

INFORME MÁS ALLÁ DEL RECONOCIMIENTO FACIAL Y DE VOZ - SEMINARIO DE TÓPICOS

Fernandez Gastelo Willy
Universidad Nacional Pedro Ruiz Gallo

12 de diciembre de 2018

Índice

1. INTRODUCCION	4
2. APLICACIONES DEL RECONOCIMIENTO FACIAL	4
2.1. Aplicaciones Actuales	4
2.2. Posibles aplicaciones futuras	5
3. FUNCIONAMIENTO	5
4. TECNICAS Y ALGORITMOS	6
4.1. Tecnicas tradicionales	6
4.1.1. Holísticos	6
4.1.2. Locales o geométricos	8
4.2. Técnicas 3D	8
4.3. Técnicas de análisis de la textura de la piel	9
5. RECONOCIMIENTO BASADO EN VIDEO	9
6. ASPECTOS TECNICOS	10
6.1. Debilidades	10
6.1.1. Problema de iluminación	10
6.1.2. Problema del cambio de pose	11
6.1.3. Efectividad	11
6.1.4. Privacidad	11
7. EVALUACION	11
8. MEJORAS RECIENTES	13
9. RECONOCIMIENTO DE VOZ	13
9.1. Diseño de un sistema de RAH	13
9.2. Decodificador acústico-fonético	14
9.3. Modelo del lenguaje	14
9.4. Características de los sistemas existentes	14
9.5. Clasificación	14
9.6. Usos y aplicaciones	14
Referencias	15

Índice de figuras

3.1. funcionamiento	6
4.1. Los primeros 10 a) Eigenfaces, b) Fisherfaces, y c) Laplacianfaces calculados a partir de imágenes de caras de la base de datos de YALE	7

1. INTRODUCCION

El reconocimiento facial se ha convertido en los últimos años en un área de investigación activa que abarca diversas disciplinas, como procesamiento de imágenes, reconocimiento de patrones, visión por ordenador y redes neuronales. Involucra tanto a investigadores del área de informática como a neurocientíficos y psicólogos. Se podría considerar también dentro del campo de reconocimiento de objetos, donde la cara es un objeto tridimensional sujeto a variaciones de iluminación, pose, etc., y ha de ser identificada basada en su proyección 2D (excepto cuando se utilizan técnicas 3D).

El objetivo de un sistema de reconocimiento facial es, generalmente, el siguiente: dada una imagen de una cara "desconocida", o imagen de test, encontrar una imagen de la misma cara en un conjunto de imágenes "conocidas", o imágenes de entrenamiento. La gran dificultad añadida es la de conseguir que este proceso se pueda realizar en tiempo real. El sistema identificará las caras presentes en imágenes o videos automáticamente. Puede operar en dos modos:

Verificación o autenticación de caras: compara una imagen de la cara con otra imagen con la cara de la que queremos saber la identidad. El sistema confirmará o rechazará la identidad de la cara.

Identificación o reconocimiento de caras: compara la imagen de una cara desconocida con todas las imágenes de caras conocidas que se encuentran en la base de datos para determinar su identidad.

Por su naturaleza amigable, este tipo de sistemas siguen siendo atractivos a pesar de la existencia de otros métodos muy fiables de identificación personal biométricos, como el análisis de huellas dactilares y el reconocimiento del iris.

2. APLICACIONES DEL RECONOCIMIENTO FACIAL

2.1. Aplicaciones Actuales

Se utiliza principalmente en sistemas de seguridad para el reconocimiento de usuarios. En estos sistemas se utiliza un lector que define las características del rostro, y cuando este solicita el acceso, se verifica comparando los datos obtenidos con la base de datos. Sin embargo, estos sistemas no son útiles a largo plazo ya que, a medida que pasan los años, los rasgos faciales varían y al solicitar el acceso ya no coinciden con la imagen en la base de datos. Para solucionar este problema se puede utilizar un algoritmo que interprete el paso de los años, aunque igualmente sigue sin ser del todo fiable), o bien, renovar frecuentemente la base de datos.

También se utiliza en aplicaciones de interacción persona-ordenador, en gestión multimedia, y en software como Google's Picasa, Apple iPhoto, Sony's Picture Motion Browser (PMB), Facebook y Asus Smart Logon.

