

INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE INGENIERÍA MECÁNICA Y ELÉCTRICA
UNIDAD CULHUACAN
INGENIERÍA EN COMUNICACIONES Y ELECTRÓNICA

SEMINARIO DE TITULACIÓN
“PROCESAMIENTO DIGITAL DE SEÑALES”

RECONOCIMIENTO DE VOZ PARA UN
CONTROL DE ACCESO
MEDIANTE RED NEURONAL DE RETROPROPAGACIÓN

T E S I N A

Que para obtener el grado de:

INGENIERO EN
COMUNICACIONES Y ELECTRÓNICA

Presenta:

Sara Eunice Rascón Montiel

ASESORES:

M. en C. ORLANDO BELTRÁN NAVARRO.
M. en C. BRAULIO SANCHEZ ZAMORA



México, D. F., Noviembre de 2009.

**INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE INGENIERÍA MECÁNICA ELÉCTRICA
UNIDAD CULHUACAN**

TESINA

Que para obtener el título de: **INGENIERO EN COMUNICACIONES Y
ELECTRÓNICA**

Por la opción de titulación: **SEMINARIO DE TITULACIÓN
“PROCESAMIENTO DIGITAL DE SEÑALES”**

Deberá desarrollar: **SARA EUNICE RASCÓN MONTIEL**

**“RECONOCIMIENTO DE VOZ PARA UN CONTROL DE ACCESO MEDIANTE
RED NEURONAL DE RETROPROPAGACIÓN”**

El trabajo presentado es un sistema que utiliza diversas herramientas y consta de varias etapas, tales como: captura y grabación de muestras de voz, entrenamiento de una red neuronal artificial de retropropagación, y captura y evaluación de la misma red para lograr el reconocimiento de la voz humana, y enviar una señal a un circuito electrónico que nos permita operar una cerradura eléctrica, y de esta manera mantener controlado un acceso, logrando el objetivo principal de mantener la seguridad física en los centros de trabajo.

CAPÍTULO I. MARCO TEÓRICO

I.1. CONTROL DE ACCESO

I.2. LA VOZ

I.3 RECONOCIMIENTO DE VOZ

I.4 REDES NEURONALES

CAPÍTULO II. DESARROLLO

CAPÍTULO III. RESULTADOS

CONCLUSIONES

ANEXOS

BIBLIOGRAFÍA

M. en C. Orlando Beltrán Navarro
Coordinador del seminario

M. en C. Braulio Sánchez Zamora
Asesor

Ing. Ignacio Monroy Ostría
Jefe de carrera de ICE

Señor, digno eres de recibir la gloria y la honra y el poder; porque tú creaste todas las cosas, y por tu voluntad existen y fueron creadas.

Apoc. 4:11

TEMA

RECONOCIMIENTO DE VOZ PARA UN CONTROL DE ACCESO MEDIANTE RED NEURONAL DE RETROPROPAGACIÓN

OBJETIVO

Presentar una aplicación de la tecnología actual para mantener la seguridad física en los centros de trabajo.

ÍNDICE

	Página.
Introducción	3
Capítulo I. Marco teórico	4
I.1. Control de Acceso	4
I.2. La Voz	11
I.3 Reconocimiento de Voz	21
I.4 Redes Neuronales	26
Capítulo II. Desarrollo	41
Capítulo III. Resultados	53
Conclusiones	57
Anexos	58
Bibliografía	63

INTRODUCCIÓN

Hoy en día es imprescindible contar con sistemas de seguridad para controlar los accesos a sitios donde se resguardan objetos o información valiosa, o simplemente en donde se requiere contar con registros de accesos del personal a un centro de trabajo.

El presente trabajo consiste en la utilización de herramientas computacionales para lograr el reconocimiento de la voz humana y ejecutar una acción específica que en este caso será mantener controlado un acceso.

Para comprender de manera general el trabajo propuesto, en el marco teórico se abordan los conceptos principales de cada una de las herramientas utilizadas para lograr la aplicación.

En primera instancia, se presenta el origen de la necesidad de mantener la seguridad en un lugar o tratar de resguardar un bien, desde la utilización de guardias hasta los sistemas biométricos más complejos, tales como el reconocimiento de la voz.

Entrando un poco en el tema central de este trabajo, el reconocimiento de la voz, se menciona que un sistema de este tipo es 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, el procesamiento de señales, la inteligencia artificial y la ciencia de la computación.

Ahora bien, para lograr el reconocimiento de la voz, la herramienta computacional utilizada es la red neuronal artificial de retropropagación programada en un simulador, que consta de varias etapas: captura y grabación de la voz, entrenamiento de la red, captura y evaluación de la red.

Una vez evaluada y en su caso, si es correcta la identificación de la voz, se obtiene una señal de salida, la cual se envía a un circuito electrónico para permitir la apertura de una cerradura eléctrica, siendo ésta la última etapa del sistema propuesto.

CAPÍTULO I

MARCO TEÓRICO

I.1 CONTROL DE ACCESO

El control de accesos consiste básicamente en la capacidad de controlar y conocer quién y cuándo gana acceso, sea a una zona de un edificio, a cierta información, a un mensaje o a un servicio. El control de acceso es posiblemente la medida de seguridad más importante que se puede establecer, dado que un control de accesos bien aplicado nos permite saber quién ha hecho qué, dónde y cuándo, y controlar quién puede hacer qué, dónde y cuándo.

Es difícil reconocer a varios cientos de otras personas; para distinguir a los amigos de los enemigos se inventaron los signos distintivos y los uniformes.

Hay pinturas rupestres que muestran a guerreros formando un frente y marchando en columna siguiendo a un líder que vestía un uniforme llamativo para distinguirse de sus tropas.

En una fortificación se agudizaba el problema de reconocer a los que deberían poder entrar o salir. No bastaba un uniforme, pues un enemigo podría hacerse de uno, y frecuentemente se colaboraba con aliados y aparecía una variedad de uniformes que daban derecho a entrar o salir.

Así se desarrolló la metodología de centinelas y "santo y seña" para otorgar acceso a una fortificación. Esta metodología persiste hoy en día en los manuales militares.

R. W. Gregory (2000) explica el *Rondín* durante el cual el Oficial de la Guardia, el Comandante, o el Oficial del día pueden, cuando quieran, acompañados por un sargento y una escolta de dos hombres, recorrer la instalación. Al acercarse a un centinela éste dirá:

Alto. ¿Quién vive?
La respuesta adecuada es:
Rondín.
A lo cual el centinela ordena:
Alto Rondín. Que avance el sargento con la contraseña.
Si el sargento desempeña bien esta acción el centinela dice:
La contraseña es la correcta.
Y el Oficial ordenará entonces:
Prosiga el rondín.

Este tipo de procedimientos datan de la prehistoria, y van de la mano con las fortificaciones, ya sean permanentes o de campo. De ellos, y de su estricto cumplimiento depende la seguridad de las tropas y el equipo resguardado.

La combinación del establecimiento de perímetros y de mecanismos de autenticación, como el de santo y seña, ha dado origen a lo que hoy se conoce como Control de Acceso, que es la actividad principal en la seguridad física y en la seguridad informática.

El control de acceso es uno de los principales aspectos que involucran la seguridad física en edificios o sitios, además de los sistemas de protección contra incendios, suministro de energía, aire acondicionado, detección de agua, telecomunicaciones, etc.

Un sistema acotado en sus componentes y sus usuarios es aquel en que sea posible hacer una lista de los unos y los otros y que puede emplear un sistema que permita a cada componente saber en forma inequívoca con quien esta interaccionando en un momento dado.

Este conocimiento es la base del funcionamiento seguro del sistema. En cada interacción es posible identificar las partes, garantizar su identidad lo que permite que la confidencialidad, autenticidad y la integridad

Los pasos indispensables para implementar un sistema de control de acceso de cualquier tipo son:

- Registro
- Identificación
- Autenticación

REGISTRO

Cualquier sistema que pretenda hacer uso de mecanismos de identificación y autenticación requiere de un procedimiento de registro de los componentes del mismo. En otras palabras hay que levantar un inventario de componentes y/o de usuarios.

El sistema se considera acotado a los elementos que consten en este inventario; aquellos elementos que no aparezcan en este inventario, por definición, no forman parte del sistema.

Claro que este inventario se modifica a lo largo del tiempo, estando sujeto a procesos de altas, bajas y cambios.

