



TRABAJO FIN DE GRADO
INGENIERÍA EN INFORMÁTICA

Reconocimiento de Expresiones Faciales

por Computador

Autor

Francisco José Fajardo Toril

Directores

Miguel García Silvente

Eugenio Aguirre Molina



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

Granada, 15 de agosto de 2017



Reconocimiento de Expresiones Faciales

por Computador.

Autor

Francisco José Fajardo Toril

Directores

Miguel García Silvente

Eugenio Aguirre Molina

Reconocimiento de Expresiones Faciales: por Computador

Francisco José Fajardo Toril

Palabras clave: Reconocimiento, expresión, facial, visión, computador, aprendizaje, automático, redes, neuronales, convolucionales.

Resumen

El objetivo de este documento es mostrar al lector el proceso seguido para el desarrollo de una aplicación software capaz de resolver el problema del reconocimiento de expresiones faciales por computador, que puede ser descrito como:

Dada una imagen de entrada, obtener la etiqueta de salida que mejor identifique la expresión facial que aparece en dicha imagen.

De entre todas las técnicas posibles para la resolución del problema, he basado mi solución en el entrenamiento una red neuronal convolucional. El modelo entrenado clasifica la imagen de entrada mediante la asignación de una (o varias) etiquetas, incluyendo un porcentaje por cada una que indica cuál es la más probable de estar presente en dicha imagen. El conjunto de posibles etiquetas es:

- | | |
|-------------|-------------|
| ■ Afraid | ■ Neutral |
| ■ Angry | ■ Sad |
| ■ Disgusted | ■ Surprised |
| ■ Happy | |

Facial Expression Recognition: by Computer

Francisco José Fajardo Toril

Keywords: Facial, expression, recognition, computer, vision, machine, learning, convolutional, neural, network.

Abstract

This document main objective is to show the reader the process I followed to develop a software application which is able to solve the face expression recognition problem through computer. It can be described as:

To obtain the tag that better describe the facial expression that appears in the given input image.

Between all the possible techniques that allow us to solve this problem, I chose to train a convolutional neural network. Trained model classify the input image by the assignment of one or few tags, including a percentage that suggest which is the most probable tag to appear in that image. Here is the whole set of possible tags:

- | | |
|-------------|-------------|
| ▪ Afraid | ▪ Neutral |
| ▪ Angry | ▪ Sad |
| ▪ Disgusted | ▪ Surprised |
| ▪ Happy | |

Yo, **Francisco José Fajardo Toril**, alumno de la titulación Grado en Ingeniería Informática de la **Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación de la Universidad de Granada**, con DNI 76424577Q, autorizo la ubicación de la siguiente copia de mi Trabajo Fin de Grado en la biblioteca del centro para que pueda ser consultada por las personas que lo deseen.

Fdo: Francisco José Fajardo Toril

Granada a 15 de agosto de 2017.

D. **Miguel García Silvente**, Profesor del Área de Ciencias de la Computación e Inteligencia Artificial del Departamento YYYY de la Universidad de Granada.

D. **Eugenio Aguirre Molina**, Profesor del Área de Ciencias de la Computación e Inteligencia Artificial del Departamento YYYY de la Universidad de Granada..

Informan:

Que el presente trabajo, titulado ***Reconocimiento de Expresiones Faciales, por Computador***, ha sido realizado bajo su supervisión por **Francisco José Fajardo Toril**, y autorizamos la defensa de dicho trabajo ante el tribunal que corresponda.

Y para que conste, expiden y firman el presente informe en Granada a 15 de agosto de 2017.

Los directores:

Miguel García Silvente Eugenio Aguirre Molina

Agradecimientos

Poner aquí agradecimientos...

Capítulo 1

Introducción

1.1. Motivación

El problema del reconocimiento de expresiones faciales es un paso adelante tras resolver el problema del reconocimiento facial en imágenes. Tal y como mencioné en el resumen puede ser descrito de la siguiente manera:

Dada una imagen de entrada, obtener la etiqueta de salida que mejor identifique la expresión facial que aparece en dicha imagen.