Áreas	Aplicaciones específicas
Biometria	Licencia de Conducir, Programas de Derecho, Inmigración, DNI, Pasaportes, Registro de Votantes, Fraude
Seguridad de la información	Inicio de Sesión, Seguridad en Aplicaciones, Seguridad en Bases de Datos, Cifrado de Información, Seguridad en Internet, Acceso a Internet, Registros Médicos, Terminales de Comercio Seguro, Cajeros Automáticos
Cumplimiento de la ley y vigilancia	Videovigilancia Avanzada, Control CCTV, Control Portal, Análisis Post-event, Hurto, Seguimiento de Sospechosos, Investigación
Tarjetas inteligentes	Valor Almacenado, Autenticación de usuarios
Control de acceso	Acceso a Instalaciones, Acceso a Vehículos

Tabla 2.1: Aplicaciones del reconocimiento facial.

2.2. Posibles aplicaciones futuras

Una aplicación de reconocimiento facial futura se basa en establecer esta técnica a nivel de usuario. Por ejemplo, en un supermercado o en un establecimiento pequeño se podría llevar a cabo un control sobre quien abre la caja registradora mediante un reconocimiento facial previo, de esta manera también se pueden evitar intentos de robo ya que, al no reconocer el rostro, la caja permanecería cerrada.

Un caso más extremo sería en los cajeros automáticos donde, para poder operar, fuese necesario un reconocimiento facial en vez del actual PIN.

3. FUNCIONAMIENTO

El proceso consta de cuatro módulos principales:

Detección de la cara: detecta que hay una cara en la imagen, sin identificarla. Si se trata de un video, también podemos hacer un seguimiento de la cara. Proporciona la localización y la escala a la que encontramos la cara.

Alineación de la cara: localiza las componentes de la cara y, mediante transformaciones geométricas, la normaliza respecto propiedades geométricas, como el tamaño y la pose, y fotometricas, como la iluminación. Para normalizar las imágenes de caras, se pueden seguir diferentes reglas, como la distancia entre las pupilas, la posición de la nariz, o la distancia entre las comisuras de los labios.

También se debe definir el tamaño de las imágenes y la gama de colores. Normalmente, para disminuir la carga computacional del sistema, se acostumbra a utilizar imágenes pequeñas en escala de grises. A veces también se realiza una ecualización del histograma.

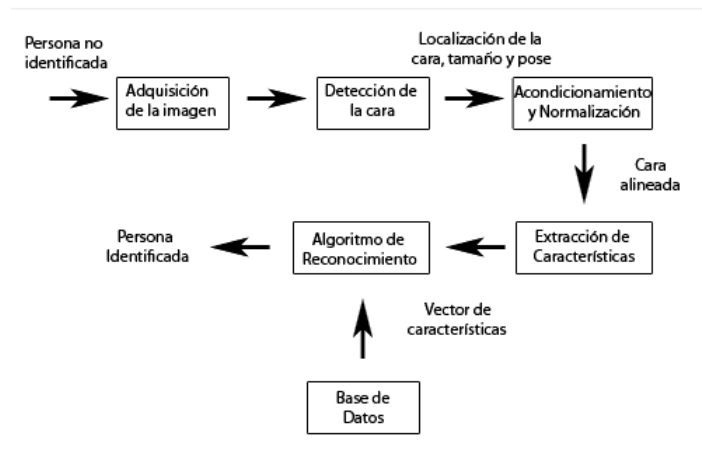


Figura 3.1: funcionamiento

Extracción de características: proporciona información para distinguir entre las caras de diferentes personas según variaciones geométricas o fotométricas.

Reconocimiento: el vector de características extraído se compara con los vectores de características extraídos de las caras de la base de datos. Si encuentra uno con un porcentaje elevado de similitud, nos devuelve la identidad de la cara; si no, nos indica que es una cara desconocida.

Los resultados obtenidos dependen de las características extraídas para representar el patrón de la cara y de los métodos de clasificación utilizados para distinguir los rostros, pero para extraer estas características apropiadamente, hace falta localizar y normalizar la cara adecuadamente.

4. TECNICAS Y ALGORITMOS

Los métodos de reconocimiento facial tradicional se pueden dividir en dos grandes grupos:

4.1. Técnicas tradicionales

Los métodos de reconocimiento facial tradicional se pueden dividir en dos grandes grupos:

4.1.1. Holísticos

Reconocen según toda la imagen facial. Son métodos basados en correlación. El esquema de clasificación más simple, donde se utilizan modelos de comparación para el reconocimiento, es el template matching.