IDENTIFICACIÓN

Durante el registro en cualquier sistema se anota un identificador del usuario.

Este puede ser su nombre, un apodo, un número (como en las cuentas numeradas de ciertos bancos, o el numero de pasaporte) o información que distinga a un usuario de los demás.

Mediante este identificador se localiza en el registro el renglón correspondiente al usuario.

Un ejemplo claro es el pasaporte, que es un documento mediante el cual el país de origen certifica la identidad de quien lo porta como ciudadano.

Al pedir un pasaporte el ciudadano presenta pruebas de su ciudadanía, que son verificadas por el gobierno. Si esta verificación es exitosa, el gobierno emite un documento de identidad que es aceptado por otros países por tratado internacional.

Como apoyo a este documento de identidad se incluyen en él tres elementos que son la fotografía, la firma y la huella digital del dueño del pasaporte. Éstos son susceptibles de ser verificados en cualquier frontera por personal de migración especializado.

AUTENTICACIÓN

Una credencial se utiliza para verificar la identidad en un momento dado de un usuario. Esta verificación de identidad se conoce como autenticación. Mediante la autenticación se facilita la comprobación automatizada de la identidad de un usuario interno o externo siempre que sea necesario.

Hay al menos tres tipos de autenticación que pueden combinarse: por algo que eres, como puede ser la biometría, por algo que tienes, como es la posesión de una tarjeta, y por algo que sabes, como es el caso de una contraseña.

Existe un número muy amplio de medios de autenticación más específicos, como:

- Contraseñas
- Frases de paso (es como una contraseña, pero más larga)
- Preguntas mágicas (es una pregunta personal en la que sólo nosotros conocemos la respuesta)
- Tarjetas “inteligentes” (tarjetas con microchip que almacenan cierta cantidad de datos)
- Biometría (de la voz, iris, letra, rostro, huellas digitales o mecanografía, característicos de cada persona)

Algunos de estos medios exigen lectores especiales que los hacen muy caros. Los medios biométricos suelen tener problemas de falsos negativos o falsos positivos por variaciones naturales.

Al registrarse el usuario debe depositar o recibir un autenticador, que es un dato que se relacione con el usuario.

La posesión de este dato se considera como evidencia incontrovertible de que el que lo exhibe es quien aparece en el registro.

La imposibilidad de exhibir el autenticador inmediatamente lleva a la denegación del uso del servicio que se solicita.

El proceso de autenticación consiste en comparar las características que exhibe el usuario con aquella que aparece en renglón correspondiente en el sistema de registro.

Históricamente la autenticación de usuarios se hizo mediante contraseñas que se registran en un archivo que además contiene los datos de identificación de cada usuario y otra información relevante.

Actualmente se emplea una gama de mecanismos de autenticación que se clasifican en cuatro grupos:

- algo que el usuario conoce
- algo que el usuario tiene
- algo que caracteriza al usuario
- algo que determina su posición sobre la Tierra.

Lo más seguro es emplear dos autenticadores que pertenezcan a dos grupos distintos.

AUTENTICACIÓN BASADA EN CONOCIMIENTOS.

Se basa en un secreto compartido por el usuario y el sistema de información.

Si alguien ajeno al sistema llega a conocer el secreto compartido puede suplantar al usuario y dejar de ser ajeno.

Si el suplantador logra enterarse del secreto porque el usuario no lo ocultó, no hay defensa posible.

AUTENTICACIÓN BASADA EN CARACTERÍSTICAS FÍSICAS (AUTENTICACIÓN BIOMÉTRICA)

La tecnología informática hoy en día permite el uso de características del cuerpo de una persona para autenticarla. Se usan características que no se alteren fácilmente y que sean distintas de una persona a otra.

Deben también ser expresables matemáticamente en forma sintética y además su visibilidad debe ser compatible con el atuendo normal de las personas.

La biometría es el uso automatizado de características fisiológicas o conductuales de una persona para determinar o verificar su identidad.

La biometría fisiológica está basada en medidas o datos de partes del cuerpo humano.

La biometría conductual se basa en la medida o datos de acciones de una persona, e indirectamente en sus características físicas características se depositan en el momento del registro, y su descripción matemática se guarda en la lista de usuarios.

Al solicitar acceso el usuario exhibe la característica que haya registrado, el sistema recalcula la descripción matemática y la compara con la que se encuentra almacenada.

El proceso de verificar la identidad de una persona biométricamente, o sea autenticarla, es mucho más directo que el proceso de identificarla.

En la autenticación se conoce la identidad de la persona (mediante una credencial o una contraseña) y solo es necesario consultar una tabla que contenga la característica biométrica para compararla con la que se capta.

En el proceso de identificación hay que buscar la característica captada entre las de todos los sujetos registrados.

Básicamente los pasos que deben llevarse a cabo para autenticar a un usuario son los siguientes:

- Se captura la característica biométrica en el momento de registro del usuario
- Se procesa la información y se da de alta el descriptor de la biometría
- Se almacena el descriptor en un acervo local, central o portátil
- Se captura la característica biométrica en el lugar de autenticación
- Se procesa la información y se extrae el descriptor de la biometría
- Se compara este descriptor con el que se ha almacenado
- Se registra el resultado en una bitácora confiable

Identificadores son los datos que caracterizan a una instancia del sistema biométrico. Se crean después de que el sistema recibe la información amplia (imágenes, grabaciones) y las analiza matemáticamente.

Los identificadores se crean durante el registro del usuario, y se almacenan para usarse posteriormente. Se generan más cuidadosamente y casi siempre después de obtener muestras repetidas de la característica biométrica.

Los identificadores no interoperan. Los que genera y utiliza un sistema no son compatibles con otros sistemas.

BIOMETRÍA FISIOLÓGICA

Entre las técnicas de identificación de usuarios mediante biometría fisiológica se pueden mencionar las siguientes:

- Verificación de huellas dactilares
- Geometría de la mano
- Análisis de la retina
- Análisis del iris
- Reconocimiento de rostros
- Verificación de voz

Al hablar se producen ondas sonoras, que son ondas de compresión del aire que nos rodea. La forma de estas ondas depende de la estructura del cuello, la boca y la nariz de la persona que habla, pues las vibraciones producidas por las cuerdas bucales pasan a través de estos órganos. Es de esta manera como formulamos los fonemas que usamos para enunciar palabras.

Se usan los mismos principios que emplean los astrónomos para analizar la composición química de las estrellas, o los que se usan en laboratorios forenses para distinguir residuos de objetos. Para analizar la voz se analiza el sonido, en vez de analizarse la luz.

Esta combinación de características físicas de cada individuo produce sonidos y palabras cuyo análisis espectroscópico es suficientemente preciso para distinguir el habla de dos personas.

A diferencia de las huellas dactilares que solo ofrecen diez registros por persona como máximo, en el caso de la voz se pueden registrar muchas más instancias.

Como la precisión inherente de este tipo de métodos se acumula multiplicativamente, la precisión de la identificación por voz puede ser mucho mayor que la de las huellas dactilares.

En el caso de la voz unas cuantas palabras contienen decenas de fonemas, cada uno de los cuales se emplea para contribuir a la identificación.

La mayor parte de las comunicaciones de negocios y personales se llevan a cabo verbalmente por teléfono o presencialmente, y es posible entonces verificar la identidad de quien habla.

Estos sistemas se usan ampliamente para controlar el acceso a sistemas de llamadas de larga distancia.

Permite combinar el método de autenticación por contraseñas verbales con la verificación biométrica de voz.

BIOMETRÍA CONDUCTUAL

Entre las técnicas de identificación de usuarios mediante biometría conductual se pueden mencionar las siguientes:

- Verificación de firmas
- Verificación de la mecanografía

Aplicación al control de acceso físico.

Las organizaciones grandes dispersas geográficamente, se enfrentan al problema de permitir el acceso a cada instalación según una lista de control de acceso.

Es típico de las organizaciones grandes que la rotación de personal sea frecuente, y que se hagan contrataciones rutinariamente.

Esto implica modificar listas de control de acceso dispersas (en instalaciones en todo el mundo).

El procedimiento resulta complejo y a veces imposible pero si se establece una jerarquía de confidencialidad y se le asigna una clase a cada persona, empleado permanente, temporal o visitante, y se establece también una jerarquía de confidencialidad de los sitios (edificios, pisos de edificios, plantas, etc.) el proceso resulta sencillo.

