



TELECOMUNICACIÓN

Campus Sur
POLITÉCNICA

MEMORIA DE PRÁCTICA EXTERNA
ETS DE INGENIERÍA Y SISTEMAS DE
TELECOMUNICACIÓN
UPM

Estefanía Ferreira Cano



ÍNDICE DE CONTENIDOS

DATOS IDENTIFICATIVOS DE LAS PRÁCTICAS EXTERNAS:.....	1
1. INTRODUCCIÓN:.....	2
2. INFORMACIÓN DE LA ENTIDAD COLABORADORA	3
3. ENMARCAR LAS PRÁCTICAS EN EL CONTEXTO DE LA ENTIDAD	4
4. OBJETIVOS DE LAS PRÁCTICAS, TAREAS Y ACTIVIDADES REALIZADAS	5
TECNOLOGÍAS Y MEDIOS TÉCNICOS UTILIZADOS.....	9
5. COMPETENCIAS Y HABILIDADES ADQUIRIDAS CON LAS PRÁCTICAS	10
6. CONCLUSIONES	11
7. DIARIO DE PRÁCTICAS,.....	12

INSTRUCCIONES:

Se recomienda completar una cara por cada uno de los epígrafes a desarrollar, a excepción del nº4, "Objetivos de las prácticas...", que debe tener mayor contenido (dos caras)

1. INTRODUCCIÓN:

Las prácticas se han desarrollado en la empresa CITSEM (Centro de Investigación en Tecnologías Software y Sistemas Multimedia para la Sostenibilidad). En concreto, el trabajo realizado está englobado en el área de investigación: Grupo de Diseño Electrónico y Microelectrónico (GDEM).

La actividad llevada a cabo en estas prácticas se enmarca en un entorno de investigación. En concreto, se ha basado en el reconocimiento de expresiones faciales de una persona a tiempo real.

Para la realización de este fin, los medios a emplear han sido Kinect y los entornos de desarrollo de aplicaciones para este sistema de captación de información.

Con el fin de llegar a estos objetivos, se ha realizado un estudio previo sobre los patrones de reconocimiento facial y de expresiones, empleando los trabajos de investigación realizados en este centro anteriormente, así como, una investigación propia de este tema. Tras este estudio, se comenzó una investigación sobre cómo Kinect reconocía estas expresiones, caracterizando todo su funcionamiento.

Cuando se alcanzó una base sólida sobre el funcionamiento de Kinect, se procedió a implementar nuevas funciones de reconocimiento facial, en concreto, mediante dos métodos de reconocimiento de expresiones faciales: una basada en puntos de coordenadas 2D en las zonas más significativas de la cara, y otra basada en el reconocimiento de expresiones por movimientos faciales, capturados por una malla 3D.

Para la implementación de estas aplicaciones, se realizó un estudio de la SDK de Kinect, entorno de desarrollo que se empleó para el desarrollo de estas prácticas.

Tras la implementación de ambos métodos y la comparación, se llegó a una conclusión final.

2. INFORMACIÓN DE LA ENTIDAD COLABORADORA

La entidad donde se han llevado las prácticas ha sido CITSEM (Centro de Investigación en Tecnologías Software y Sistemas Multimedia para la Sostenibilidad). Este centro de investigación tiene como objetivo el aprovechamiento de las tecnologías básicas en las que sus miembros tienen experiencia, tecnologías software y multimedia, en un contexto de sostenibilidad.

Presenta cinco misiones principales, en las que se enfocarán todas las distintas áreas de trabajo que se llevan a cabo en dicha entidad. Estas misiones son:

- La realización por propia iniciativa de proyectos y otras tareas de investigación, desarrollo e innovación en el ámbito de las tecnologías software y multimedia con especial orientación hacia la sostenibilidad, así como la difusión de los resultados alcanzados con el fin de incrementar el conocimiento científico y los recursos tecnológicos al alcance de su entorno socio-económico.
- Colaborar con otras entidades (públicas o privadas) en la realización de proyectos de investigación, desarrollo y/o innovación en el ámbito de las tecnologías software y multimedia que contribuyan al logro de sus objetivos empresariales o de servicio público.
- Realizar trabajos específicos, informes o labores de asesoría para empresas privadas o instituciones públicas en el ámbito de las tecnologías software y multimedia.
- Impartir formación de postgrado tanto para salida profesional como investigadora (doctorado).
- Organizar e impartir cursos de especialización y perfeccionamiento, seminarios, conferencias y otras actividades de análoga naturaleza, en el ámbito de las tecnologías software y multimedia.