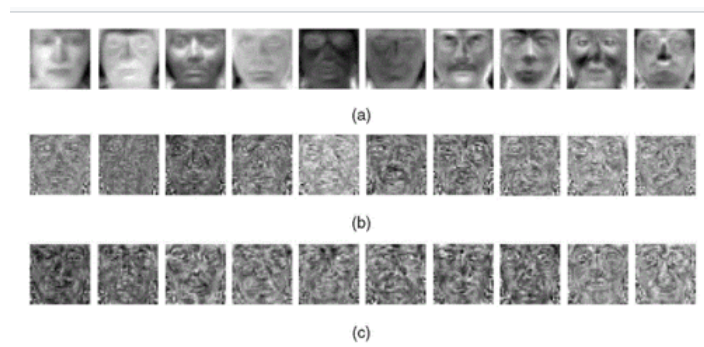


Figura 4.1: Los primeros 10 a) Eigenfaces, b) Fisherfaces, y c) Laplacianfaces calculados a partir de imágenes de caras de la base de datos de YALE

El problema del template matching es que ha de comparar muchas características (para él, un pixel es una característica), y si tenemos en cuenta que en la base de datos encontramos M personas, con N imágenes por persona, observamos que este método no se puede implementar en tiempo real. Por lo tanto, se trabaja con otros métodos que decorrelacionan las características entre sí para conseguir reducir el espacio facial en un número menor de coeficientes, que tengan un alto poder discriminatorio entre las personas.

Es lo que se denomina subespacio facial. Ejemplos de métodos que trabajan a partir de subespacios son el Análisis de Componentes Principales (PCA - Principal Component Analysis) a partir de eigenfaces, el Análisis Linear Discriminant (LDA - Linear Discriminant Analysis) o el Discriminante Linear de Fisher (FLD - Fisher Linear Discriminant) a partir de fisherfaces.

La técnica PCA se considera una de las que proporciona un mayor rendimiento. Funciona proyectando las imágenes faciales sobre un espacio de facciones que engloba las variaciones significativas entre las imágenes faciales conocidas.

Las facciones significativas se llaman eigenfaces, ya que son los eigenvectores, o componentes principales, del conjunto de caras. La proyección caracteriza la imagen facial de un individuo como la suma de los diferentes pesos de todas las facciones y, de la misma manera, para reconocer una imagen facial determinada sólo hará falta comparar estos pesos con aquellos de los individuos conocidos previamente.

No tiene en cuenta la información de qué imágenes pertenecen a un mismo individuo. Es muy sensible a cambios en las condiciones de iluminación en diferentes imágenes de una misma persona.

El método LDA permite utilizar la información entre miembros de la misma clase (imágenes de la misma persona) para desarrollar un conjunto de vectores de características donde las variaciones entre las diferentes caras se enfatizan mientras que los cambios debidos a la

iluminación, expresión facial y orientación de la cara no.

Es decir, maximiza la variancia de las muestras entre clases, y la minimiza entre muestras de la misma clase.

La técnica FLD es equivalente al LDA. Los resultados obtenidos con FLD son bastante mejores que los que podemos obtener con PCA, sobre todo cuando las condiciones lumínicas varían entre el conjunto de imágenes de entrenamiento y de test, y también con cambios de expresión facial, dando más peso a zonas como los ojos, la nariz o las mejillas que a la boca, porque son zonas más invariables en las diferentes expresiones que puede tener una persona.

Otros métodos, en vez de utilizar subespacios faciales, siguen una clasificación por redes neuronales y plantillas deformables, como EGM - Elastic graph matching.

4.1.2. Locales o geométricos

Se comparan diferentes características geométricas de las caras. Existen dos divisiones, la basada en los vectores característicos extraídos del perfil, y la basada en los extraídos a partir de una vista frontal. Se utilizaba mucho anteriormente pero sus resultados no son óptimos.

4.2. Técnicas 3D

Últimamente ha incrementado la tendencia del reconocimiento facial tridimensional, donde se utilizan imágenes 3D tanto en el entrenamiento como en el reconocimiento. Esta técnica utiliza sensores en 3D para captar información sobre la forma de la cara.