El control de acceso es efectivo si:

- Se tiene un sistema de identificación y autenticación de los usuarios para que uno de ellos no pueda adquirir los derechos de otro
- La información acerca de las características de los objetos y sujetos está protegida en contra de modificaciones no autorizadas.

Cabe señalar que debe haber para el efecto una interrelación entre cada área que responsabiliza a un individuo a través del uso de una clave, y el área de control que asigna y controla las claves emitidas a los usuarios. El no conocer sobre el despido, renuncia o ausencia de un personal determinado durante un tiempo específico o permanente provocaría que el proceso de control de acceso no tenga éxito por daños provocados por actos de sabotaje, robos, asaltos, etc.

I.2 LA VOZ

De entre la amplia gama de sonidos con finalidad comunicativa podemos distinguir los sonidos del habla y estudiarlos por separado, para analizar sus características acústicas y su función. Por lo tanto, consideramos sonido del habla cualquier unidad sonora producida por los órganos de fonación en tanto ésta tenga finalidad comunicativa y sea posible su diferenciación.

La generación del habla por parte de las personas, consiste en la creación de una onda de presión sonora que se propaga a través del aire a una velocidad de unos 340 metros por segundo. La concatenación de unos determinados sonidos y en un orden prefijado, característicos de cada idioma, constituyen el mensaje.

FONÉTICA

La fonética es una rama científica que estudia los sonidos del lenguaje (entre otros) desde un punto de vista articulatorio y acústico; se propone determinar las características físicas de la fonación, y la ciencia en la que se inserta es la física.

EL APARATO FONADOR

Se conoce como aparato fonador al conjunto de órganos que intervienen en la producción de un sonido articulado; aunque tales órganos no tienen como finalidad fisiológica primaria la función fonética, pues en realidad no podemos hablar de la existencia “natural” de un aparato fonador, es indiscutible que el hombre ha “adaptado” los órganos de la fonación para hablar.

Algunos órganos intervienen apoyando indirectamente tanto el aprendizaje como la realización cotidiana del habla, tales como el sistema nervioso, la vista y el oído.

Los órganos que constituyen el conjunto total del aparato fonador se puede clasificar en tres grupos bien definidos:

- Cavidades infraglóticas o aparato respiratorio
- Cavidad laríngea
- Cavidades supraglóticas

Ver la Figura I.2.1 para mayores detalles de la partes que constituyen el aparato fonador.

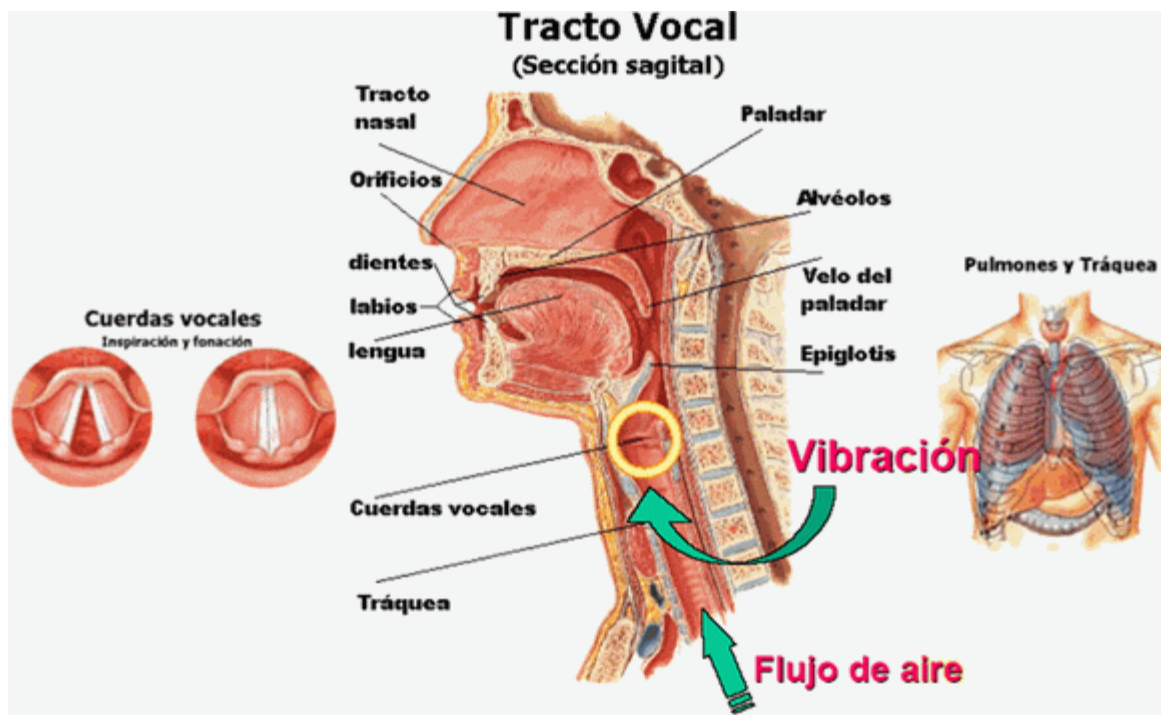


Figura I.2.1. Aparato fonador humano

El sonido se produce mediante la expulsión del aire pulmonar a través de las cuerdas vocales; si al paso del aire las cuerdas se aproximan y vibran, el sonido que se articula es sonoro, y si al aproximarse las cuerdas no vibran al paso del aire, el sonido articulado es sordo.

Entre los sonidos articulados sonoros se pueden establecer dos grupos:

- El de las tradicionalmente llamadas vocales, en cuya fonación las cuerdas se encuentran más tensas, los golpes vibratorios son más fuertes, la frecuencia es mayor y la glotis está más abierta.
- El de las consonantes sonoras producidas con una menor tensión de las cuerdas, menor fuerza de los golpes vibratorios, la frecuencia y la abertura de la glotis también menores.

La vibración de las cuerdas vocales genera la formación de una onda sonora conocida como tono fundamental. A partir del tono fundamental se producen una serie de armónicos que, acompañados al tono fundamental constituyen el timbre de la voz.

Otro aspecto sonoro que se origina en la cavidad laríngea es la intensidad que corresponde a la mayor o menor fuerza del aire al pasar por las cuerdas vocales.

De los tres aspectos del sonido originados en la laringe: tono, timbre e intensidad, los dos últimos se filtran y modifican sustancialmente en las cavidades supraglóticas.

Un cuarto aspecto sonoro es la duración, y corresponde a la cantidad en la emisión de un sonido, ya sea por interrupción en la corriente de aire o por la modificación del curso de ésta para hacerle rebotar o conducirla de una manera distinta.

EL FONEMA

El fonema es la abstracción de un sonido del habla.

Los fonemas son unidades mínimas de articulación y constituyen los elementos de la descripción fonológica; se distinguen esencialmente de los demás sonidos del habla en que, a pesar de no poseer una carga de significado propio, tienen la capacidad de distinguir parejas de palabras en un contexto; por ejemplo, en el caso de la pareja /peso/ /beso/, el elemento que varía en la pareja son los fonemas /p/ /b/.

La ciencia que estudia los fonemas se llama Fonología y la que estudia los sonidos, Fonética.

Son siete los rasgos distintivos que nos permiten clasificar los fonemas del español:

1. Oral
2. Nasal
3. Sonoro
4. Sordo
5. Labial
6. Velar
7. Dental o alveolar

FONEMAS VOCÁLICOS

Se consideran fonemas vocálicos o vocales aquellos que se realizan por la libre vibración de las cuerdas vocales al paso de la corriente de aire pulmonar; el chorro de aire sonoro no encuentra más oposición en las cavidades bucal y nasal que la indispensable para modificar su dirección, de modo que la corriente rebote, para finalmente salir por la boca.