Las áreas de trabajo en las que está dividido el CITSEM para el desarrollo de su investigación son: Grupo de Tecnología Software y Sistemas (SYST), Grupo de Diseño Electrónico y Microelectrónico (GDEM), y Grupo de Redes y Servicios de Próxima Generación (GRyS); el primero de la ETSISI y los otros dos de la ETSIST. En total cuenta con más de cuarenta investigadores a tiempo completo y se integra, como un actor más, en la estrategia de promover y fortalecer la I+D+i que se hace en el Campus Sur de la UPM.

3. ENMARCAR LAS PRÁCTICAS EN EL CONTEXTO DE LA ENTIDAD

La entidad donde se han desarrollado las prácticas, cubre tres áreas de investigación:

- Grupo de Tecnología Software y Sistemas (SYST).
- Grupo de Diseño Electrónico y Microelectrónico (GDEM).
- Grupo de Redes y Servicios de Próxima Generación (GRyS) que abarcan distintas temáticas: Tecnologías de Imagen y Vídeo, Redes y servicios y Tecnologías de software y servicios.

En el presente caso, se ha trabajado en el área GDEM. En este área se ha realizado un estudio y pruebas de reconocimiento facial.

El espacio de trabajo se ha situado en un laboratorio del centro de investigación CITSEM. Para la puesta en común con el resto del grupo de trabajo y tutora, se realizó en el mismo centro.

Las personas con las que se han trabajado en dicha entidad han sido:

- Juan Manuel Meneses Chaus: Director del CITSEM.
- Martina Eckert: Doctor del CITSEM, tutora de prácticas.

4. OBJETIVOS DE LAS PRÁCTICAS, TAREAS Y ACTIVIDADES REALIZADAS

En las prácticas cursadas, el objetivo principal se ha centrado en conseguir un reconocimiento de expresiones faciales a través de la información recogida por Kinect.

Para conseguir este objetivo principal, se han presentado una serie de objetivos a realizar:

- Estudio de reconocimiento de expresiones faciales.
- Estudio del funcionamiento de Kinect.
- Estudio del entorno de desarrollo SDK.
- Estudio de las implementaciones realizadas para reconocimiento de expresiones faciales.
- Análisis del reconocimiento de expresiones faciales de Kinect.
- Estudio de lenguaje C#.
- Implementación de nuevas posibilidades para el reconocimiento facial y de expresiones.

En el primer periodo de estas prácticas, se realizó el estudio del reconocimiento facial de expresiones. Para alcanzar los conocimientos necesarios, se estudió documentación realizada con anterioridad en esta entidad, sobre el tema que concierne a estas prácticas. También se realizó una investigación para ampliar esta información y profundizar en el tema.

Tras asentar los conocimientos necesarios sobre reconocimiento facial de expresiones, se centró la actividad de las prácticas en saber cómo funciona Kinect. Para este fin, se volvió a acudir a documentación realizada en CITSEM anteriormente, así como una investigación propia. Cuando se obtuvo una idea de cómo detecta un cuerpo y, en concreto, la cabeza del usuario, se llevó a cabo una investigación más exhaustiva sobre la detección de las partes integrantes de la cara con Kinect y de las implementaciones de reconocimiento facial, en SDK.

Con los anteriores objetivos conseguidos, se procedió a comenzar con las implementaciones de reconocimiento de expresiones faciales en tiempo real. Para comenzar con esta actividad, se decidió trabajar con el lenguaje C# y se realizó un breve estudio del mismo. Centrándose en un programa de ejemplo aportado por SDK, donde se empleaba la librería de SDK para reconocimiento facial, se comenzó a implementar pequeñas pruebas, para familiarizarse con el entorno.