Esta información se utiliza posteriormente para identificar rasgos característicos del rostro como por ejemplo la barbilla, el contorno de los ojos, la nariz o los pómulos, y reteniendo información espacial, a parte de la textura y la profundidad.

Una ventaja del reconocimiento facial en 3D es que no les afectan los cambios de iluminación, como pasa en el caso de otras técnicas. Además, otro punto a favor es que pueden reconocer una cara en diferentes ángulos, incluso de perfil. El problema es que es difícil obtener imágenes 3D fidedignas en la fase de reconocimiento, ya que los sensores 3D tienen que estar muy bien calibrados y sincronizados para adquirir la información correctamente.

Es por eso que se utiliza el método de Anàlisi de Components Principals Parcial (P2CA - Partial Principal Component Analysis), derivado del PCA, donde se utilizan imágenes en 3D en la fase de entrenamiento y en la base de datos, pero en la fase de test puede utilizar tanto imágenes en 2D como en 3D. La técnica intenta reconstruir modelos faciales en 3D a partir de múltiples imágenes de la misma persona adquiridas mediante un sistema multicámara o a partir de aparatos 3D. Las imágenes 3D son imágenes de 180° en coordenadas cilíndricas. Otros ejemplos de técnicas 3D son 3-D Morphable Model i 3-D Face Recognition.

4.3. Técnicas de análisis de la textura de la piel

Esta tendencia utiliza los detalles visuales de la piel. Analiza las líneas únicas, patrones y detalles evidentes como manchas y/o cicatrices del rostro del sujeto.

Al utilizar este algoritmo nos ahorramos tener que recorrer toda la base de datos ya que podemos descartar imágenes fácilmente. Hay estudios que demuestran que utilizando esta técnica, juntamente con el reconocimiento facial, el rendimiento puede aumentar hasta un 25 por ciento.

5. RECONOCIMIENTO BASADO EN VIDEO

Se han desarrollado algunos sistemas de reconocimiento facial basados en video, por ejemplo, salas inteligentes que pueden reconocer a las personas e iniciar automáticamente las acciones apropiadas. Otro ejemplo son los sistemas que detectan la fatiga de un conductor, monitorizando las expresiones de la cara y los movimientos de cabeza.

Pero en aplicaciones de videovigilancia, el reconocimiento y la identificación facial todavía es una tarea difícil debido a:

- Baja calidad de video: normalmente la grabación se hace en la calle, con una iluminación inadecuada y, además, las personas no son colaborativas, en el sentido de que no ayudan a ser reconocidas. Para solucionarlo, se pueden aplicar técnicas de superresolución.
- Imagen pequeña de la cara: debido a las condiciones de adquisición de la imagen, el tamaño de la cara suele ser más pequeño que los que hay en la base de datos. Eso no sólo hace que la tarea sea más difícil, si no que también afecta a la exactitud del proceso de segmentación, así como a la detección de los puntos más importantes utilizados en el reconocimiento.

Hay aplicaciones en las que da buenos resultados, como en sistemas de control de acceso y en cajeros, donde el video se adquiere en un ambiente relativamente controlado y el tamaño de la cara es aceptable. En estos casos, el reconocimiento basado en video ofrece algunas ventajas respecto al basado en imágenes:

- El video proporciona más información, ya que tenemos más frames para elegir en cual realizar la clasificación.
- El video proporciona continuidad temporal, permitiendo reutilizar la información de clasificación obtenida de los frames de alta calidad para procesar los frames de baja calidad.

- El video permite hacer un seguimiento de las caras, así los cambios de pose y las expresiones faciales se pueden compensar.

6. ASPECTOS TECNICOS

6.1. Debilidades

El sistema de reconocimiento facial no es perfecto y a veces difícil de implementar en ciertas condiciones. Una de las principales debilidades de este sistema es debida a el ángulo en el que se encuentra el rostro que queremos reconocer. Estudios han confirmado que el reconocimiento actúa correctamente hasta los 20 una vez superado este ángulo comienzan a surgir problemas.

Es por este motivo que se está investigando el reconocimiento en 3D con el cual este inconveniente desaparecería. Otro inconveniente es el mal funcionamiento en situaciones de poca luz, además, llevar el pelo largo, gafas de sol o otros objetos que cubran parte del rostro dificulta mucho la tarea. El algoritmo no siempre es capaz de distinguir los rostros si la expresión de este es diferente a la almacenada en la base de datos.