A priori la solución para un ser humano es trivial. Tenemos muchas maneras de adivinar el estado de ánimo de una persona, pero si tenemos que basarnos en nuestra visión, una posible solución sería: *"Simplemente observa los rasgos faciales del sujeto e intenta adivinar su estado de ánimo basándote en tu experiencia pasada con otros seres humanos."*

Imagine que deseamos desarrollar un agente basado en inteligencia artificial para que interactúe con otros seres humanos. Sería de gran utilidad que este agente basara sus acciones en función del estado de ánimo de la persona/as con las que está comunicándose para lograr de esta forma un mayor grado de comprensión o empatía. De la misma manera podríamos pensar en una aplicación que escoge la música idónea para nosotros en función de nuestro estado de ánimo en ese momento a través de una captura mediante la cámara del computador.

1.1.1. Trabajos relacionados

Como se puede ver, la solución a este problema puede ser de utilidad en diversas aplicaciones del mundo real, sin embargo, ¿Cómo podría un computador resolver este problema? Hay diversas propuestas que han tratado de resolver este problema.



Figura 1.1: Puntos de Referencia extraídos con ASM.

Facial Expression Analysis using Active Shape Model[1]

Active Shape Model(ASM)[3] es una técnica que permite, dada una imagen del objeto para el que fue entrenado, localizar un conjunto de puntos de referencia previamente definidos tal y como se puede ver en la Figura 1.1. En el caso de un rostro humano estos puntos pueden estar los principales indicadores de la expresión facial, como pueden ser las cejas, boca, ojos... En este paper [1] en concreto, se tienen en cuenta para cada cara, 68 puntos de referencia. Se procesan los puntos extraídos y se obtiene un modelo geométrico para ese rostro concreto. Acto seguido, haciendo uso de la técnica de clasificación conocida como Support Vector Machine(SVM), se clasifica ese modelo geométrico para predecir a qué expresión facial pertenece obteniendo un porcentaje de acierto del 92,1 %.

Facial expression recognition and synthesis based on an appearance model[2]

En este caso se hace uso de la técnica *Active Appearance Model (AAM)*[4], un método estadístico que se emplea para emparejar, un modelo geométrico, a una nueva forma del mismo tipo para el que fue entrenado, pero que el modelo no ha visto antes, como por ejemplo un nuevo rostro humano. Esta técnica (AAM) es una generalización de la ya antes mencionada ASM[3]. Además, lo que en este caso se intenta es que la información usada esté exenta de redundancia por lo que se aplican técnicas de aprendizaje no supervisado en una fase de preprocesamiento de la información como *Análisis de Componentes Principales (PCA)* o *Análisis de Componentes Independientes(ICA)* para reducir la dimensionalidad de la información en una fase de preprocesamiento. En este caso obtienen alrededor de un 84 % de acierto.

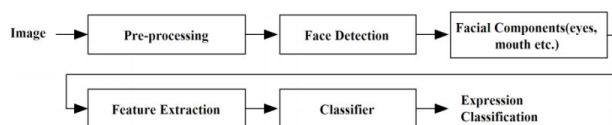


Figura 1.2: Diagrama general del proceso de reconocimiento.

1.1.2. ¿Qué se pretende conseguir con este proyecto?

Las técnicas de ASM[3] y AAM[4] son sensibles a rotaciones de la imagen, es decir, será difícil extraer los puntos de referencia en imágenes que disten mucho del concepto de imagen frontal, como por ejemplo una la fotografía de una persona mirando hacia un lado con una expresión facial determinada. Además, en la Figura 1.2 extraída del siguiente paper[5], se puede ver el proceso general que se sigue a la hora de clasificar una imagen, en nuestro caso reconocimiento de expresiones faciales.

El objetivo de este proyecto es simplificar dicho diagrama para que se asemeje al de la Figura 1.3 de forma que no sea necesario extraer características de la imagen, y que sea más robusto con imágenes inclinadas o rotadas. Para lograr este objetivo y tras observar el auge que han tenido en la resolución de problemas de visión artificial, basaremos nuestra solución en el uso de Redes Neuronales Convolucionales.