Podemos clasificar las vocales de acuerdo con las tres variables articulatorias:

- El grado de abertura de la boca
- La posición de la lengua
- El grado de sonoridad

A partir de cada una de estas variables se puede esquematizar la clasificación de las vocales como sigue:

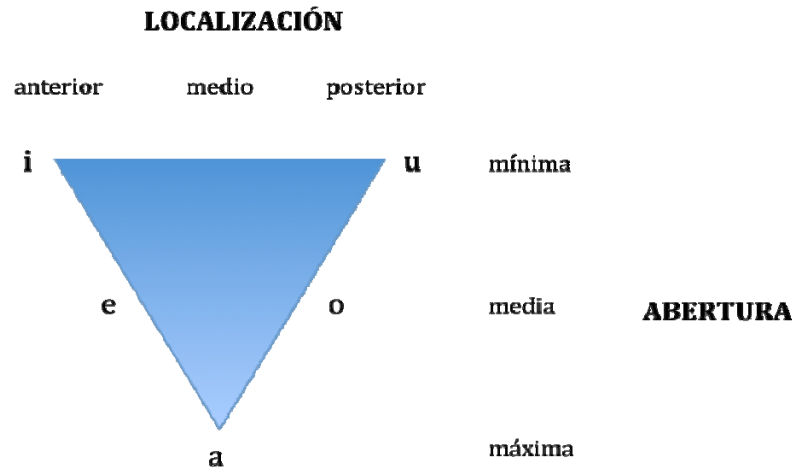


Figura I.2.2. Triángulo vocálico del español (HELWAG)

FONEMAS CONSONÁNTICOS

En la articulación de los sonidos consonánticos siempre hay un obstáculo más o menos grande que impide salir el aire desde los pulmones al exterior. Según las circunstancias que rodean esta salida del aire, existen ciertos factores que debemos tener en cuenta a la hora de clasificarlos:

- Zona o punto de articulación. Es el lugar donde toman contacto los órganos que intervienen en la producción del sonido. Por ejemplo, si para producir un sonido entran en contacto los dos labios, se crearán sonidos bilabiales como es el caso de las realizaciones de los fonemas /p/, /b/ y /m/.
- Modo de articulación. Es la postura que adoptan los órganos que producen los sonidos. Por ejemplo, si los órganos cierran total y momentáneamente la salida del aire, los sonidos serán oclusivos. Ese es el caso de los sonidos /p/, /t/ y /k/.
- Actividad de las cuerdas vocales. Cuando producimos sonidos, las cuerdas vocales pueden vibrar o no vibrar. Si las cuerdas vocales no vibran, los sonidos se llaman sordos. Así producimos /p/, /f/, /k/. Cuando, por el contrario, las cuerdas vocales vibran se llaman sonoros. Esto pasa al pronunciar /a/, /b/, /d/.
- Actividad de la cavidad nasal. Si al producir sonidos, parte del aire pasa por la cavidad nasal, los sonidos se llaman nasales. Son de esta clase /m/, /n/, /ñ/. Si todo el aire pasa por la cavidad bucal se llaman orales. De este tipo son /f/, /e/, /s/.

Rasgo	Órganos	Ejemplos
Bilabial	Los dos labios.	/p/, /b/, /m/
Labiodental	Labio inferior y dientes superiores.	/f/
Interdental	Lengua entre los dientes.	/z/
Dental	Lengua detrás de los dientes superiores.	/t/, /d/
Alveolar	Lengua sobre la raíz de los dientes superiores.	/s/, /l/, /r/, /rr/, /n/
Palatal	Lengua y paladar.	/ch/, /y/, /ll/, /ñ/
Velar	Lengua y velo del paladar.	/k/, /g/, /j/

Tabla I.2.1. Rasgos motivados por el punto de articulación

Rasgo	Órganos	Ejemplos
Oclusivo	Cierre total y momentáneo del paso del aire.	/p/, /b/, /t/, /d/, /k/, /g/, /n/, /m/
Fricativo	Estrechamiento por donde pasa el aire rozando.	/f/, /z/, /j/, /s/
Africado	Se produce una oclusión y después una fricación.	/ch/, /ñ/
Lateral	El aire pasa rozando los lados de la cavidad bucal.	/l/, /ll/
Vibrante	El aire hace vibrar la punta de la lengua al pasar.	/r/, /rr/

Tabla I.2.2. Rasgos motivados por el modo de articulación

Rasgo	Órganos	Ejemplos
Sordo	No vibran las cuerdas vocales.	/p/, /t/, /k/, /ch/, /z/, /s/, /j/, /f/
Sonoro	Vibran las cuerdas vocales.	/b/, /z/, /d/, /l/, /r/, /rr/, /m/, /n/, /ll/, /y/, /g/

Tabla I.2.3. Rasgos motivados por la intervención de las cuerdas vocales

Rasgo	Órganos	Ejemplos
Nasal	Parte del aire pasa por la cavidad nasal.	/m/, /n/, /ñ/
Oral	Todo el aire pasa por la boca.	El resto

Tabla I.2.4. Rasgos motivados por la intervención de la cavidad nasal

LA VOZ COMO SEÑAL

Mediante un micrófono es posible captar la onda de presión sonora emitida por un orador y convertirla en una señal eléctrica. En un osciloscopio podría obtenerse una imagen semejante a la mostrada en la Figura I.2.3 al pronunciar la frase “el golpe de timón fue sobrecogedor”.

Se trata de una representación de la amplitud de la señal de voz en función del tiempo, que suele denominarse oscilograma.

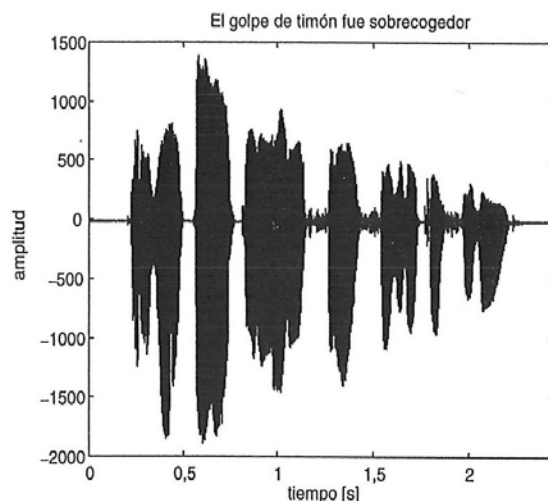


Figura I.2.3. Oscilograma de una frase

En la Figura I.2.3 se observa la existencia de zonas con distintas amplitudes. Si colocamos suavemente los dedos sobre nuestra “manzana de Adán”, notaremos que mientras hablamos existen sonidos en los cuales nuestras cuerdas vocales vibran y otros en los que permanecen en reposo, tal como se revisó anteriormente:

- Sonidos sonoros: Las cuerdas vocales vibran y el aire pasa a través del tracto vocal sin impedimentos importantes.
- Sonidos sordos: Las cuerdas vocales no vibran y existen restricciones importantes al paso del aire que proviene de los pulmones, por lo que son de amplitud menor y normalmente de naturaleza más ruidosa que los sonoros.

Para apreciar esta diferencia entre ambos tipos de sonidos se muestra una ampliación de una pequeña parte de la frase de la Figura I.2.4 correspondiente a los fonemas /f/ (sordo) y /u/ (sonoro) extraídos de la palabra “fue”. En este oscilograma se muestra el carácter ruidoso de la /f/ y periódico de la /u/, así como sus diferentes amplitudes.

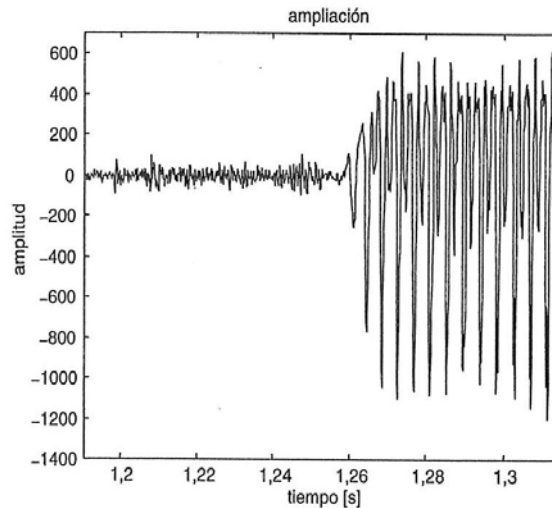
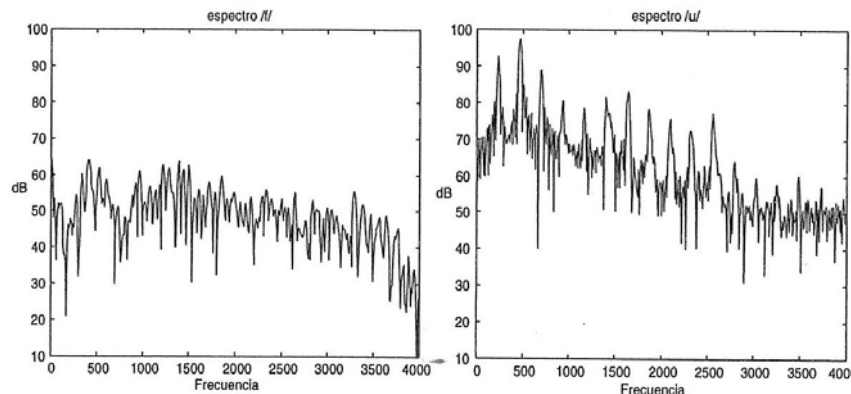


Figura I.2.4. Ampliación correspondiente a /fu/

La representación en el dominio transformado mediante la FFT (transformada rápida de Fourier) supone una visualización alternativa. Las figuras I.2.5 muestran el resultado de aplicar el logaritmo al módulo de dicha transformación, para cada uno de los fonemas.



