facial nerve grading system. *Otolaryngol Head Neck Surg*. 93, 146-147.
- [26] Kang T. S., Vrabec, J. T., Giddings, N., Terris, D. J. (2002). Facial nerve grading systems (1985-2002): beyond the House-Brackmann scale. *Otol Neurotol*. 23, 767-771.
- [27] Burres S. & Fisch U. (1986). The comparison of facial grading systems. *Arch Otolaryngol Head Neck Surg*. 112, 755-758.
- [28] Murty G. E., Diver J. P., Kelly P. J., O'Donoghue G. M. & Bradley P. J. (1994). The Nottingham system: objective assessment of facial nerve function in the clinic. *Otolaryngol Head Neck Surg*. 110, 156-161.
- [29] Ross B. G., Frader G., Nedzelski J. M. (1996). Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg*. 114, 380-386.
- [30] Kinect for Windows Sensor (2015). Recuperado el 17 de marzo del 2015, de <https://msdn.microsoft.com/es-mx/library/hh855355.aspx>.
- [31] Ruiz R. (2014). Reconocimiento de expresiones faciales mediante imágenes adquiridas con cámara Kinect. Universidad Pontificia Comillas. Escuela Técnica Superior de Ingeniería (ICAI). Madrid, España.

- [32] Viola P. & Jones M. J. (2003). Robust Real-Time Face Detection. *International Journal of Computer Vision*. 57(2), 137-154.
- [33] Smolyanskiy N., Huitema C., et al. (2014). Real-time 3D face tracking based on active appearance model constrained by depth data. *Image and Vision Computing*. 32, 860-869.
- [34] Zhou M., Liang L., et al. (2010). AAM based face tracking with temporal matching and face segmentation. *Computer Vision and Pattern Recognition*. 701-708.
- [35] Philippe M. (2012). Tratamiento de la Parálisis Facial en medicina, fisioterapia y acupuntura. Recuperado el 23 de junio del 2015, de <http://www.efisioterapia.net>.
- [36] Khan H., Yoon-ho H., et al. (2007). Agreement between two different methods for facial nerve conduction study in patients with acute Bell's palsy. Department of Neurology, Seoul National University Hospital.
- [37] Chu E. A., Farrag T. Y., et al. (2011). Threshold of Visual Perception of Facial Asymmetry in a Facial Paralysis Model. *Arch Facial Plast Surg*. 13(1), 14-19.

ANEXOS I: PUBLICACIONES

International Conference on Intelligent Computing
Online Submission & Review System

International

Home

Guidelines

Paper's Decision

Final Paper Submission

Update Paper

Paper Submission

Submitted materials

Personal Info

Log out

Dear Author:

The papers you have submitted are listed as follows:

Paper ID	Paper Title	Paper's Status	Submitted Date	Presentation	Publication Type	Update Paper	Comments of Reviewers
505		rejected for repeated submission or low quality	2015-3-30 22:50:07			Rejected	View Comments
507	Face Recognition using SURF	Accepted	2015-3-30 22:58:40	oral		Update	View Comments

Dear Author:

The papers you have submitted are listed as follows:

Paper ID	Paper Title	Paper's Status
505		rejected for repeated submission or low quality
507	Face Recognition using SURF	Accepted

Face Recognition Using SURF

Raúl Cid Carro¹, Juan-Manuel Ahuactzin Larios²,
Edmundo Bonilla Huerta¹(✉), Roberto Morales Caporal¹,
and Federico Ramírez Cruz¹

¹ Instituto Tecnológico de Apizaco, 90300 Apizaco, Tlaxcala, Mexico
edbonn@hotmail.com, edbonn@hotmail.fr

² Probayes Americas SA de CV, Vía Atlixcáyotl 2301,
Parque Tecnológico de Puebla, 72453 Puebla, Puebla, Mexico

Abstract. In recent years, several scale-invariant features have been proposed in literature, this paper analyzes the usage of Speeded Up Robust Features (SURF) as local descriptors, and as we will see, they are not only scale-invariant features, but they also offer the advantage of being computed very efficiently. Furthermore, a fundamental matrix estimation method based on the RANSAC is applied. The proposed approach allows to match faces under partial occlusions, and even if they are not perfectly aligned. Thus based on the above advantages of SURF, we propose to exploit SURF features in face recognition since current approaches are too sensitive to registration errors and usually rely on a very good initial alignment and illumination of the faces to be recognized.

1 Introduction

Face recognition has been an active area of research over the last two decades due to wide applications. It involves computer recognition of personal identity based on geometric or statistical features derived from face images. Nowadays, illumination invariance, facial expressions, and partial occlusions are one of the most challenging problems in face recognition [1], where face images are usually analyzed locally to cope with the corresponding transformations. Local feature descriptors describe a pixel in an image through its local neighborhood content, their purpose is to provide a representation that allows to efficiently match local structures between images. They should be distinctive and at the same time robust to changes in viewing conditions. Many different descriptors and interest-point detectors have been proposed in the literature, and the descriptor performance often depends on the interest point detector [2].

The Scale Invariant Feature Transform (SIFT) proposed by David G. Lowe [3] has been widely used in object detection and recognition. Nonetheless, despite the variety of works on the use of SIFT features in face recognition, these methods still cannot satisfy the speed requirement of on-line applications. On the other hand, SURF suggested by Herbert Bay [4] is a scale and in-plane rotation invariant detector and descriptor with comparable or even better performance than SIFT. Its feature is also personal specific. Just like SIFT, in SURF, detectors are first employed to find the interest points in an image, and then the descriptors are used to extract the feature vectors at each interest point. However, instead of Difference of Gaussians (DoG) filter

used in SIFT, SURF uses Hessian-matrix approximation operating on the integral image to locate the interest points, which reduces the computation time drastically.

The benefits of facial recognition are that it is not intrusive, can be done from a distance even without the user being aware they are being scanned. What sets apart facial recognition from other biometric techniques is that it can be used for surveillance purposes; as in searching for wanted criminals, suspected terrorists, and missing children. Facial recognition can be done from far away so with no contact with the subject so they are unaware they are being scanned.

2 Speed-Up Robust Features (SURF)

SURF is a scale and in-plane rotation invariant feature. It contains interest point detector and descriptor. The detector locates the interest points in the image, and the descriptor describes the features of the interest points and constructs the feature vectors of the interest points. Conceptually similar to the SIFT descriptor, the 64-dimensional SURF descriptor also focusses on the spatial distribution of gradient information within the interest point neighborhood. The SURF descriptor is invariant to rotation, scale, brightness and, after reduction to unit length, contrast. It outperforms to existing schemes in terms of repeatability, distinctiveness and robustness, with a faster performance.

2.1 Interest Point Detection

In computer vision, the concept of interest points, also called keypoints or feature points, has been largely used to solve many problems in object recognition, image registration, visual tracking, 3D reconstruction, and more. It relies on the idea that instead of looking at the image as a whole, it could be advantageous to select some special points in the image and perform a local analysis on these ones. The main difference between SURF and SIFT descriptors is mainly speed and accuracy. Since SURF descriptors are mostly based on intensity differences, they are faster to compute. However, SIFT descriptors are generally considered to be more accurate in finding the right matching feature [3]. In order to detect characteristic points on a scale invariantly SIFT approach it uses cascaded filters, where the difference Gaussian, DoG, is calculated on rescaled images progressively.

2.2 Integral Image

The detection is speeded up by the summed area table known as an integral image, which is an algorithm for generating the sum of values in a rectangular sum of grid. The integral image can be given as:

$$I(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y)$$

2.3 Points of Interest in the Hessian Matrix

Different from SIFT using DoG to detect interest points, The SURF detector focuses its attention on blob-like structures in the image. These structures can be found at corners of objects, but also at locations where the reflection of light on specular surfaces is maximal. Integral images are used in Hessian matrix approximation, which reduce computation time drastically. Given a point $X = (x, y)$ in an image I , the Hessian matrix $H(x, \sigma)$ in X at scale σ is defined as follows:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix}$$

where $L_{xx}(X, \sigma)$, $L_{xy}(X, \sigma)$ and $L_{yy}(X, \sigma)$ are the convolutions of the second order Gaussian derivatives $\partial^2 g(\sigma)/\partial x^2$ with the image at point $X = (x, y)$. These derivatives are known as the Laplacian of Gaussians. The local maxima of this filter response occurs in regions where both L_{xx} and L_{yy} are strongly positive, and where L_{xy} is strongly negative. Therefore, these extrema occur in regions in the image with large intensity gradient variations in multiple directions, as well as at saddle points. Visually, this means that blob-like structures refer to corners and speckles. The calculation of all of these derivatives at different scales is computationally costly. The objective of the SURF algorithm is to make this process as efficient as possible.