Figura 1.3: Diagrama del proceso objetivo.

Capítulo 2

Objetivos

Objetivo general a conseguir: se enuncia lo que se quiere hacer sin entrar en detalles.

Objetivos específicos: Es dividir el objetivo general en los pasos a seguir con sub-objetivos más simples. Dicho de otro modo, son cada uno de los pasos a realizar para alcanzar el objetivo general, es decir solucionando todos y cada uno de los objetivos específicos se resuelve el objetivo general. (Poner entre 3 y 5 como mucho).

Se puede decir en qué apartado se tratará cada objetivo específico.

Poner también los aspectos formativos previos utilizados, por ejemplo si se han usado técnicas de visión concretas como Transformada de Hough, Método de detección de rostros Viola-Jones, Filtro de Partículas, o técnicas de aprendizaje por SVM, etc. Se explica un poco cada método.

Capítulo 3

Resolución del trabajo

Como método de ingeniería del software decir que vamos a seguir la técnicas de modelo de prototipos rápido o también llamado modelado de prototipado rápido.

Ver http://www.ecured.cu/index.php/Modelo_de_Prototipos

3.1. Recursos

Decir los recursos humanos (autor y directores), hardware y software que se van a utilizar.

3.2. Especificación de requisitos

Decir que se partió de una especificación inicial de requisitos que a medida que se fueron implementando los prototipos se fue refinando posteriormente. Se puede poner la inicial y la final o solo la final indicando que se están poniendo los requisitos que finalmente tiene que tener el sistema.

Los requisitos se pueden referir a las necesidades del usuario del sistema (requisitos del usuario), a lo que tiene que hacer la aplicación (requisito funcional) o a cómo tiene que hacerlo (requisito no funcional). Ejemplo:

En este sistema un robot tiene que coger con sus pinzas un envase de medicamento y llevárselo a una persona anciana que por sí misma no puede recordar su medicación.

Requisitos del usuario:

RU1. La persona puede moverse libremente por una habitación donde está el robot.

RU2. La persona es capaz de coger el medicamento cuando se lo ofrece el robot.

RU3. La persona es capaz de tomarse el medicamento por sí misma.

Requisitos funcionales.

RF1. El robot mediante la cámara kinect debe poder localizar a la persona.

RF2. El robot conoce la posición de la mesa pues tiene un mapa de la habitación.

RF3. El robot debe identificar el medicamento correcto según un plan de medicación previamente establecido.

RF4. El robot debe poder coger el medicamento con sus pinzas.
etc...

Requisitos no funcionales

RNF1. El robot no puede atropellar ni dañar a la persona en ningún momento.

RNF2. La aplicación debe ejecutarse en entornos linux

RNF3. La aplicación debe utilizar pocos recursos para reaccionar con rapidez.

algo de la interfaz, como tratar posibles fallos, etc.

3.3. Planificación

Poner una tabla de tiempos con las planificación del proyecto diciendo cuando se tiene previsto alcanzar cada subobjetivo planteado. Con su correspondiente división en fases y tareas, y la posterior comparación con los datos reales obtenidos tras realizar el proyecto. Entre las fases está la realización de los diferentes prototipos I, II y III por ejemplo.

Poner presupuesto según horas de trabajo estimadas.

3.4. Análisis funcional

A partir de aquí nos referimos solamente al prototipo final que da lugar a la aplicación final.

Hay que describir la funcionalidad que debe poseer el sistema para poder cumplir con los objetivos y requisitos que se han dicho previamente. La descripción de esta funcionalidad puede hacerse analizando las tareas (que aparecerán en la planificación) y estudiando la inter-relación entre ellas y sus conexiones.