Figuras I.2.5. Espectrograma de /f/ y espectrograma de /u/

Estas representaciones reciben el nombre de espectrogramas, o representaciones de la amplitud en función de la frecuencia. La señal de voz es limitada en banda, a unos 8 kHz. Sin embargo, la mayor parte de la información se encuentra en los primeros 4 kHz, que es aproximadamente el ancho de banda utilizado en las comunicaciones por vía telefónica.

Es posible observar que la periodicidad del fonema /u/ correspondiente a la segunda mitad de la Figura I.2.4, se aprecia en el espectrograma mediante la existencia de una frecuencia fundamental y una serie de armónicos equiespaciados una misma frecuencia. Esta frecuencia recibe el nombre de pitch (tono), y está directamente relacionada con la frecuencia de vibración principal de las cuerdas vocales. El margen habitual para locutores masculinos adultos del valor del pitch es de 50 a 250 Hz (periodo de 20 a 4 ms o 160 a 32 muestras al

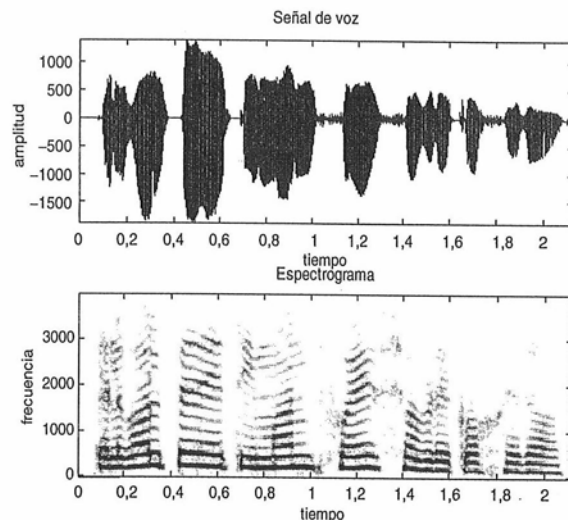
trabajar con frecuencias de muestreo $F_m=8$ kHz), mientras que para locutoras se encuentra entre 120 y 500 Hz (periodo de 8.3 a 2 ms o 66 a 16 muestras con $F_m=8$ kHz).

Otro aspecto destacable de los espectrogramas correspondientes a fonemas sonoros es la presencia de zonas enfatizadas, o resonancias, y zonas desenfáticas, o antirresonancias. Estas zonas se deben a la configuración física de las diferentes cavidades del tracto configurado por el sistema articulador. La posición de las frecuencias de resonancia depende de la forma, tamaño y características físicas del tracto vocal al producir cada sonido. De este modo, cada forma del tracto vocal puede modelarse o formarse por un filtro con un conjunto de frecuencias de resonancia que lo caracterizan. Por ello a cada una de estas frecuencias de resonancia se le denomina “frecuencia formante” o simplemente formante. Los distintos formantes se denominan F_1 , F_2 , F_3 , F_4 , etc., indicando las frecuencias de resonancia ordenadas en sentido ascendente. Habitualmente se utilizan entre 3 y 5 formantes al considerar como frecuencia de muestreo $F_m=8$ kHz. Normalmente las señales sonoras presentan una característica típica de paso bajo. Por ello los formantes más significativos son los primeros. En castellano existen cinco sonidos vocálicos cuyos primeros formantes (según Antonio Quilis) son aproximadamente los mostrados en la Tabla I.2.5.

Vocal/Formante (Hz)	F_1	F_2
/i/	284	2430
/e/	527	2025
/a/	689	1458
/o/	608	1215
/u/	243	770

Tabla I.2.5. Valores frecuenciales de los primeros formantes de las vocales castellanas.

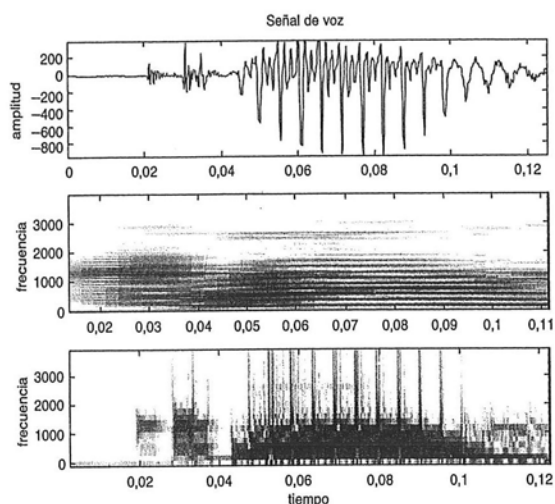
Dado que el tracto vocal evoluciona en el tiempo para producir los distintos sonidos, la caracterización espectral de la señal de voz también es variante en el tiempo. Esta evolución temporal puede representarse mediante un espectrograma de la señal de voz o sonograma. Éste es una representación bidimensional que muestra la evolución temporal de la caracterización espectral. En la Figura I.2.6 se representan el oscilograma y el sonograma de la frase: “el golpe de timón fue sobrecogedor”.



Figuras I.2.6. Oscilograma y sonograma de la señal correspondiente a la frase “el golpe del timón fue sobrecogedor”, pronunciada por una locutora.

Los formantes aparecen como franjas horizontales, mientras que los valores de amplitud en función de la frecuencia se representan en tonalidades de grises en sentido vertical.

Existen dos tipos de sonogramas: de banda ancha y de banda estrecha, en función del ancho de banda del filtro que se haya utilizado para realizar el análisis frecuencial. En el caso de los espectrogramas de banda ancha, se obtendrá una buena resolución temporal. En cambio en el caso de espectrogramas de banda estrecha se obtendrá una buena resolución frecuencial, dado que el filtro utilizado es de banda estrecha, y permite obtener estimaciones espectrales más precisas. La figura I.2.7 muestra un fragmento de señal y su correspondiente espectrograma de banda ancha (obtenido con ventanas de 40 muestras). Obsérvese que en el de banda estrecha se diferencian claramente los formantes, mientras que en el de banda ancha se observan los detalles temporales con mayor claridad de pitch).



Figuras I.2.7. Oscilograma y espectrograma de banda estrecha y banda ancha

Los sonidos de la voz se pueden representar mediante un espectrograma, que indica las componentes frecuenciales de la señal de voz.

El espectro nos proporciona información acerca de los parámetros del modelo de producción de voz, tanto de la excitación (tono) como del filtro que representa el tracto vocal (formantes)

DIGITALIZACIÓN DE LA SEÑAL

El procesado digital de señal mediante un DSP, ordenador, etc., requiere previamente la conversión de la señal acústica a eléctrica mediante un micrófono, y la conversión de la señal analógica resultante a señal digital. Por otra parte, para restaurar o generar señal audible a partir de un sistema digital, será necesaria la conversión digital a analógica, su amplificación y su radiación mediante altavoz.

Para realizar la conversión de analógico a digital es necesario realizar un muestreo o discretización de los valores de la señal cada cierto intervalo de tiempo, denominado periodo de muestreo T_m cuyo inverso es la frecuencia de muestreo f_m . En este punto hay que tener en cuenta que el teorema de Nyquist establece que, para evitar fenómenos de aliasing en señales de paso bajo, es necesario muestrear como mínimo al doble del ancho de banda de la señal. Esto supone incorporar antes del muestreador un filtro de paso bajo antialiasing que limite al ancho de banda de la señal analógica, y elimine el ruido presente más allá del ancho de banda de la señal útil.