La primera implementación de reconocimiento de expresiones faciales que se finalizó, se basaba en el reconocimiento de expresiones faciales mediante 5 puntos 2D estratégicos en ciertas partes de la cara. Los puntos escogidos fueron los mostrados en la figura 1.

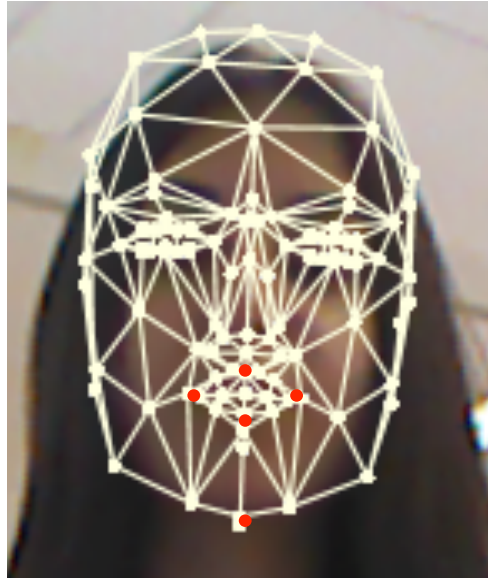


Figura 1. Puntos escogidos para reconocimiento de expresiones.

Las expresiones reconocidas en esta implementación son: sonrisa, abrir boca, subir cabeza y bajar cabeza. El programa muestra un mensaje al realizar alguna de estos movimientos por parte del usuario. Para la detección de estas expresiones, se captura la información de las posiciones de los puntos al arrancar el programa, donde el usuario debe mostrar una cara neutral. A lo largo del funcionamiento de la aplicación, se comparan las posiciones espaciales de los puntos capturados al comienzo, con los del instante actual, para saber la variación espacial y con ello el movimiento producido. El resultado de la aplicación se muestra en la figura 2.



Figura 2. Aplicación basada en puntos.

La siguiente aplicación realizada ha sido el reconocimiento de emociones mediante unidades de acción faciales. SDK para Kinect, en su librería especializada en reconocimiento facial, implementa métodos para este fin. Mediante la utilización de éstos, se realizó una aplicación que detectaba las 6 emociones más estudiadas en reconocimiento facial: felicidad, disgusto, tristeza, enfado, sorpresa y miedo. Basándose en trabajos anteriores del CITSEM y en la documentación propia de las librerías de la SDK, se pudo llevar a cabo esta implementación, relacionando las unidades de acción proporcionadas por SDK (figura 3) con las emociones (tabla 1).

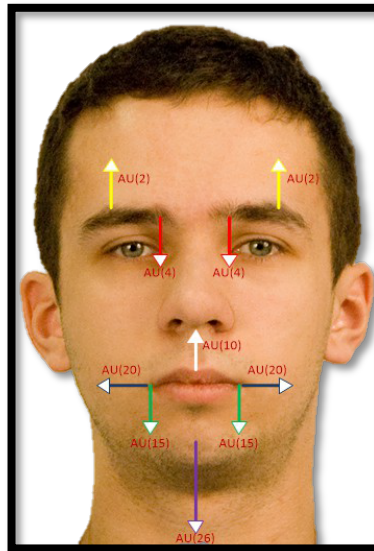


Figura 3. Action Units detectadas mediante librerías de SDK.

Tabla 1. Relación entre AUs y emociones.

Emoción	AUs (Imagen)
Fear	4,2,15,20
Anger	4,15, 20
Sadness	4,15
Disgust	2,4,10,26,20
Happy	2, 4, 15,20, 26
Surprise	2,4,26

La aplicación final se basa en la comparación del rostro neutral del usuario al arranque del programa, con los cambios de emociones que sufra en el instante actual. Se muestra un ejemplo de detección en la figura 4.