Para solucionar algunos de estos problemas, a parte de hacer el reconocimiento sobre video, también se pueden utilizar técnicas multimodales donde, a parte de la imagen de la persona, se incorpora también información de voz y audio, si esta se encuentra dentro del contenido.

Principalmente, podemos distinguir dos problemas que causan una seria degradación del rendimiento en la mayoría de los sistemas existentes

6.1.1. Problema de iluminación

Los cambios provocados por la iluminación son normalmente más grandes que las diferencias entre personas, causando a los sistemas basados en comparación a equivocarse al clasificar las imágenes de entrada. Se han propuesto algunas soluciones basadas en el conocimiento, en particular teniendo en cuenta que todas las caras pertenecen a una misma clase. Estas técnicas están divididas en 4 tipos:

- Métodos heurísticos: por ejemplo, cuando utilizamos los subespacios de eigenfaces, descartando las componentes principales.
- Métodos de comparación de imágenes: se utilizan representaciones apropiadas de la imagen y medidas de distancia.
- Métodos basados en la clase: utilizan múltiples imágenes de la misma cara en una pose fija pero bajo diferentes condiciones lumínicas.
- Métodos basados en el modelo: utilizan modelos 3D.

6.1.2. Problema del cambio de pose

El rendimiento de un sistema de reconocimiento facial también baja significativamente cuando hay presentes cambios en la pose. Hay diferentes métodos propuestos para solucionarlo:

- Métodos donde la base de datos incluye imágenes de una persona en diferentes poses.
- Métodos híbridos, donde hay disponibles diferentes imágenes por persona durante el entrenamiento, pero sólo una por persona en el reconocimiento. Es la más utilizada.
- Métodos basados en una única imagen, donde no hay entrenamiento. No es popular.

6.1.3. Efectividad

Según un estudio reciente, el reconocimiento facial automático da mejores resultados que en los humanos, pero todavía no hay una técnica que proporcione una solución robusta para todo tipo de situaciones y para las diferentes aplicaciones que lo puedan necesitar. Según la revista 'Science', utilizar imágenes compuestas por varias fotografías adquiridas en diferentes ángulos de una misma persona permite que el rendimiento del algoritmo llegue hasta un 100 en el mejor de los casos.

Este es el resultado de estudios de psicólogos de la Universidad de Glasgow, donde se está tratando el tema y se ha llegado a la conclusión de que este nuevo sistema emula una de las claves de la capacidad humana para reconocer rostros. Aun así continúa sin poder distinguir entre diferentes expresiones faciales.

6.1.4. Privacidad

Aunque esta tecnología nos aporta grandes beneficios, se cuestiona moralmente la privacidad del sujeto. Lo que a simple vista se considera como un avance tecnológico se puede convertir en un arma. Dependiendo de su uso se puede controlar toda una sociedad entera sabiendo en todo momento donde se encuentran y qué están haciendo.

7. EVALUACION

Dada la gran cantidad de teorías y técnicas aplicables a el reconocimiento facial, son necesarias una clara evaluación y una comparativa para estos algoritmos. Para que sean factibles, se han de utilizar grandes cantidades de imágenes para una evaluación adecuada. También es muy importante que la muestra sea estadísticamente lo más similar posible a las imágenes que surgen en la aplicación que se está considerando.

La puntuación ha de realizarse de manera que refleje el coste de los errores de reconocimiento. Recordemos que los resultados dependen mucho de la aplicación que tratamos, así que no podemos extrapolar los resultados para otras aplicaciones.

Existen extensas bases de datos públicas disponibles, así como protocolos de test, para probar las aplicaciones del reconocimiento facial. Uno de ellos es el protocolo FERET, cuyo objetivo es proporcionar un marco de actuación que modele una configuración en tiempo real y reunir una extensa base de datos con imágenes de caras para poder desarrollar algoritmos y evaluarlos. Es el que se utiliza en el FRVT (Face Recognition Vendor Test).