Una vez realizada la discretización en el tiempo, es necesario discretizar los valores de amplitud mediante un convertidor analógico-digital, que deberá ser de un número suficientemente elevado de bits. Por ejemplo, en telefonía digital se utilizan 8 bits, en audio con calidad de compact disc 16 bits, para mezclas digitales de CD 20 bits, etc. La secuencia de muestras resultante recibe el nombre de muestras PCM.

Las siguientes figuras resumen los pasos necesarios para realizar la conversión Analógica/Digital (A/D) y Digital/Analógica (D/A) respectivamente.



Figura I.2.8. Conversión A/D



Figura I.2.9. Conversión D/A

I.3 RECONOCIMIENTO DE VOZ

Sin duda alguna, la voz es la forma más natural y eficiente de comunicación entre los seres humanos. Sin embargo, cada vez son más frecuentes las situaciones en las que la comunicación se establece con una máquina.

El reconocimiento de voz permite la comunicación entre los seres humanos y las computadoras.

En la producción del habla intervienen elementos como la semántica, lingüística, articulación y la acústica. Todos ellos condicionan las propiedades acústicas de la señal de la voz. Las diferencias debidas intrínsecamente al locutor son el resultado de una combinación de diferencias anatómicas del tracto vocal y de la forma de hablar adquirida por cada persona. En el reconocimiento de locutor se utilizan estas diferencias para discriminar entre locutores.

APLICACIONES

Si bien algunas aplicaciones prácticas pueden requerir reconocimiento del habla y del locutor simultáneamente, en principio se pueden dividir en dos grandes grupos: las del reconocimiento del habla y las del reconocimiento de locutor. Otro grupo de menor importancia estaría formado por otros temas, como reconocimiento del idioma de una determinada frase hablada, variantes dialectales, etc.

Ahora, entre las principales aplicaciones del reconocimiento del locutor cabe destacar:

- Acceso a recintos o informaciones: mediante la voz es posible identificar a las personas y permitirles acceder o no a fábricas, hogares, informaciones bancarias privadas o de determinados grupos de trabajo, etc.
- Reconocimiento de sospechosos: a partir de grabaciones policiales, es posible identificar a qué individuo pertenece la voz de la grabación.
- Transcripción automática de reuniones: en reuniones con diversos oradores interesa disponer de la transcripción de las frases habladas (reconocimiento del habla), así como de las identidades de las personas que han realizado las alocuciones.

Existen dos grandes grupos de aplicaciones dentro del reconocimiento del locutor:

a) Identificación

Consiste en determinar la identidad de un locutor mediante una máquina (ordenador). Para ello, la voz de entrada debe ser conocida previamente por el ordenador, de la misma forma que sucede entre las personas. El conocimiento de las diferentes voces se adquirirá mediante un proceso de entrenamiento, en el cual se recogerá un conjunto suficientemente amplio de

frases habladas de cada uno de los locutores que formarán parte de una base de locutores conocidos por el sistema. Posteriormente, en el proceso de identificación de un locutor de entrada, se realizará un test consistente en comparar los datos de entrada con todos los existentes en la memoria del sistema. A partir de la comparación entre los datos de entrada y los datos almacenados en memoria durante el proceso de entrenamiento, se decidirá la identidad del locutor de entrada.

b) Verificación

En este caso, el locutor de entrada suministra su identidad, y el sistema deberá comprobar los datos de entrada con los almacenados en memoria de ese mismo locutor, para comprobar si es quien dice ser. Para prevenir suplantaciones de personalidad mediante grabaciones, será necesario realizar un reconocimiento dependiente del texto, que será distinto en cada sesión, y comprobar mediante reconocimiento del habla que el texto pronunciado es el correcto.

Una segunda clasificación de los sistemas de reconocimiento de locutor hace referencia a la dependencia o no del texto pronunciado por el locutor.

a) Sistemas independientes del texto

El proceso de identificación o verificación debe funcionar correctamente para cualquier texto usado en los procesos de entrenamiento y test.

b) Sistemas dependientes del texto

El texto debe ser el mismo en los procesos de entrenamiento y test, o es conocido. Lógicamente la dependencia del texto puede permitir mejorar el comportamiento, pero en determinadas aplicaciones puede ser inconveniente, o simplemente inviable, por desconocer el locutor que está siendo sometido a un proceso de verificación/identificación y, por tanto, ser incontrolable el contenido del mensaje.

En algunos casos, la dependencia con el texto está implícita, como por ejemplo, cuando se realiza el entrenamiento y el test con cadenas de dígitos. En este caso, aunque el orden de los dígitos varíe, el texto pronunciado contiene aproximadamente la misma información temporal (aunque no en la misma secuencia). Especialmente si el número de dígitos pronunciado es suficientemente elevado.

Una tercera clasificación hace referencia a cómo se ha tomado el conjunto de locutores. Atendiendo a ésta clasificación, se puede dividir en dos grupos:

a) Conjunto cerrado

Consiste en identificar a un locutor de entre un conjunto de P locutores conocidos. Lógicamente, cuanto mayor sea el número de locutores posibles (P) mayor es la probabilidad de equivocarse.

b) Conjunto abierto

Consiste en decidir si un locutor determinado pertenece a un conjunto de P locutores conocidos. En este caso, no es necesario decidir cuál es de los P locutores.

En verificación del locutor los resultados se expresan en función de dos tipos de errores:

- Tasa de falsa aceptación: Es la probabilidad de dar como válido aun impostor (verificar afirmativamente un locutor diferente de quien dice ser).
- Tasa de falso rechazo: Es la probabilidad de dar como no válido a un usuario correcto (verificar negativamente a un locutor que realmente es quien dice ser).

Se trata de tasas contrapuestas, puesto que reducir la tasa falsa de aceptación implica aumentar la de falso rechazo, haciendo más restrictivo el sistema, y reducir la tasa de falso rechazo implica aumentar la de falsa aceptación, haciendo menos restrictivo el sistema. Normalmente se adopta un compromiso consistente en igualar ambas tasas.

Los diferentes pasos involucrados en el proceso de identificación de locutor son los siguientes:

1. Adquisición de señal de la voz.
2. Extracción de características (coeficientes de predicción lineal (LPC), cepstrum, etc.).
3. Comparación de las características con los modelos de locutor pertinentes almacenados en memoria y toma de decisiones (aceptación, rechazo o identificación).

Las técnicas de verificación e identificación de locutores pretenden obtener la autenticación de una determinada persona, o la identificación de su identidad dentro de un conjunto de locutores posibles. Las principales aplicaciones de este campo se encuentran en la seguridad: control de accesos a recintos, informaciones restringidas, etc.

ADQUISICIÓN DE LA SEÑAL DE VOZ

Consiste en captar mediante un micrófono la onda acústica producida por el locutor y digitalizarla para poderla tratar en un ordenador. Es importante destacar que tanto el micrófono usado como la sala en la que se realiza la grabación, puede afectar las tasas de reconocimiento. Especialmente si no son los mismos en los procesos de entrenamiento (etapa de aprendizaje del sistema) y de test (etapa de funcionamiento del sistema).

EXTRACCIÓN DE CARACTERÍSTICAS

Consiste en dividir la señal de voz en tramas de típicamente 10 a 30 ms de forma de onda y transformarlas en un vector N -dimensional de características. El resultado es una secuencia de vectores x_i que se compararán con los diferentes modelos, para decidir a qué locutor pertenecen.

Mediante la parametrización se consigue reducir la cantidad de datos a tratar. Por ejemplo, si la señal de voz se adquiere a 8000 muestras por segundo cuantificadas a 16 bits/muestra, parametrizando tramas de 20 ms a 14 coeficientes se obtiene una reducción de datos de 11.4 (160/14). De esta forma se reduce la complejidad computacional del proceso de reconocimiento. Por otra parte mediante la parametrización se consigue transformar la señal de voz a un nuevo espacio de características, en el cual es más sencillo diferenciar entre locutores.

Entre las características más usadas cabe destacar:

- Coeficientes LPC y sus derivados: Log Area Ratios (LAR), Line Spectrum Frequencies (LSF), etc.
- Coeficientes Cepstrum y sus derivados: Mel-Cepstrum, derivadas cepstrales, coeficientes cepstrales ponderados, etc.

A partir de los vectores de características se puede extraer un modelo de locutor, que pertenecerá a una de las siguientes clases:

- Modelo paramétrico: supone una estructura caracterizada por unos parámetros.
- Modelo no paramétrico: no supone ningún modelo de la función de densidad de probabilidad.