Bay and many others have found that the number of results in scale-space decreases exponentially with increasing scale size. One potential explanation is that large amounts of Gaussian blurring can average out nearly all useful information in images. Therefore, searching through the scale space linearly can be very wasteful computationally. As an alternative, SURF introduces the notion of scale octaves: each octave linearly samples a small region of the scale space, but this region size is proportional to the scale magnitudes used. Each adjacent octave doubles the region size and sampling increment used in the previous one, to reduce the amount of search necessary at larger scales. The only negative side-effect of using octaves is that results found at larger scales can potentially have more error, due to the larger increments used in their corresponding octaves. To compensate, the SURF detector interpolates the coordinates of any local maxima found into the sub-pixel and sub-scale range.

Finally, Bay propose to make the distinction between bright blobs found on a dark background, versus dark blobs found on a bright background. This property can be represented by the sign of the Laplacian, as shown below:

$$\text{sgn}\{L_{xx}(X, \sigma) + L_{yy}(X, \sigma)\} = \begin{cases} +1 & \Rightarrow \text{bright blob over dark background} \\ -1 & \Rightarrow \text{dark blob over bright background} \end{cases}$$

To find blobs, we need to convolve the source image with various Gaussian-related filters. Because continuous Gaussian filters have non-integer weights, these convolutions require floating-point operations, which can severely hamper processing speed. Thus, SURF's authors propose to approximate these filters by using a set of 9×9 box filters (Fig. 1) this is used as the approximations of a Gaussian with $\sigma = 1.2$ and

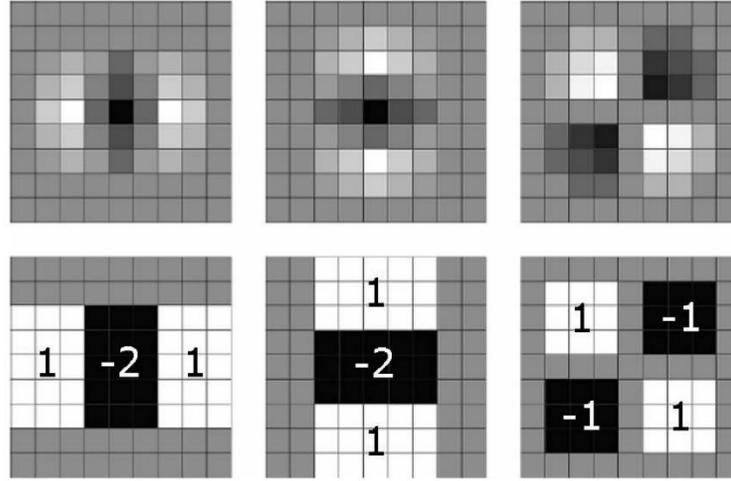


Fig. 1. Top row (left to right): The Gaussian second order partial derivative in L_{xx} , L_{yy} , and L_{xy} . Bottom row (left to right): The approximation for the second order Gaussian partial derivative D_{xx} , D_{yy} and D_{xy}

represents the lowest scale for computing the blob response maps. We will denote them by D_{xx} , D_{yy} and D_{xy} . The weights applied to the rectangular are kept simple for computational efficiency. This yields:

$$\det(H_{approx}) = D_{yy}D_{xx} - (wD_{xy})^2$$

where w is a weight for the energy conservation between the Gaussian kernels and the approximated Gaussian kernels, and

$$w = \frac{|L_{xy}(1.2)|_F |D_{yy}(9)|_F}{|L_{yy}(1.2)|_F |D_{xy}(9)|_F} = 0.912 \approx 0.9$$

$|X|_F$ is the Frobenius norm.

In order to detect keypoints using the determinant of Hessian it is necessary to introduce the notion of a scale space. The SURF constructs a pyramid scale space, like the SIFT. Different from the SIFT to repeatedly smooth the image with a Gaussian and then subsample the image, the SURF directly changes the scale of box filters to implement the scale space due to the use of the box filter and integral image [7].

2.4 Features Descriptors

In general, feature descriptors describe a pixel (or a position) in an image through its local content. They are supposed to be robust to small deformations or localization

errors, and give us the possibility to find the corresponding pixel locations in images which capture the same amount of information about the spatial intensity patterns under different conditions.

In feature matching, feature descriptors are usually N-dimensional vectors that describe a feature point, ideally in a way that is invariant to change in lighting and to small perspective deformations. In addition, good descriptors can be compared using a simple distance metric (for example, Euclidean distance). Therefore, they constitute a powerful tool to use in feature matching algorithms [5].

To describe each feature, SURF summarizes the pixel information within a local neighborhood. The first step is determining an orientation for each feature, by convolving pixels in its neighborhood with the horizontal and the vertical Haar wavelet filters. Shown in Fig. 2, these filters can be thought of as block-based methods to compute directional derivatives of the image's intensity. By using intensity changes to characterize orientation, this descriptor is able to describe features in the same manner regardless of the specific orientation of objects or of the camera. This rotational invariance property allows SURF features to accurately identify objects within images taken from different perspectives.

In fact, the intensity gradient information can also reliably characterize these pixel regions. By looking at the normalized gradient responses, features in images taken in a dark room versus a light room, and those taken using different exposure settings will all have identical descriptor values. Therefore, by using Haar wavelet responses to generate a unit vector representing each feature and its neighborhood, the SURF feature framework inherits two desirable properties: lighting invariance and contrast invariance.

For the extraction of the descriptor, the first step consists of constructing a square region centered at the interest point and oriented along the orientation decided by the orientation selection method introduced by Bay. The region is split up equally into smaller 4×4 square sub-regions (as shown in Fig. 3). This preserves important spatial information.

For each sub-region, we compute Haar wavelet responses at 5×5 equally spaced sample points. For simplicity, we call d_x the Haar wavelet response in horizontal direction and d_y the Haar wavelet response in vertical direction. To increase the robustness towards geometric deformations and localization errors, the responses d_x and d_y are first weighted with a Gaussian centered at the interest point. These values are computed (with a kernel size of 4σ) within a circular neighborhood of radius 6σ at locations regularly spaced by intervals of σ . For a given orientation, the responses inside a certain angular interval $\pi/3$ are summed, and the orientation giving the longest vector is defined as the dominant orientation.

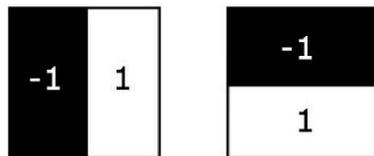


Fig. 2. Horizontal and vertical Haar wavelet filters

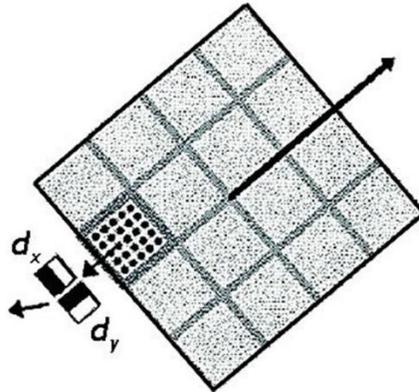


Fig. 3. Structure breakdown of each feature's neighborhood

Then, the wavelet responses d_x and d_y are summed up over each sub-region and form a first set of entries in the feature vector. In order to bring in information about the polarity of the intensity changes, we also extract the sum of the absolute values of the responses, $|d_x|$ and $|d_y|$. Hence, each sub-region has a four-dimensional descriptor vector v for its underlying intensity structure $v = \sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|$. Concatenating this for all 4×4 sub-regions, this results a descriptor vector of length 64. The wavelet responses are invariant to a bias in illumination (offset). Invariance to contrast (a scale factor) is achieved by turning the descriptor into a unit vector.