En la siguiente tabla se muestran una serie de bases de datos disponibles actualmente, donde se puede observar el número de elementos utilizados y el tiempo de reconocimiento:

Base de datos	Npersonas	Poses	Iluminacion	Expresiones faciales	Tiempo
AR	116	1	4	4	2
AR	208	1	?	1	12
BANCA	66-10402	1	9-15	6	2
CAS-PEAL	54	1	4	1	2
CMU Hyper	68	13	43	3	1-5
CMU PIE	91	1	3	3	1
Equinox IR	1199	9-20	2	4	1
FERET	10	1	77-84	2	2
Harvard RL	1000	7	16	1	1
KFDB	15	3	3	5	1
MIT	200	3	3	1	1
MPI	300+	1	3	1	1
ND HID	1573	2	1	2	10/13
NIST MID	10	1	?	?	1
ORL	20	?	1	?	?
UMIST	284	?	1	?	1
U. Texas	125	1	?	?	1
U. Oulu	295	?	16	1	1
XM2VTS	116	1	1	?	4
Yale	15	1	3	6	1
Yale B	15	9	64	1	1

Tabla 7.1: datos disponibles actualmente.

8. MEJORAS RECIENTES

A mediados del 2006, en el Face Recognition Grand Challenge (FRGC) se evaluaron los últimos algoritmos de reconocimiento facial. Escanners faciales 3D, imágenes de rostros de alta definición y del iris fueron utilizadas en los tests. Los resultados indicaron que estos nuevos algoritmos tienen 10 veces más exactitud que los algoritmos con fecha anterior a 2002, y 100 veces más que los de 1995. Estos algoritmos son tan precisos que son capaces de reconocer a dos gemelos idénticos.

Otra mejora reciente es que las imágenes de baja resolución actualmente ya no son un problema porque pueden ser tratadas a partir de la superresolución del rostro, conocido como Face Hallucination. Igualmente, la instalación de cámaras de alta definición está en proceso.

9. RECONOCIMIENTO DE VOZ

El reconocimiento automático del habla (RAH) o reconocimiento automático de voz es una disciplina de la inteligencia artificial que tiene como objetivo permitir la comunicación hablada entre seres humanos y computadoras. El problema que se plantea en un sistema de este tipo es el de hacer cooperar un conjunto de informaciones que provienen de diversas fuentes de conocimiento (acústica, fonética, fonológica, léxica, sintáctica, semántica y pragmática), en presencia de ambigüedades, incertidumbres y errores inevitables para llegar a obtener una interpretación aceptable del mensaje acústico recibido.

Un sistema de reconocimiento de voz es una herramienta computacional capaz de procesar la señal de voz emitida por el ser humano y reconocer la información contenida en ésta, convirtiéndola en texto o emitiendo órdenes que actúan sobre un proceso. En su desarrollo intervienen diversas disciplinas, tales como: la fisiología, la acústica, la lingüística, el procesamiento de señales, la inteligencia artificial y la ciencia de la computación.

9.1. Diseño de un sistema de RAH

Un aspecto crucial en el diseño de un sistema de RAH es la elección del tipo de aprendizaje que se utilice para construir las diversas fuentes de conocimiento. Básicamente, existen dos tipos:

Aprendizaje deductivo: Las técnicas de Aprendizaje Deductivo se basan en la transferencia de los conocimientos que un experto humano posee a un sistema informático. Un ejemplo paradigmático de las metodologías que utilizan tales técnicas lo constituyen los Sistemas Basados en el Conocimiento y, en particular, los Sistemas Expertos.

Aprendizaje inductivo: Las técnicas de Aprendizaje Inductivo se basan en que el sistema pueda, automáticamente, conseguir los conocimientos necesarios a partir de ejemplos reales sobre la tarea que se desea modelizar. En este segundo tipo, los ejemplos los constituyen aquellas partes de los sistemas basados en los modelos ocultos de Márkov o en las redes neuronales artificiales que son configuradas automáticamente a partir de muestras de aprendizaje. En la práctica, no existen metodologías que estén basadas únicamente en el Aprendizaje Inductivo, de hecho, se asume un compromiso deductivo-inductivo en el que los aspectos generales se suministran deductivamente y la caracterización de la variabilidad inductivamente.

9.2. Decodificador acústico-fonético

Las fuentes de información acústica, fonética, fonológica y posiblemente léxica, con los correspondientes procedimientos interpretativos, dan lugar a un módulo conocido como decodificador acústico-fonético (o en ocasiones a un decodificador léxico). La entrada al decodificador acústico-fonético es la señal vocal convenientemente representada; para ello, es necesario que ésta sufra un preproceso de parametrización. En esta etapa previa es necesario asumir algún modelo físico, contándose con modelos auditivos y modelos articulatorios.