MODELO LPC (Coeficientes de Predicción Lineal)

Se trata de una de las técnicas más potentes de análisis de voz, y uno de los métodos más útiles para codificar voz con buena calidad a una tasa de bits relativamente baja.

Proporciona unas aproximaciones a los parámetros de la voz muy precisas.

Se concentra en modelar las resonancias de la garganta al generar la voz.

FUNCIONAMIENTO DEL LPC

Su función es representar la envolvente espectral de una señal digital de voz de forma comprimida, utilizando la información de un modelo lineal.

Analiza el habla estimando los formantes, eliminando sus efectos de la señal, y a continuación estimando la intensidad y frecuencia de la señal restante.

Como las señales de voz varían con el tiempo, este proceso se realiza en pequeños segmentos de la señal de voz, llamados “frames”

En general, 30 a 50 frames por segundo resultan en una señal de voz inteligible y con una buena compresión.

APLICACIONES CON LPC

- LPC se utiliza generalmente para el sintetizador de voz.
- También se utiliza por las compañías telefónicas como método de compresión de voz, como por ejemplo en el estándar GSM.
- Otra aplicación es en redes wireless con seguridad, donde la voz debe digitalizarse, y encriptarse para ser enviada por un canal de capacidad limitada.
- En la música para combinar el sonido de instrumentos con la voz, en la que una señal enviada por un instrumento se combina con un filtro estimado a partir de la voz del cantante.

I.4 REDES NEURONALES

Las redes neuronales artificiales RNA (o en inglés Artificial Neural Systems ANS) son sistemas, hardware o software, de procesamiento, que copian esquemáticamente la estructura neuronal del cerebro para tratar de reproducir sus capacidades. Las RNA son capaces de aprender de la experiencia a partir de las señales o datos provenientes del exterior, dentro de un marco de computación paralela y distribuida, fácilmente implementable en dispositivos hardware específicos.

Para entender mayormente las RNA es importante explicar los conceptos básicos de los sistemas neuronales biológicos y el paralelismo entre ambos.

Una neurona biológica es una célula especializada en procesar información.

Se estima que el sistema nervioso contiene alrededor de cien mil millones de neuronas. Las neuronas presentan múltiples formas, aunque muchas de ellas presentan un aspecto similar, con un cuerpo o soma (de entre 10 y 80 micras de longitud), del que surge un denso árbol de ramificaciones (árbol dendrítico) compuesto por las dendritas, y del cual parte una fibra tubular denominada axón (cuya longitud varía desde las 100 micras hasta el metro en el caso de las neuronas motoras), que también se ramifica en su extremo final para conectar con otras neuronas. Ver la figura I.4.1.

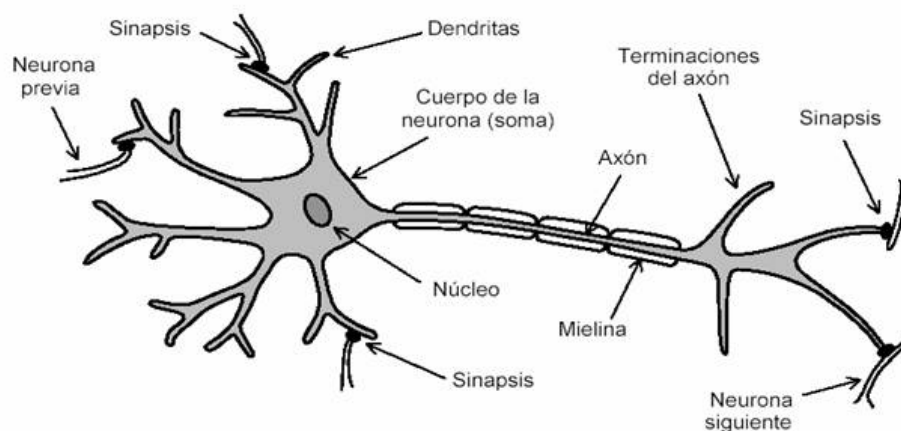


Figura I.4.1. Estructura de una neurona biológica típica

Desde un punto de vista funcional, las neuronas constituyen procesadores de información sencillos. Como todo sistema de este tipo, poseen un canal de entrada de información, las dendritas, un órgano de cómputo, el soma, y un canal de salida, el axón.

Ahora bien, la unión entre dos neuronas se denomina sinapsis. En el tipo de sinapsis más común no existe un contacto físico entre las neuronas, sino que éstas permanecen separadas

por un pequeño vacío de unas 0.2 micras. En relación a la sinapsis, se habla de neuronas presinápticas (la que envía las señales) y postsinápticas (las que la recibe). Las sinapsis son direccionales, es decir, la información fluye siempre en un único sentido.

Las señales nerviosas se pueden transmitir eléctrica o químicamente. La transmisión química prevalece fuera de las neuronas, mientras que la eléctrica lo hace en el interior. La transmisión química se basa en el intercambio de neurotransmisores, mientras que la eléctrica hace uso de descargas que se producen en el cuerpo celular, y que se propagan por el axón.

Las RNA imitan la estructura hardware del sistema nervioso, con la intención de construir sistemas de procesamiento de la información paralelos, distribuidos y adaptativos, que pueden presentar un cierto comportamiento “inteligente”.

Como puede apreciarse en la tabla siguiente, el cerebro y un computador convencional son mucho más diferentes de lo que suele suponerse cuando se habla de “cerebros electrónicos”. Debe recordarse que un computador convencional es, en esencia, una máquina de von Neumann, construida en torno a una única CPU o procesador, que ejecuta de un modo secuencial un programa almacenado en memoria. Por el contrario, el cerebro no está compuesto por un único procesador, sino por miles de millones de ellos (neuronas), aunque muy elementales. Curiosamente, las neuronas son mucho más simples, lentas y menos fiables que una CPU, y a pesar de ello, existen problemas difícilmente abordables mediante un computador convencional, que el cerebro resuelve eficazmente (reconocimiento del habla, visión de objetos inmersos en ambiente natural, respuesta ante estímulos del entorno, etc.).

	Cerebro	Computador
Velocidad de proceso	= 10^{-2} seg. (100 Hz)	= 10^{-9} seg. (1000 MHz)
Estilo de procesamiento	paralelo	secuencial
Número de procesadores	10^{11} - 10^{14}	pocos
Conexiones	10 000 por procesador	pocas
Almacenamiento del conocimiento	distribuido	direcciones físicas
Tolerancia a fallos	amplia	nula
Tipo de control del proceso	auto-organizado	centralizado

Tabla I.4.1. Cerebro frente a computador convencional

Los tres conceptos clave de los sistemas nerviosos, que se pretende emular en los artificiales, son: paralelismo de cálculo, memoria distribuida y adaptabilidad al entorno.

Las neuronas, en definitiva, se autoorganizan, aprendiendo del entorno y adaptándose a él, y de esta autoorganización emergen ricas propiedades de procesamiento. Nuestra capacidad de percepción y, en última instancia, nuestro pensamiento, son producto de ello.

ESTRUCTURA DE UN SISTEMA NEURONAL ARTIFICIAL

Los elementos básicos de un sistema neuronal biológico son las neuronas que se agrupan en conjuntos compuestos por millones de ellas organizadas en capas, constituyendo un sistema con funcionalidad propia. Un conjunto de estos subsistemas da lugar a un sistema global (el sistema nervioso, en el caso biológico). En la realización de un sistema neuronal artificial puede establecerse una estructura jerárquica similar. El elemento esencial de partida será la neurona artificial, que se organizará en capas, varias capas constituirán una red neuronal; y, por último, una red neuronal (o un conjunto de ellas), junto con las interfaces de entrada y salida, más los módulos convencionales necesarios, constituirán en sistema global de proceso. Ver Figura I.4.2.