2.5 Recognition by Matching

With the SURF features and descriptors, scale-invariant matching can be achieved. For example, (Fig. 4) shows two faces of the same person in different moments where the lines indicate the corresponding matched interest points. This is accomplished by first detecting features on each image, and then extracting the descriptors of these features. Each feature descriptor vector in the first image is then compared to all feature descriptors in the second image. The pair that obtains the best score (that is, the lowest distance between the two vectors) is then kept as the best match for that feature. This process is repeated for all features in the first image. Good feature descriptors must be invariant to small changes in illumination, in viewpoint, and to the presence of image noise. Therefore, they are often based on local intensity differences.

The matching is carried out by a nearest neighbor matching strategy $m(X, Y)$: the descriptor vectors $X = \{x_1, \dots, x_I\}$ extracted at keypoints $1, \dots, I$ in a test image X are compared to all descriptor vectors $Y = \{y_1, \dots, y_J\}$ extracted at keypoints $1, \dots, J$ from the reference images $Y_n, n = \{1, \dots, N\}$ by the Euclidean distance. Additionally, a ratio constraint is applied only if the distance from the nearest neighbor descriptor is less than α times the distance from the second nearest neighbor descriptor, a matching pair is detected. Finally, the classification is carried out by assigning the class

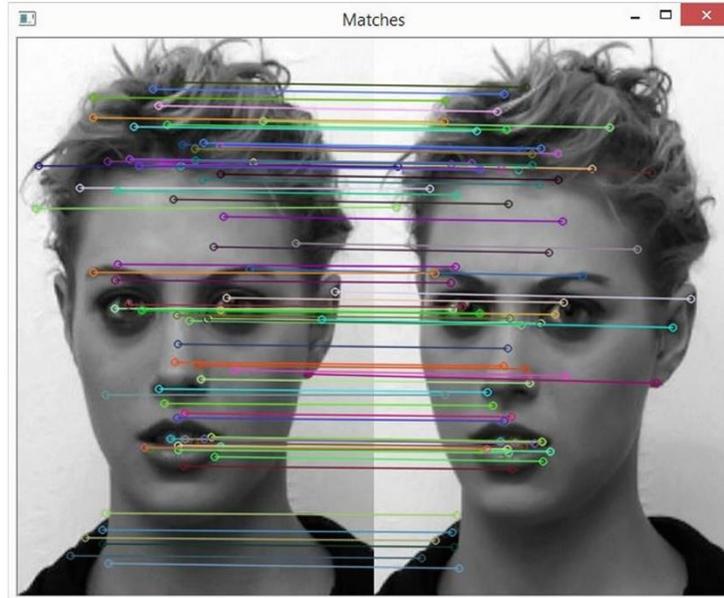


Fig. 4. Most of the matches correctly link a point on the left with its corresponding image point on the right, despite some differences in head position and facial expression between the images

$c = \{1, \dots, C\}$ of the nearest neighbor image $Y_{n,c}$ which achieves the highest number of matching correspondences to the test image X .

2.6 Matching Images Using RANSAC

The use of the spatial information about matching points can help to reduce the amount of falsely matched correspondences, i.e. outliers. With the assumption that many parts of a face nearly lie on a plane, with only small viewpoint changes for frontal faces, a given homography (transformation) between the test and train images can reject outlier matches which lie outside a specified inlier radius.

The RANSAC (Random Sample Consensus) algorithm aims at estimating a given mathematical entity from a data set that may contain a number of outliers. The idea is to randomly select some data points from the set and perform the estimation only with these. The number of selected points should be the minimum number of points required to estimate the mathematical entity. Once a transformation has been estimated, all points of a test image will be projected to the train image. If a projected point lies in a given radius to its corresponding point it is classified as inlier for that particular homography, otherwise it is declared as an outlier. After a given number of iterations the maximum amount of inliers of all estimated homographies will be used as a measurement to determine the likelihood of the similarity between the test and the train image.

The central idea behind the RANSAC algorithm is that the larger the support set is, the higher the probability that the computed matrix is the right one. Obviously, if one (or more) of the randomly selected matches is a wrong match, then the computed fundamental matrix will also be wrong, and its support set is expected to be small. This process is repeated a number of times, and at the end, the matrix with the largest support will be retained as the most probable one [5].

2.7 Face Recognition Algorithm

The first step is detect the feature point and computing their descriptors. Next, we proceed to feature matching in order to find the best matching points for each feature. We now have two relatively good match sets, one from the first image to second image and the other one from second image to the first one. From these sets, we will now reject matches that do not obey the homography constraint. This test is based on the RANSAC (Fig. 5).

3 Experimental Results

We have used the openCV framework, this is a C++ framework for computer vision. It ships with its own implementation of SURF and several other computer vision algorithms. It was chosen as it provides different good low level routines for working with images and easy loading and saving of different image formats.

The algorithm was tested using two free face database images that contains about ten different images of each of 79 distinct subjects [8, 9]. For some subjects, the images

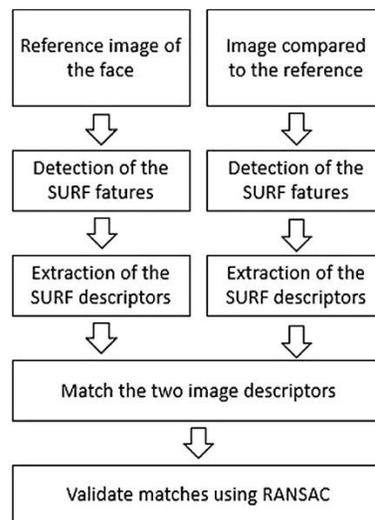


Fig. 5. Overview of the proposed approach

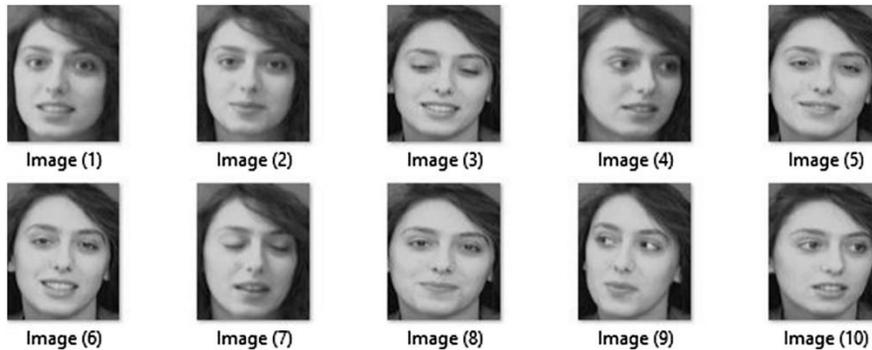


Fig. 6. Example of the images under different conditions

were taken at different times, varying the lighting, facial expressions (open, closed eyes, smiling, not smiling) and facial details (glasses, no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position and its total number of keypoints was calculated (Fig. 6).

Metrics like Correct Recognition Rate (*CRR*) for identification and the Equal Error Rate (*EER*) for recognition are used to measure the performance of SURF against SIFT. *CRR* of the system is defined as

$$CRR = \frac{NC}{TN} \times 100$$

where *NC* denotes the number of correct (Non-False) recognitions of face images and *TN* is the total number of face images in the testing set. At a given threshold, the probability of accepting the imposter, known as False Acceptance Rate (*FAR*) and the probability of rejecting the genuine person, known as False Rejection Rate (*FRR*) are obtained. Equal Error Rate (*EER*) is the error where $FAR = FRR$ (Table 1).

As it can see, the algorithm SURF evinces a better performance and accuracy as expected. On the other hand SIFT turned out to be more time consuming. The range for matching points between each of the images of the same person is 7 to 45. Therefore, by the number of matching points we can determine a certain degree of similarity between two images. The Fig. 7 shows the comparison and matching between two images having a minimal difference. As can be seen, all points are matched correctly, therefore we can conclude with a high degree of certainty that both images are from the same person or someone very similar.