Para la realización de este análisis se pueden utilizar Diagramas de Flujo para poder conocer generalmente un único punto de inicio y un único punto de término o en varios.

https://es.wikipedia.org/wiki/Diagrama_de_flujo

Se pueden plantear también casos de uso. Los diagramas de casos de uso sirven para describir la inter-relación entre el sistema y el usuario del mismo. Se pueden utilizar para plantear diferentes casos de interacción entre el robot y la persona y cómo tiene que reaccionar el sistema en cada caso.

https://es.wikipedia.org/wiki/Diagrama_de_casos_de_uso

3.5. Implementación y pruebas

Decir qué lenguaje de programación se ha utilizado y las tecnologías implicadas aunque se hayan comentado en el apartado de recursos. Justificar su uso (rendimiento, disponibilidad, etc.). Si se ha usado open source decirlo y explicar las ventajas.

Las pruebas se realizan para comprobar la verificación y validación del producto software. La verificación consiste en comprobar que el producto realiza lo que está programado, es decir la programación no tiene errores y funciona en todos los casos cumpliendo los requisitos. La validación tiene que ver con que cumpla con lo que espera el usuario.

Verificación y Validación: Conjunto de procesos de comprobación y análisis que aseguran que el software que se desarrolla está acorde a su especificación y cumple las necesidades de los clientes.

Verificación: ¿Estamos construyendo el producto correctamente? Se comprueba que el software cumple los requisitos funcionales y no funcionales de su especificación.

Validación: ¿Estamos construyendo el producto correcto? Comprueba que el software cumple las expectativas que el cliente espera. Importante: Nunca se va a poder demostrar que el software está completamente libre de defectos

las pruebas que pueden utilizarse son muy diversas. Aconsejo centrarnos en pruebas de caja blanca y de caja negra.

Las pruebas de la caja blanca se centran en la estructura interna del programa para elegir los casos de prueba. El objetivo de estas pruebas consiste en probar todos los posibles casos de ejecución de la aplicación para comprobar que los datos se comportan de manera correcta internamente.

Decir que se han hecho las pruebas de caja blanca.

Las pruebas de caja negra son aquellas que se centran en las salidas y entradas de los módulos, sin atender a su comportamiento interno (comprobando mediante las pruebas de caja blanca). Las pruebas de caja negra garantizan la interconectividad entre los diferentes módulos de la aplicación, así como su correcto funcionamiento final.

Poner algunos casos de prueba de caja negra.

Capítulo 4

Conclusiones y trabajo futuro

Decir lo que se ha conseguido realizar comentando sus puntos fuertes y débiles. Decir si se han alcanzado los objetivos específicos y el general propuesto y en qué grado.

Indicar las asignaturas del grado más relacionadas con la ejecución del TFG y cómo el TFG ha ayudado a afianzar los conocimientos adquiridos en el Grado.

Valoración personal si se quiere.

Avanzar algunas líneas de trabajo futuro para solucionar las debilidades detectadas o para conseguir nuevas funcionalidades interesantes.

Capítulo 5

Bibliografía

Bibliografía

- [1] R. Shbib, and S. Zhou. *Facial Expression Analysis using Active Shape Model*, School of Engineering, University of Portsmouth, United Kingdom 2015.
- [2] B. Abboud, F. Davoine, and M. Dang. *Facial expression recognition and synthesis based on an appearance model*, Heudiasyc Laboratory CNRS, University of Technology of Compiègne, BP 20529, 60205 Compiègne Cedex, France 2004.
- [3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. *Active Shape Models - Their Training and Application*, Department of Medical Biophysics, University of Manchester, Oxford Road, Manchester M13 9PT, England 1994.
- [4] T. F. Cootes, G. J. Edwards, and C. J. Taylor. *Active Appearance Models*, Wolfson Image Analysis Unit, Department of Medical Biophysics, University of Manchester, Manchester M13 9PT, UK 1998.
- [5] J. Kumari, R. Rajesh, and KM. Pooja. *Facial expression recognition: A survey*, Dept. of Computer Science, Central University of South Bihar, India 2015.

Capítulo 6

Anexo

Al final de la memoria hay que añadir un anexo de una página o dos, explicando como se usa el software a modo de manual de usuario. Es decir, como se llaman los comandos, qué parámetros hay que darle, como se llaman los ficheros de datos de entrada, etc.