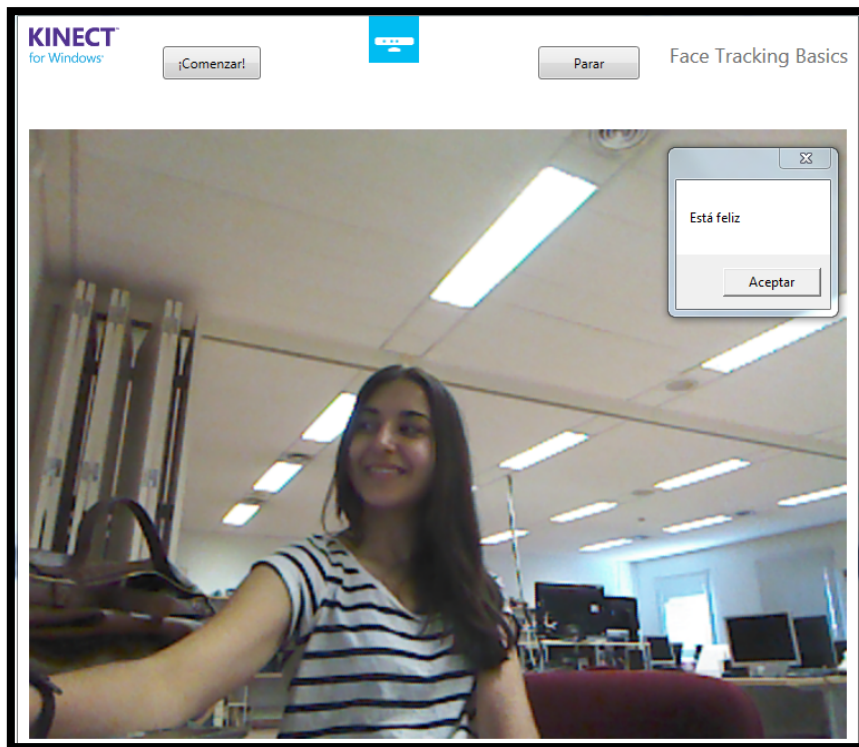


Figura 4. Aplicación basada en AUs.

Tras la implementación de ambas aplicaciones, se comparó su precisión y los resultados satisfactorios que aportaban cada una de ellas, llegando a la conclusión, que la mejor opción para reconocimiento de expresiones faciales era el reconocimiento basado en puntos.

TECNOLOGÍAS Y MEDIOS TÉCNICOS UTILIZADOS

- Ordenador.
- Entorno de desarrollo SDK v 1.8.
- Kinect versión 1.5.
- Microsoft Visual Studio, para poder desarrollar en C#.
- Bases de datos científicas para obtener la información correspondiente al estado del arte de las tecnologías concretas. Esta información incluye papers, surveys...
- Windows Office, puesto que quincenalmente se realizan presentaciones del trabajo desarrollado hasta el momento, además de una documentación pertinente a los progresos realizados.

5. COMPETENCIAS Y HABILIDADES ADQUIRIDAS CON LAS PRÁCTICAS

Las competencias adquiridas en las prácticas realizadas son:

- Conocimientos avanzados del área de Ingeniería de Telecomunicación con capacidad para aplicarlos de una forma profesional y para innovar y desarrollar elementos y sistemas en ese área.
- Capacidad de búsqueda y selección de información, de razonamiento crítico y de elaboración y defensa de argumentos dentro del área.
- Capacidad para expresarse correctamente de forma oral y escrita y transmitir información mediante documentos y exposiciones en público.
- Capacidad de abstracción, de análisis y de síntesis y de resolución de problemas.
- Capacidad de trabajo en equipo y en entornos multidisciplinares.
- Habilidades para la utilización de las Tecnologías de la Información y las Comunicaciones.
- Conocimientos básicos sobre el uso y programación de los ordenadores y programas informáticos con aplicación en ingeniería.
- Capacidad para aprender de manera autónoma nuevos conocimientos y técnicas adecuados para la concepción, el desarrollo o la explotación de sistemas y servicios de telecomunicación.
- Capacidad de utilizar la programación orientada al procesado digital de imagen.
- Capacidad de utilizar aplicaciones de comunicación e informáticas.

Las asignaturas impartidas en la ETSIST relacionadas con el trabajo que se ha llevado a cabo son:

- Programación I.
- Programación II.
- Comunicación Profesional.
- Técnicas de Búsqueda y Sistemas de Información
- Fundamentos de Sonido e Imagen.
- Tecnologías de Imagen y Vídeo.
- Tratamiento Digital de la Imagen.