The Fig. 8 shows an example of the effectiveness in both algorithms at the time to matching the keypoints correctly that allow us to identify the resemblance or similarity

Table 1. Comparative results between SIFT and SURF

	CRR	ERR
SIFT	0.8734	0.3170
SURF	0.9897	0.2962

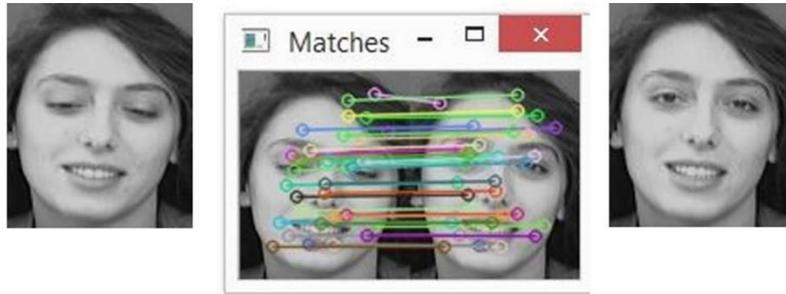


Fig. 7. Image(3) and Image(6) have the highest number of matched points. The difference between the two is minimal

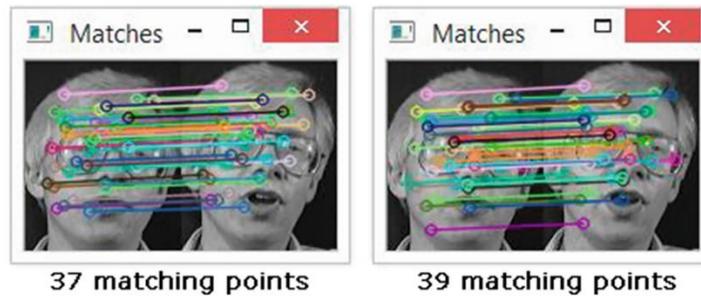


Fig. 8. Comparison between SIFT (left image) and SURF (right image)

present in images despite the different circumstances. Nonetheless, SURF obtained a greater number of correct matching points than SIFT in a shorter period of time.

4 Conclusions

Both algorithms can be used for face recognition, but nevertheless SURF algorithm is proven to be more suitable. However, we believe that facial recognition is more beneficial to use for facial authentication than for identification purposes, as it is too easy for someone to alter their face, features with a disguise or mask, etc. Environment is also a consideration as well as subject motion and focus on the camera.

References

1. Ekenel, H.K., Stiefelwagen, R.: Why is facial occlusion a challenging problem? In: International Conference on Biometrics, Sassari, Italy (2009)
2. Mikolajczyk, K., Schmid.: A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1615–1630 (2005)

3. Lowe, D.G.: Object recognition from local scale-invariant features. In: International Conference on Computer Vision, Corfu, Greece pp. 1150–1157 (1999)
4. Bay, H., Ess, A., Tuytelaars, T., Van-Gool, L.: Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **110**, 346–359 (2008)
5. Lagani-Re, R.: *OpenCV 2 Computer Vision Application Programming Cookbook*. Packt Publishing, Birmingham, Birmingham UK (2011)
6. Li, J., Zhang, Y.: Learning SURF Cascade for Fast and Accurate Object Detection. Intel Labs China (2013)
7. Hung-Fu, H., Shen-Chuan, T.: Facial expression recognition using new feature extraction algorithm. *Electron. Lett. Comput. Vis. Image Anal.* **11**, 41–54 (2012)
8. Lee, K.C., Ho, J., Kriegman, D.: Acquiring linear subspaces for face recognition under variable lighting. *IEE Trans. Pattern Anal. Mach. Intell.* **27**, 684–698 (2005)
9. Samaria, F., Harter, A.: Parameterisation of a stochastic model for human face identification. In: *Proceeding of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota FL (1994)

Paper Submission #3467 for RevistaIEEE-AL

De : editor-in-chief-r9@ieee.org
Envoyé : mer. 07/10/15 12:06
À : editor-in-chief-r9@ieee.org, edbonn@hotmail.com

THIS IS AN AUTOMATIC MESSAGE. PLEASE, DON'T REPLY.
If you notice any problems, please contact the program chair at editor-in-chief-r9@ieee.org.

Thank you for your submission to RevistaIEEE-AL. Below is a copy of the information you submitted for your records.

Paper ID: 3467

Title: Facial Expression Analysis with Kinect for the Diagnosis of Paralysis Using Nottingham Grading System

Title: Análisis de expresión facial con Kinect para el diagnóstico de parálisis utilizando el sistema de gradiente de Nottingham

Author1

Name: Raúl Cid Carro
Org: Instituto Tecnológico de Apizaco
Country: Mexico
Email:raul.cid.carro@gmail.com

Author2

Name: Edmundo Bonilla Huerta
Org: Instituto Tecnológico de Apizaco
Country: Mexico
Email:edbonn@hotmail.com

Author3

Name: Federico Ramírez Cruz
Org: Instituto Tecnológico de Apizaco
Country: Mexico
Email:

Author4

Name: Roberto Morales Caporal
Org: Instituto Tecnológico de Apizaco
Country: Mexico
Email:

Author5

Name: Carlos Pérez Corona
Org: Instituto Tecnológico de Apizaco
Country: Mexico
Email:

Other Authors:

Contact Author #: 2

Contact Alt Email: edbonn@hotmail.fr

Contact Phone: 0124172010

Keywords: Kinect, Facial Nerve Paralysis, Facial Nerve Palsy, Nottingham Grading System, Image Analysis.

Abstract: Facial palsy is a neuromuscular disorder caused by damage to motor and sensory pathways of the facial nerve, which causes a movement disorder in the facial muscles. In clinical diagnosis of facial nerve palsy, simple and quantitative assessment of facial motion is strongly required. A systematic evaluation of facial nerve paralysis allows the clinician to determine objectively the severity of disability and record this information, in order to evaluate the response to therapy. The proposed method evaluates the degree of facial paralysis by using Kinect as a feature extraction tool and Nottingham grading system. The proposed grading system measure the distance between the eyebrows to infra orbital and the distance between the edges of mouth and lateral canthus. We divided the grading scale into three levels: normal, mild and severe by using the estimated parameters.

Comments:

Paper: included



Registro ponencia -- Al menos uno de los ponentes debe prestar sus servicios en el extranjero o en los estados de Puebla y Tlaxcala

Gracias

Hemos registrado su propuesta y le notificaremos en unos cuantos días de la decisión del Comité de Arbitraje del Congreso. La lista de ponencias aceptadas se encuentra en el sitio del congreso: AcademiaJournals.com

[Modificar tu respuesta](#)

[Enviar otra respuesta](#)

Este formulario se creó con Formularios de Google.
[Crea uno propio](#)

 Google Forms

Facial Expression Analysis with Kinect for the Diagnosis of Paralysis Using Nottingham Grading System

R. Cid, E. Bonilla, C. Pérez, J. F. Ramírez and R. Morales

Abstract— Facial palsy is a neuromuscular disorder caused by damage to motor and sensory pathways of the facial nerve, which causes a movement disorder in the facial muscles. In clinical diagnosis of facial nerve palsy, simple and quantitative assessment of facial motion is strongly required. A systematic evaluation of facial nerve paralysis allows the clinician to determine objectively the severity of disability and record this information, in order to evaluate the response to therapy. The proposed method evaluates the degree of facial paralysis by using Kinect as a feature extraction tool and Nottingham grading system. The proposed grading system measure the distance between the eyebrows to infra orbital and the distance between the edges of mouth and lateral canthus. We divided the grading scale into three levels: normal, mild and severe by using the estimated parameters.

Keywords— Kinect, Facial Nerve Paralysis, Facial Nerve Palsy, Nottingham Grading System, Image Analysis.

I. INTRODUCCIÓN

ACTUALMENTE, las ciencias de la salud, y la medicina en particular, son uno de los campos del saber más evolucionados y beneficiados por el uso de tecnología, al tiempo que registran un crecimiento exponencial tanto en el número de usuarios, como en el de instituciones que se han incorporado a la búsqueda de diferentes medios que permitan tener un mejor nivel de vida. Según David Banta, una tecnología (médica) es cualquier técnica o herramienta, producto o proceso, método o aparato, que permita ampliar las capacidades humanas [1]. Una forma común para clasificar a las tecnologías médicas es la siguiente:

- Tecnologías preventivas: Su propósito es proteger al individuo, impidiendo la presentación y/o el avance de la enfermedad.
- Tecnologías de diagnóstico: Sirven para determinar los momentos y/o procesos patológicos por los que pasa un paciente.
- Tecnologías terapéuticas o de rehabilitación: Su función es eliminar a la enfermedad o minimizar sus efectos (aquí cabe una subdivisión, para especificar que existen tecnologías con un sentido realmente curativo y otras de carácter paliativo, llamadas por Lewis Thomas *half-way technologies* [2]).

- Tecnologías de administración y organización: Son estrategias para el otorgamiento correcto y oportuno de los servicios de salud [3].