9.3. Modelo del lenguaje

Las fuentes de conocimiento sintáctico, semántico y pragmático dan lugar al modelo del lenguaje del sistema. Cuando la representación de la Sintaxis y de la Semántica tiende a integrarse, se desarrollan sistemas de RAH de gramática restringida para tareas concretas.

9.4. Características de los sistemas existentes

Los sistemas comerciales han estado disponibles desde 1990. A pesar del aparente éxito de estas tecnologías, muy pocas personas utilizan el sistema del reconocimiento del habla en sus computadoras. Parece ser que muchos de los usuarios utilizan el ratón y el teclado para guardar o redactar documentos, porque les resulta más cómodo y rápido a pesar del hecho de que todos podemos hablar a más velocidad de la que tecleamos. Sin embargo, mediante el uso de ambos, el teclado y el reconocimiento del habla, nuestro trabajo sería mucho más efectivo.

Este sistema donde está siendo más utilizado es en aplicaciones telefónicas: agencias de viajes, atención al cliente, información etc. La mejora de estos sistemas de reconocimiento del habla ha ido aumentando y su eficacia cada vez es mayor.

9.5. Clasificación

Los sistemas de reconocimiento de voz pueden clasificarse según los siguientes criterios:

Entrenabilidad: determina si el sistema necesita un entrenamiento previo antes de empezar a usarse. Dependencia del hablante: determina si el sistema debe entrenarse para cada usuario o es independiente del hablante. Continuidad: determina si el sistema puede reconocer habla continua o el usuario debe hacer pausas entre palabra y palabra. Robustez: determina si el sistema está diseñado para usarse con señales poco ruidosas o, por el contrario, puede funcionar aceptablemente en condiciones ruidosas, ya sea ruido de fondo, ruido procedente del canal o la presencia de voces de otras personas. Tamaño del dominio: determina si el sistema está diseñado para reconocer lenguaje de un dominio reducido (unos cientos de palabras p. e. reservas de vuelos o peticiones de información meteorológica) o extenso (miles de palabras).

9.6. Usos y aplicaciones

Aunque en teoría cualquier tarea en la que se interactúe con un ordenador puede utilizar el reconocimiento de voz, actualmente las siguientes aplicaciones son las más comunes:

Dictado automático: El dictado automático es, hasta hoy, el uso más común de las tecnologías de reconocimiento de voz. En algunos casos, como en el dictado de recetas médicas y diagnósticos o el dictado de textos legales, se usan corpus especiales para incrementar la precisión del sistema.

Control por comandos: Los sistemas de reconocimiento de habla diseñados para dar órdenes a un computador (p.e. "Abrir Firefox", "cerrar ventana") se llaman Control por comandos. Estos sistemas reconocen un vocabulario muy reducido, lo que incrementa su rendimiento. Telefonía: Algunos sistemas PBX permiten a los usuarios ejecutar comandos mediante el habla, en lugar de pulsar tonos. En muchos casos se pide al usuario que diga un número para navegar un menú.

Sistemas portátiles: Los sistemas portátiles de tamaño reducido, como los relojes o los teléfonos móviles, tienen unas restricciones muy concretas de tamaño y forma, así que el habla es una solución natural para introducir datos en estos dispositivos.

Sistemas diseñados para discapacitados: Los sistemas de reconocimiento de voz pueden ser útiles para personas con discapacidades que les impidan teclear con fluidez, así como para personas con problemas auditivos, que pueden usarlos para obtener texto escrito a partir de habla. Esto permitiría, por ejemplo, que los aquejados de sordera pudieran recibir llamadas telefónicas.

Referencias

- [1] S.LI, A.JAIN , *Handbook of Face Recognition*, Springer , 2004.
- [2] GARY BRADSKI y ADRIAN KAEHLER, «Learning OpenCV», *O'Reilly*, **2008**
- [3] FRANCISCO CASACUBERTA NOLLA, «Congreso de la Lengua Española», SEVILLA, **1992**
- [4] [HTTP://WWW.TLDP.ORG/HOWTO/SPEECH-RECOGNITION-HOWTO/](http://www.tldp.org/HOWTO/SPEECH-RECOGNITION-HOWTO/)