6. CONCLUSIONES

Los objetivos y el plan de trabajo planteado al comienzo de las prácticas se ha cumplido. Para llegar a este fin, se han adquirido habilidades de búsqueda de información, así como un aprendizaje del proceso de investigación científica.

Se ha aprendido un lenguaje nuevo de programación, como es el C#.

Para el desarrollo de aplicaciones para Kinect, se ha familiarizado con la misma, además de aprender a trabajar con entornos de desarrollo especializados. En concreto, se ha adquirido conocimientos sobre el desarrollo de aplicaciones en SDK para Kinect.

Se ha adquirido una base de reconocimiento facial mediante tratamiento digital de la imagen, tanto fija como en tiempo real, mediante la captación continua de las mismas.

Además de los conocimientos técnicos, se ha adquirido un aprendizaje sobre el trabajo de investigación en grupo, y la comunicación de conocimientos a compañeros y tutor, mediante presentaciones y documentación.

Particularizando en la labor del tutor y su función desempeñada en estas prácticas, ha sido fundamental en esta investigación. Su continuo interés y ayuda, tanto a la hora del estudio de las herramientas y conocimientos a adquirir, como a la hora de realizar las implementaciones posteriores.

La tutora ha estado en todo momento pendiente del trabajo realizado y de los problemas que han ido surgiendo a lo largo de las prácticas. En todo momento ha prestado su ayuda, o se ha encargado de buscar ayuda externa para solventar cualquier duda o complicación en el proyecto.

7. DIARIO DE PRÁCTICAS,

9-13 de Febrero

- Reunión con tutora para ver progresos.
- Estudio de anteriores trabajos relacionados con el reconocimiento facial en el CITSEM.

16-20 de Febrero

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Investigación de funcionamiento de Kinect.

23-27 de Febrero

- Reunión con tutora para ver progresos.
- Investigación de librerías de desarrollo para Kinect.
- Investigación reconocimiento facial en Kinect.

2-6 de Marzo

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Estudio de aplicaciones ya implementadas para Kinect anteriormente en CITSEM.

9-13 de Marzo

- Reunión con tutora para ver progresos.
- Estudio de lenguaje C#.
- Estudio de librerías de SDK.

16-20 de Marzo

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Implementación de una aplicación para detectar puntos faciales en 2D.

23-27 de Marzo

- Reunión con tutora para ver progresos.
- Análisis de la información que detecta Kinect con las librerías implementadas en SDK.

30-3 de Abril

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Implementación de una aplicación que detecte ciertas emociones mediante puntos.

6-10 de Abril

- Reunión con tutora para ver progresos.
- Implementación de una aplicación que detecte ciertas emociones mediante puntos.
- Análisis de resultados de la aplicación.

13-17 de Abril

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Estudio detenido de las unidades de acción faciales.
- Estudio de los métodos y clases para detectar unidades de acción en las librerías implementadas en Kinect.

20-24 de Abril

- Reunión con tutora para ver progresos.
- Implementación de aplicación para detección de emociones mediante unidades de acción.

27-1 de Mayo

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Implementación de aplicación para detección de emociones mediante unidades de acción.

4-8 de Mayo

- Reunión con tutora para ver progresos.
- Implementación de aplicación para detección de emociones mediante unidades de acción.

11-15 de Mayo

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Análisis de resultados y precisión de la detección de emociones implementada.

18-22 de Mayo

- Reunión con tutora para ver progresos.
- Comparación de resultados obtenidos a lo largo del periodo de trabajo.
- Elaboración de una documentación recogiendo los resultados obtenidos.

25-29 de Mayo

- Reunión con resto de integrantes del mismo área de investigación y tutores.
- Reunión con tutora para ver progresos.
- Realización de presentación final de prácticas.
- Presentación del trabajo realizado en Jornada de presentaciones en la ETSIST.

1-5 de Junio

- Reunión con tutora para ver progresos y resultados finales.
- Realización de informe final.
- Mejora de la implementación de reconocimiento de expresiones mediante puntos.