En México la necesidad por los diferentes tipos de tecnología médica va en aumento. Según el censo realizado por el Instituto Nacional de Estadística y Geografía (INEGI) en el año 2010, existen en México 5 millones 739 mil 270 personas con discapacidad, lo que representa el 5.1% de la población total del país, en donde la parálisis facial ocupa uno de los diez primeros lugares de atención en diversas unidades de rehabilitación del país; esto debido en gran parte a que el nervio facial es el que más se lesiona de todos los pares craneales.

II. PARÁLISIS FACIAL

La parálisis facial es la debilidad o parálisis de la musculatura inervada por el nervio facial (VII par craneal) debida a una lesión del VII par craneal en cualquier lugar de su recorrido desde el núcleo de origen a las estructuras que inerva. La lesión impide el movimiento normal de diferentes músculos como: frente, párpados y boca; así como, la expresión de las emociones. La parálisis de los músculos faciales causa asimetría, ocasionando problemas psicológicos y emocionales.

Existen dos tipos de parálisis facial: *Parálisis Facial Periférica* (PFP) y *Parálisis Facial Central* (PFC) (ver Figura 1). La PFC es causada por afección de la neurona motora superior, y se acompaña con frecuencia de hemiplejía del mismo lado de la lesión. Afectando al movimiento voluntario de la parte interior de la parte inferior de la cara, pero no a los músculos frontal y orbicular de los párpados. Además, suelen ser normales los movimientos faciales reactivos a emociones. La PFP, también es conocida como parálisis de Bell, de neurona motora inferior o parálisis idiopática; es causada por la afección aguda del nervio facial a nivel periférico; origina pérdida del movimiento voluntario en todos los músculos faciales del lado afectado y es la más común de las parálisis. La PFP puede ser unilateral o bilateral, completa si abarca todos los músculos del lado afectado, o parcial si sólo afecta a un grupo de ellos. La forma de parálisis más común es unilateral, afectando la hemicara del nervio lesionado [4].

R. Cid, Instituto Tecnológico de Apizaco (ITA), Apizaco, Tlaxcala, México, raul.cid.carro@gmail.com

E. Bonilla, Instituto Tecnológico de Apizaco (ITA), Apizaco, Tlaxcala, México, edbonn@hotmail.com

C. Pérez, Instituto Tecnológico de Apizaco (ITA), Apizaco, Tlaxcala, México, cperezcorona@gmail.com

J. F. Ramírez, Instituto Tecnológico de Apizaco (ITA), Apizaco, Tlaxcala, México, federico_ramirez@yahoo.com.mx

R. Morales, Instituto Tecnológico de Apizaco (ITA), Apizaco, Tlaxcala, México,

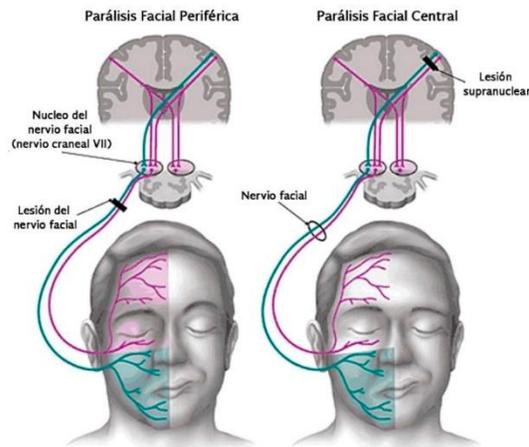


Figura 1. Clasificación de la parálisis facial.

III. DIAGNÓSTICO

La orientación diagnóstica incluye una cuidadosa evaluación clínica y exploración física. Partiendo de estas evaluaciones se realizan una serie de estudios que tienen como objetivo identificar la causa de la parálisis, los cuales son: análisis de sangre del fluido cerebro-espinal, rayos X, estudios de neuroimagen de conducción del nervio, electromiografía y electroneurografía. Estos estudios se solicitan de acuerdo a sospecha clínica posterior a la evaluación clínica de signos y síntomas empleando algún método de diagnóstico.

MÉTODOS DE DIAGNÓSTICO

Los métodos usados actualmente para clasificar el nivel de la función del nervio facial se dividen en dos enfoques: *tradicionales* o *informáticos*. El término “enfoque tradicional” es usado para referirse a los métodos que no requieren ningún equipo de cómputo especializado. Una gran parte de estos enfoques se basan en la utilización de las evaluaciones subjetivas del observador. Por otra parte en los métodos con un “enfoque informático” se implica el uso de equipo especializado para medir y cuantificar los datos digitales de manera objetiva. El mayor inconveniente con este enfoque es el costo y la operatividad.

Los enfoques tradicionales para la evaluación de la función facial han evolucionado considerablemente durante el último medio siglo. Los informes iniciales en la literatura eran de carácter descriptivo. La escala de House-Brackmann, aunque resulta útil en la normalización de los informes de la parálisis del nervio facial, es imperfecta. En primer lugar, debido a la amplia gama de movimientos faciales comprendida en diferentes grados, no puede distinguir los cambios sutiles en la función facial. Como resultado, la diferenciación entre los mejores y peores resultados de procedimientos de injerto de nervio puede ser imperfecta para discernir. En segundo lugar,

los pacientes pueden tener disfunción en una región facial que encaja en un grado, sin embargo, ajustarse en un grado diferente en otra región. Este aspecto puede dar lugar a inconsistencia en la puntuación [5]. Informes anteriores han señalado que la escala House-Brackmann tiene limitaciones significativas de concordancia entre observadores. Estas preocupaciones han llevado a la propuesta de sistemas adicionales.

El primero de ellos fue el sistema Burres-Fisch. El sistema Burres-Fisch es un sistema objetivo sobre la base de mediciones lineales del desplazamiento de puntos en la cara de referencia. Las distancias se midieron con calibradores de mano, así como con electrodos de superficie que producían registros electromiográficos. Este enfoque proporciona un índice de medición lineal que es una escala graduada continua. El sistema Burres-Fisch fue diseñado para minimizar el prejuicio del observador y la subjetividad. Entre las limitaciones de este enfoque están la incapacidad para hacer simultáneamente registros de diferentes regiones faciales, la falta de evaluación de defectos secundarios y la más importante, la excepcionalmente ardua y lenta tarea de medir y realizar los cálculos requeridos [6].

Dos enfoques adicionales son el sistema de clasificación Nottingham y la escala Sunnybrook. El sistema Nottingham mantuvo el concepto de Burres-Fisch de medición objetiva pero permitió una evaluación más rápida y agregó criterios para defectos secundarios. Las diferencias en las mediciones de los puntos de referencia mostrados en la Figura 2 en un estado de reposo y tras el máximo esfuerzo son sumadas y utilizadas para generar un porcentaje cuando se compara contra el lado contralateral. La presencia de defectos secundarios está marcada por separado. Esta escala ha sido identificada por varios investigadores como posiblemente ser la más fácil y reproducible de todas, por esta razón fue elegida como el sistema de clasificación del nivel de parálisis facial para este trabajo. Las limitaciones de este método incluyen la incapacidad para evaluar la disfunción del nervio facial bilateral y el uso de sólo modificadores descriptivos para describir defectos secundarios [7]

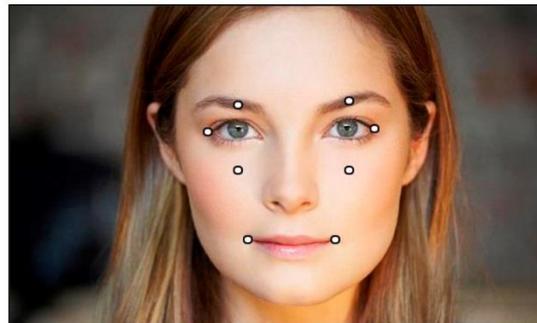


Figura 2. Puntos de referencia facial designados en el método Nottingham.

IV. KINECT

El *Kinect* es un dispositivo físico creado por Alex Kipman y desarrollado por *Microsoft* para la consola *Xbox 360*, fue lanzado en noviembre del 2010 y desde febrero del 2012 para PC a través de *Windows 7* y *Windows 8*. *Kinect*. Permite a los usuarios controlar e interactuar con la consola sin necesidad de tener contacto físico con un control de videojuegos tradicional, mediante una interfaz natural de usuario que reconoce gestos, comandos de voz, objetos e imágenes.

Los componentes principales del *Kinect*® son indicados en la Figura 3 [8], ellos son:

1. Micrófono multidireccional: Es un arreglo de cuatro micrófonos que pueden aislar la voz del usuario del ruido en la habitación. Mediante la comparación de la demora en cada micrófono, la fuente de la voz puede ser localizada.
2. Emisor de láser infrarrojo: Emite activamente un espectro infrarrojo cercano, el cual puede ser distorsionado por una superficie irregular y luego formado aleatoriamente como manchas reflejadas. Las manchas pueden ser recibidas por el sensor de profundidad (Número 3 en la figura).
3. Sensor de profundidad (cámara infrarroja): Captura la señal de infrarrojos que se puede convertir en un mapa de profundidad.
4. Inclinador motorizado: El motor puede ser programado con el fin de lograr el mejor ángulo de visión.
5. Cable USB: Transmite el flujo de video, de profundidad y de audio. La fuente de alimentación adicional debe conectarse para obtener todas las funciones del *Kinect*. (La fuente de poder del *Kinect* es de 12W mientras que la fuente de poder de un cable USB normal es de 2.5W).
6. Cámara RGB: Captura el flujo de video de color.

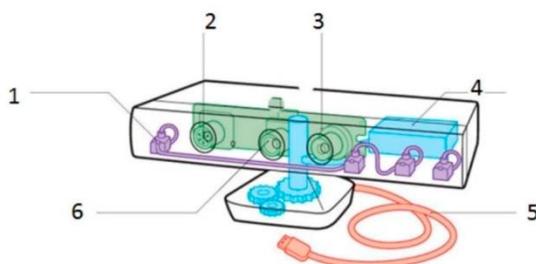


Figura 3. Estructura del Kinect y sus componentes principales.

ESPECIFICACIONES TÉCNICAS

Sensor de profundidad

El sensor de profundidad permite tener una visión 3D del entorno de *Kinect*. El emisor genera rayos infrarrojos y el sensor

detecta los rayos que rebotan en los objetos o personas de la habitación. Los rayos reflejados son convertidos en información de profundidad para poder medir la distancia entre un objeto y el sensor, lo que permite generar la imagen de profundidad. Esto es especialmente relevante para el trabajo, al hacer posible conocer con precisión la distancia de los diferentes puntos del rostro a la cámara. El sensor de profundidad cuenta con un rango válido que va desde 0.8 hasta 4 metros de distancia, tal como se muestra en la Figura 4.

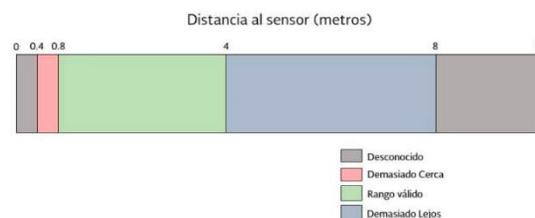


Figura 4. Rango válido del sensor de profundidad.

En una imagen de profundidad, el color de cada píxel nos dirá qué tan lejos está esa parte de la imagen de la cámara. Puesto que estos valores corresponden directamente al lugar donde se encuentran los objetos en el espacio, estos son mucho más útiles en la determinación para saber en dónde empieza un objeto, en donde acaba otro y si hay alguna persona alrededor.

Cámara RGB

Kinect cuenta con un sensor óptico que permite obtener imágenes a color en tiempo real. Dicho sensor cuenta con un campo de visión de 43° en vertical por 57° en horizontal y es capaz de capturar fotogramas a razón de 30 imágenes por segundo.

INTERFAZ DE SEGUIMIENTO FACIAL DE KINECT

El Kit de Desarrollo de Software de Seguimiento Facial de Windows (SDK de Seguimiento Facial), junto con el Kit de Desarrollo de Software de *Kinect* para Windows (SDK de *Kinect para Windows*), contienen una serie de librerías e interfaces que permiten seguir rostros humanos en tiempo real. El motor del seguimiento facial del SDK de Seguimiento Facial analiza los datos de entrada capturados por la cámara *Kinect*, calcula la posición de la cara y expresiones faciales, y hace esa información disponible para cualquier aplicación en tiempo real.

El SDK de Seguimiento Facial usa el sistema de coordenadas del *Kinect* para obtener sus resultados de seguimiento en 3D. El origen de coordenadas se localiza en el centro óptico de la cámara (ver Figura 5 [8]); el eje Z apunta hacia el usuario, el eje Y apunta hacia arriba y el eje X hacia la izquierda del sensor. Las medidas se obtienen en metros para las distancias y en grados para los ángulos de rotación.

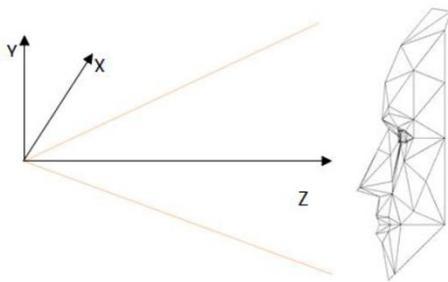


Figura 5. Espacio de la cámara.

El SDK de Seguimiento Facial acepta como entrada imágenes de color y profundidad del *Kinect*. La calidad de seguimiento puede verse afectada por la calidad de esas imágenes de entrada, ya que las imágenes oscuras o borrosas dificultan la tarea del seguimiento y reconocimiento facial. Por otro lado, las caras más grandes o cercanas son más fáciles de seguir.

V. DESCRIPCIÓN DEL SISTEMA

El sistema desarrollado consta de tres fases principales: adquisición de los datos, almacenamiento de los datos y evaluación de los datos.

En la primera fase, la adquisición de los datos se realizó obteniendo la información visual y facial proporcionada por el *Kinect* en tiempo real. Para lograr esta tarea, se crearon diferentes herramientas, las cuales resultan esenciales dentro del trabajo, ya que la exactitud y fiabilidad con que son recogidos los datos dependen todo su posterior análisis y los resultados obtenidos.

Para la fase de almacenamiento de los datos, se consideró guardar una imagen del paciente en formato JPEG al momento de realizar alguna de las expresiones faciales previamente seleccionadas. Al mismo tiempo, se creará un archivo de texto asociado a la imagen recién guardada que contiene las distancias entre los puntos de interés para el sistema de clasificación Nottingham.

La última fase toma como entrada los archivos de texto generados por la fase anterior y realiza sobre ellos los cálculos necesarios para obtener un resultado (diagnóstico). La Figura 6 muestra las tres fases anteriormente descritas que forman parte del sistema.



Figura 6. Estructura del sistema.

ADQUISICIÓN DE LOS DATOS

El cometido principal de esta fase es permitir la visualización en tiempo real de las imágenes RGB captadas por el *Kinect* e identificar la información de los puntos faciales de interés para el sistema de clasificación Nottingham. De esta manera el usuario puede determinar la adecuación de las condiciones del procedimiento y modificar cualquiera de ellas en caso de ser necesario. Además, también permite apreciar si el *Kinect*® está detectando un rostro mediante la superposición en la interfaz de los puntos faciales de interés detectados con el rostro de la persona. Esta visión aporta información adicional acerca de la validez del procedimiento, ya que la no detección del rostro o su detección defectuosa podrían deberse a unas condiciones no óptimas en cuanto a luminosidad o distancia de la persona a la cámara del *Kinect*.

La obtención de los datos constituye la fase inicial del ciclo de vida del sistema. Su funcionamiento se descompone en las siguientes etapas que serán descritas posteriormente:

1. Establecimiento de la comunicación con el sensor Kinect.
2. Detección y seguimiento del rostro de una persona.
3. Discriminación de la información facial y visual de la persona

ALMACENAMIENTO DE LOS DATOS

En esta fase del sistema se guarda tanto la información facial como visual del paciente al momento de realizar alguna de las cuatro expresiones faciales establecidas. Para esto se utilizarán dos tipos de archivos; el primero, es un archivo de texto que contiene la información facial necesaria para realizar el diagnóstico objetivo por medio del sistema de clasificación Nottingham. Dentro del archivo se guardan las distancias específicas entre pares de puntos de interés. La Figura 7 muestra la relación entre los puntos que forman las distancias que utiliza el sistema para evaluar el nivel de movilidad facial.

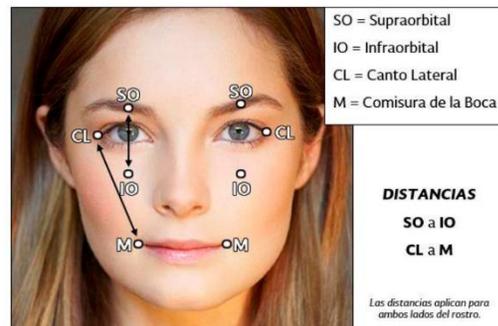


Figura 7. Distancias entre los puntos de interés.

El segundo archivo, es una fotografía del paciente, ésta sirve para complementar su expediente y llevar un registro visual del avance del paciente. La fotografía es guardada en un formato RGB con una resolución de 1280x920 píxeles. Ambos archivos

son creados simultáneamente en la misma carpeta y con el mismo nombre que es designado desde el sistema por defecto.

EVALUACIÓN DE LOS DATOS

En esta última fase se calcula un porcentaje que representa el nivel de simetría facial que tiene una persona, donde el 100% o valores muy cercanos a él representan un rostro sano (sin parálisis). Este valor se obtiene mediante la evaluación de los datos faciales que se encuentran en los archivos de textos generados en la fase anterior. Por tal motivo, la interfaz del sistema le permite al médico seleccionar los cuatro archivos de texto referentes a las cuatro expresiones faciales que se le pidió al paciente realizar. La Figura 8 [9] muestra ejemplos visuales de las expresiones faciales que el paciente debe realizar y resalta en rojo el área de los músculos mayormente involucrados en el movimiento que se esté realizando. El rostro en reposo es tomado como referencia en términos de movilidad; las distancias entre los puntos al realizar algún movimiento son comparadas con las distancia de esos puntos pero con el rostro en reposo.



Figura 8. Expresiones faciales que debe realizar el paciente en el sistema.

Una vez seleccionados los archivos, el sistema extrae los valores de las distancias de cada uno de ellos y sigue el procedimiento clásico del sistema de clasificación Nottingham. La Figura 9 [19] muestra la guía utilizada en las clínicas para diagnosticar la parálisis facial siguiendo el sistema de clasificación Nottingham. En la primera mitad de la guía, se muestra una imagen con la disposición de los ocho puntos de interés y las cuatro distancias importantes entre ellos, en la segunda mitad, se visualiza el procedimiento para calcular el porcentaje de simetría.

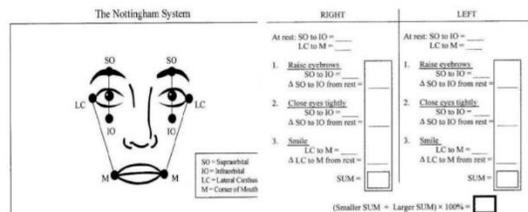


Figura 8. Guía para el diagnóstico de parálisis facial mediante el sistema de clasificación Nottingham.

DESCRIPCIÓN DE LA INTERFAZ DE USUARIO

La interfaz de usuario se divide en cinco partes. A continuación se describen el funcionamiento secuencial que tienen cada una de estas partes (ver Figura 9).

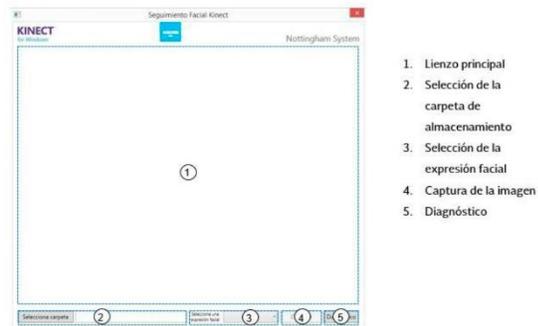


Figura 9. Partes que compone la interfaz.

1) Lienzo principal

Se presenta en la parte central de la interfaz y muestra la imagen en tiempo real que capta la cámara del *Kinect*, de no ser así es posible que exista algún problema de comunicación entre el *Kinect* y la computadora. Una vez que se detecte la imagen, el paciente debe ser colocado siguiendo las especificaciones técnicas del dispositivo, vistas en el capítulo anterior. Para esto, se recomienda crear un escenario fijo, donde no se modifiquen elemento del entorno como la luz o la ubicación de la cámara o del paciente, esto con el fin de disminuir el número de variables que puedan alterar los resultados. En este escenario, se debe dar preferencia a la ubicación del rostro, tratando de colocarlo en el centro del lienzo para que sea lo más visiblemente posible.

2) Selección de la carpeta de almacenamiento

Consta del botón "Seleccionar carpeta" que al hacer click, despliega un cuadro de dialogo donde el médico debe seleccionar la carpeta donde desea guardar los archivos con la información facial y visual del paciente. Como sugerencia se propone que el nombre de la carpeta sea único y representativo de cada paciente. Una vez seleccionada la carpeta, la ubicación de ésta será mostrada en el cuadro de texto que se encuentra a lado derecho del botón "Seleccionar carpeta".

3) Selección de la expresión facial

Para seleccionar la expresión facial que se va a capturar se cuenta con un cuadro de lista desplegable. Este cuadro contiene un listado con las cuatro expresiones faciales que se requiere que el paciente realice. Es importante prestar atención a esta selección para generar los archivos correctamente y no haya duplicidad.

4) Captura de imagen

El botón "Capturar" se encuentra por defecto deshabilitado, éste se habilita hasta que el médico haya seleccionado una carpeta de almacenamiento y el sistema haya reconocido un rostro. Una vez que se da click sobre el botón, se muestra un mensaje indicando que la captura fue realizada correctamente y

por consiguiente, la creación de los archivos con la información facial y visual del paciente fue exitosa.

5) Diagnóstico

Para utilizar el botón “Diagnóstico” de la última parte de la interfaz, es necesario contar antes con los archivos generados en la parte de *Captura de imagen*. Al dar click sobre el botón, se despliega un cuadro de dialogo, donde el médico debe seleccionar sólo los 4 archivos relacionados a las expresiones faciales que hizo un paciente en un día determinado. Si los archivos seleccionados son correctos, se mostrará un mensaje con el porcentaje de simetría facial que presenta la persona. El botón “Diagnóstico” permanece siempre habilitado, lo que permite al médico acceder a esta parte del sistema de manera automática, sin necesidad de capturar nuevamente las imágenes del rostro de un paciente y, de esta manera; realizar el cálculo del nivel de parálisis sobre los archivos ya existentes las veces que quiera.

VI. PRUEBAS

Se realizaron dos pruebas diferentes con propósitos específicos. La primera prueba se hizo para corroborar la consistencia del resultado en el proceso de diagnóstico, la segunda prueba consistió en la aplicación del procedimiento de diagnóstico usando el sistema en personas que tienen cierto nivel de parálisis facial. Ambas pruebas contribuyen en la validación de los resultados obtenidos por el sistema. El proceso para el desarrollo de cada una de ellas es descrito a continuación.

CONSISTENCIA EN EL DIAGNÓSTICO

Para comprobar la solidez de los resultados obtenidos durante el proceso de diagnóstico por parte del sistema, se ideó calcular el porcentaje de simetría facial de sólo una persona sana por un lapso de 10 días, con el fin de demostrar que todos los resultados obtenidos son iguales o muy similares y las diferencias que llegan a existir entre ellos es debido, en su mayoría, por variables como la iluminación o el posicionamiento del rostro. Gracias a esta prueba, se detectó que la presencia de barba en abundancia en el sujeto, pueden llegar a interferir en la ubicación correcta de algunos de los puntos de interés del sistema. Sin embargo, este factor no representa un obstáculo significativo en la aplicación del procedimiento, por lo que no es necesario afeitarse por completo el rostro, pero se obtienen resultados más precisos. La Figura 10 muestra una tabla con los resultados obtenidos de esta prueba.



Figura 10 Resultados de la prueba de consistencia.

Como se puede observar en la figura anterior, todos los valores son similares entre sí; su media aritmética es de 97.60 y tienen una desviación estándar de sólo 1.05. Con los resultados obtenidos se puede asegurar que la prueba cumplió su objetivo al demostrar que los valores presentan sólo una pequeña variabilidad esperada debido a variables externas descritas con anterioridad. Los porcentajes altos obtenidos confirman la simetría natural de un rostro sano y del mismo modo evidencian la pequeña asimetría natural del rostro explicada anteriormente, esto confirma una gran solidez por parte del sistema.

APLICACIÓN DEL SISTEMA EN PACIENTES CON PARÁLISIS FACIAL

Tal como se observó en la prueba anterior, el nivel de parálisis facial puede concebirse como el grado de simetría en los movimientos que se puede alcanzar en ambos lados del rostro. En este sentido, el sistema de clasificación Nottingham expresa este grado o nivel en forma de porcentaje, el cual se puede definir como el cociente del parámetro de movilidad del lado afectado entre el parámetro de movilidad del lado sano.

En otras palabras, se considera que el paciente se recupera de la parálisis cuando ambos lados del rostro (afectado y sano) adquieren un nivel de movilidad igual o muy similar entre sí. La determinación de dicho nivel de simetría facial es un reto, ya que es diferente para cada individuo, esto se debe a que no todas las personas con parálisis presentan un mismo grado de movilidad en su lado sano del rostro. En la práctica tradicional esto resulta muy complicado; ya que tal como lo muestra un estudio reciente, se determinó que por lo menos se debe tener 3 milímetros de asimetría facial en la comisura de la boca, en la supraorbital o en ambas partes del rostro para que una persona pueda llegar a percibir la asimetría [11].

La Figura 11 muestra evidencias del funcionamiento del sistema ante el caso de una mujer de 75 años que presenta parálisis facial en el lado izquierdo del rostro.



Figura 11 Fotografías de las cuatro expresiones faciales obtenidas por el sistema

Después de efectuar el procedimiento necesario, el sistema dio como resultado un valor de 37.46% de simetría facial, este resultado fue corroborado por un médico especialista el cual aprobó el correcto funcionamiento del sistema, al ser un valor que representa con certeza el nivel de parálisis.

VII. CONCLUSIONES

Actualmente la tecnología es un factor vital para el desarrollo y evolución de una sociedad con anhelo de crecimiento. La tecnología se ha convertido en un componente esencial en todas las áreas, pero se hace imprescindible en el área de la salud. Hoy por hoy, se llegan a utilizar los avances tecnológicos no sólo para curar sino también para prevenir enfermedades; y posteriormente para todo tipo de investigación médica, la cual gracias a la tecnología ha realizado importantes descubrimientos. Con el trabajo realizado se comprobó la multifuncionalidad del *Kinect*; la variedad en sus características permitió modelar una solución al problema de manera objetiva y sencilla, demostrando ser una herramienta multiusos en más de un área.

REFERENCIAS

- [1] Banta H. D. & Behney C. J. (1981). Policy Formulation and technology assessment. *Milbank Mem Fund Quartely*, 59(3), 445-79.
- [2] Lewis T. (1995). *The lives of a cell: Notes of a biology watcher*. Penguin Books.
- [3] Retting R. A. & Hartman A. J. (1979). The development of medical technology: a policy perspective. *Medical Technology*. National Center for Health Services Research. Research Proceedings Series.
- [4] Rodríguez M. D., Mangas S., Ortiz M. G., Rosete H. S., Vales O. & Hinojosa R. (2011). Parálisis facial periférica. Tratamientos y consideraciones. *Arch Neurocién (Mex) INNN*. 16(3). 148-155.
- [5] Kang T. S., Vrabec, J. T., Giddings, N., Terris, D. J. (2002). Facial nerve grading systems (1985-2002): beyond the House-Brackmann scale. *Otol Neurotol*. 23, 767-771.
- [6] Burres S. & Fisch U. (1986). The comparison of facial grading systems. *Arch Otolaryngol Head Neck Surg*. 112, 755-758.
- [7] Murty G. E., Diver J. P., Kelly P. J., O'Donoghue G. M. & Bradley P. J. (1994). The Nottingham system: objective assessment of facial nerve function in the clinic. *Otolaryngol Head Neck Surg*. 110, 156-161.

- [8] Kinect for Windows Sensor (2015). Recuperado el 17 de marzo del 2015, de <https://msdn.microsoft.com/es-mx/library/hh855355.aspx>.
- [9] Philippe M. (2012). Tratamiento de la Parálisis Facial en medicina, fisioterapia y acupuntura. Recuperado el 23 de junio del 2015, de <http://www.efisioterapia.net>.
- [10] Khan H., Yoon-ho H., et al. (2007). Agreement between two different methods for facial nerve conduction study in patients with acute Bell's palsy. Department of Neurology, Seoul National University Hospital.
- [11] Chu E. A., Farrag T. Y., et al. (2011). Threshold of Visual Perception of Facial Asymmetry in a Facial Paralysis Model. *Arch Facial Plast Surg*. 13(1), 14-19.

ANEXOS II: VINCULACIÓN



Instituto Tecnológico de Apizaco

A quien corresponda

31 de julio de 2014

PRESENTE

Por este conducto reciba un atento saludo de parte de quien abajo firma y en términos de este escrito hago constar que el Ing. Raúl Cid Carro participó en el desarrollo del proyecto denominado "IDENTIFICACIÓN DE UNA ENTIDAD MEDIANTE EL ANÁLISIS DE SUS PUNTOS CARACTERÍSTICOS". El desarrollo de este proyecto se llevó a cabo en nuestra oficina ubicada en el Parque Tecnológico de Puebla.

La fecha de realización de dicha estancia de investigación en acuerdo con el director de tesis; el Dr. Edmundo Bonilla Huerta, fue del 3 de febrero de 2014 al 31 de julio del mismo año.

El objetivo inicialmente planteado fue logrado satisfactoriamente. Por tal motivo se extiende la presente carta de liberación de estancias.

Quedo a su disposición para cualquier pregunta o comentario en los números aquí mencionados.

Atentamente

Dr. Juan Manuel Ahuactzin Larios

Probayes Director General ProbaYes Américas
PROBAYES AMERICAS S.A. DE C.V.
Via Atlxclayotl 2301 • Parque Tecnológico de Puebla
San Andrés Cholula, Puebla, C.P. 72800, México
Tel: +52 (222) 303 22 81 • www.probayes.com

+52(222)3032281

PROBAYES SAS
345, rue Lavoisier - Inovallée
38330 - Montbonnot - France
Phone +33 (0)4 76 42 64 13
Fax +33 (0)4 76 42 49 90

PROBAYES AMERICAS S.A. DE C.V.
Via Atlxclayotl 2301
Parque Tecnológico de Puebla
Puebla, Puebla,
C.P. 72453, Mexico
Phone: +52 222 303 22 81



H. AYUNTAMIENTO PANOTLA

Administración 2014-2016



PANOTLA TLAXCALA A 21 DE OCTUBRE DE 2015

M. A. D. ACELA DÁVILA JIMÉNEZ
JEFA DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
DEL INSTITUTO TECNOLÓGICO DE APIZACO
PRESENTE

Sirva la presente por una parte para hacer constar que el C. Ingeniero RAUL CID CARRO se presentó en nuestras instalaciones, en carácter de estudiante de Maestría en Sistemas Computacionales del Instituto Tecnológico de Apizaco, para solicitar se le permitiera poner a prueba un sistema de carácter tecnológico, útil en el diagnóstico y valoración de avances en la rehabilitación de personas con parálisis facial.

Considerando la importancia que tiene este tipo de investigación, no se encontró inconveniente para proporcionarle las facilidades para su trabajo. Dicho sistema fue probado favorablemente con distintas personas del municipio de Panotla que acudieron al servicio de terapia de recuperación por parálisis facial durante los meses de septiembre y octubre del presente año, de las cuales se logró que la Señora Lucina Juárez de Pelcastre y el Señor Roberto Ordoñez dieran su autorización para el uso de sus datos particulares, lo que no sucedió en otros casos.

Por otra parte; otorgamos al mismo tiempo la presente carta de satisfacción por el servicio recibido solicitando que de ser posible se siga considerando nuestro departamento en investigaciones posteriores,

Sin otro particular manifiesto mi reconocimiento al trabajo que se realiza por su Institución

Atentamente,

Médico General: *Graciela García Pérez*
Responsable de UBR Panotla.
Ced. Prof. 09126164.

"Juntos por el rescate del Municipio"

Palacio Municipal S/N, Panotla, Tlax. C.P. 90140 Tel.: (01246) 41-6-82-97 y 41-6-85-55
e-mail: panotla2014-2016@hotmail.com