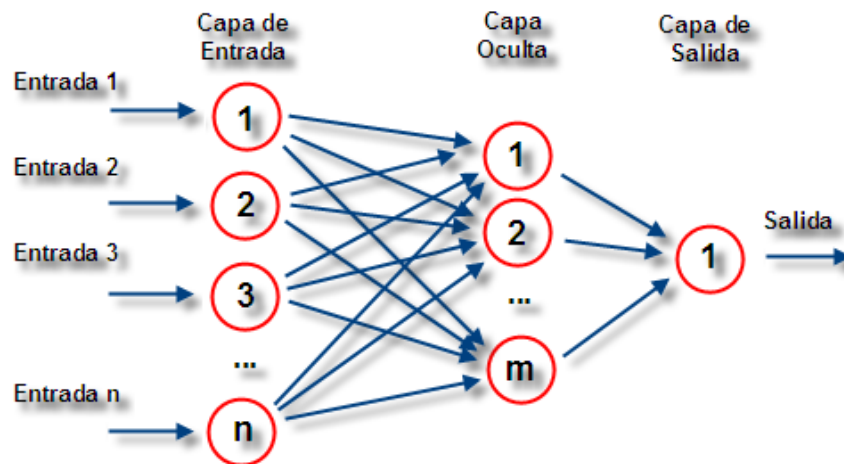


Figura I.4.2. Representación de una red neuronal artificial

MODELO GENERAL DE NEURONA ARTIFICIAL

Se denomina procesador elemental o neurona a un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas, proporciona una única respuesta o salida. Los elementos que constituyen la neurona de etiqueta i son los siguientes:

- Conjunto de entradas, $x_j(t)$
- Pesos sinápticos de la neurona i , w_{ij} que representan la intensidad de interacción entre cada neurona presináptica j y la neurona postsináptica i .
- Regla de propagación $\sigma(w_{ij}, x_j(t))$, que proporciona el valor del potencial postsináptico $h_i(t) = \sigma(w_{ij}, x_j(t))$ de la neurona i , en función de sus pesos y entradas.
- Función de activación $f_i(a_i(t-1), h_i(t))$, que proporciona el estado de activación actual $a_i(t) = f_i(a_i(t-1), h_i(t))$ de la neurona i , en función de su estado anterior $a_i(t-1)$ y de su potencial postsináptico actual.
- Función de salida $F_i(a_i(t))$, que proporciona la salida actual $y_i(t) = F_i(a_i(t))$ de la neurona i en función de su estado de activación.

De este modo, la operación de la neurona i puede expresarse como

$$y_i(t) = F_i(f_i[a_i(t-1), \sigma_i(w_{ij}, x_j(t))])$$

Este modelo de neurona formal se inspira en la operación de la biológica, en el sentido de integrar una serie de entradas y proporcionar cierta respuesta, que se propaga por el axón.

MODELO ESTÁNDAR DE NEURONA ARTIFICIAL

En la neurona estándar se considera la regla de propagación como la suma ponderada y la función de salida es la identidad. De esta forma, la neurona estándar consiste en:

- Un conjunto de entradas, $x_j(t)$ y pesos sinápticos w_{ij}
- Una regla de propagación $h_i(t) = \sigma(w_{ij}, x_j(t))$; $h_i(t) = \sum w_{ij} x_j(t)$ es la más común
- Una función de activación $y_i(t) = f_i(h_i(t))$ que representa simultáneamente la salida de la neurona y su estado de activación.

Con frecuencia se añade al conjunto de pesos de la neurona un parámetro adicional θ_i , que denominaremos umbral (bias), que se resta del potencial postsináptico, por lo que el argumento de la función de activación queda:

$$\sum_j w_{ij} x_j - \theta_i$$

Lo que representa añadir un grado de libertad adicional a la neurona.

En conclusión, el modelo de neurona que denominamos estándar queda:

$$y_i(t) = f_i(\sum_j w_{ij} x_j - \theta_i)$$

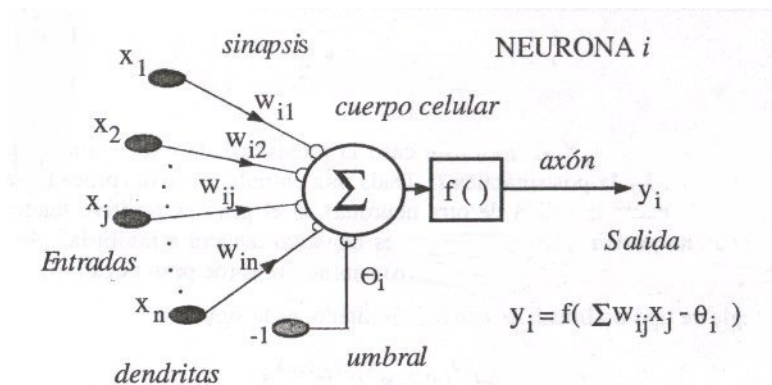


Figura I.4.3. Representación de una neurona artificial

CARACTERÍSTICAS DE LAS REDES NEURONALES

Las características de operación de las RNA son las siguientes:

Pesos

Las RNA pueden tener factores de pesos fijos o adaptables. Las que tienen pesos adaptables emplean leyes de aprendizaje para ajustar el valor de la fuerza de una interconexión con otras neuronas. Si las neuronas utilizan pesos fijos, entonces su tarea deberá estar previamente definida. Los pesos serán determinados a partir de una descripción completa del problema. Por otra parte, los pesos adaptables son esenciales si no se conoce previamente cual deberá de ser su valor correcto.

Tipos de aprendizaje

Existen dos tipos de aprendizaje: supervisado y no supervisado. El primero ocurre cuando se le proporciona a la red tanto la entrada como la salida correcta, y la red ajusta sus pesos tratando de minimizar el error de su salida calculada. Este tipo de entrenamiento se aplica por ejemplo, en el reconocimiento de patrones. El entrenamiento no supervisado se presenta cuando a la red se le proporcionan únicamente los estímulos, y la red ajusta sus interconexiones basándose únicamente en sus estímulos y la salida de la propia red. Las leyes de aprendizaje determinan como la red ajustará sus pesos utilizando una función de error o algún otro criterio. La ley de aprendizaje adecuada se determina en base a la naturaleza del problema que se intenta resolver.

Fases de operación

Las RNA adaptables tienen dos fases en su operación:

1. Entrenamiento de la red. El usuario proporciona a la red un número "adecuado" de estímulos de entrada, y de salida, la red entonces ajusta su pesos de interconexión o sinapsis hasta que la salida de la red esta "lo suficientemente cerca" de la salida correcta.
2. Recuperación de lo aprendido. A la red se le presenta un conjunto de estímulos de entrada y esta simplemente calcula su salida. Cuando la red emplea entrenamiento no supervisado, algunas veces será necesario que reajuste su sinápsis durante la fase de recuperación.

No son algorítmicas

La gran diferencia del empleo de las redes neuronales en relación con otras aplicaciones de la computación radica en que no son algorítmicas, esto es no se programan haciéndoles seguir una secuencia predefinida de instrucciones. Las RNA generan ellos mismos sus propias

"reglas", para asociar la respuesta a su entrada; es decir, aprende por ejemplos y de sus propios errores.

El conocimiento de un RNA se encuentra en la función de activación utilizada y en los valores de sus pesos.

Asociar y generalizar sin reglas como en el cerebro humano

Las redes neuronales formadas por los perceptrones se interconectan en forma muy similar a como las neuronas humanas se disponen en la corteza cerebral humana, y lo más importante, son capaces de asociar y generalizar sin reglas. Han sido utilizadas con gran éxito para reconocer retornos de sonar bajo el agua, escritura a mano, voz, topografía de terrenos, controlar brazos de robots, evaluar datos personales, modelar fenómenos cognoscitivos, y predecir tendencias financieras.

Requieren de algún tipo de patrón

La clase de problemas que mejor se resuelven con las redes neuronales son los mismos que el ser humano resuelve mejor: Asociación, evaluación, y reconocimiento de patrones. Las redes neuronales son perfectas para problemas que son muy difíciles de calcular pero que no requieren de respuestas perfectas, sólo respuestas rápidas y buenas. Tal y como acontece con el escenario bursátil en el que se quiere saber ¿compro?, ¿vendo?, ¿mantengo?, o en el reconocimiento cuando se desea saber ¿se parece? ¿es el mismo pero tienen una ligera modificación?

Por otra parte, las redes neuronales son muy malas para cálculos precisos, procesamiento serie, y no son capaces de reconocer nada que no tenga inherentemente algún tipo de patrón. Es por esto, que no pueden predecir la lotería, ya por definición es un proceso al azar.

Existen varias formas de hacer las conexiones en una RNA, así como existen varias formas de conectar neuronas biológicas en el cerebro. Cada tipo sirve para diferentes procesos, el elegir la correcta topología y sus características, es imprescindible para lograr fácilmente la solución del problema.

APLICACIONES DE LAS REDES NEURONALES

Desde el punto de vista de las aplicaciones, la ventaja de las RNA, reside en el procesamiento paralelo, adaptativo y no lineal. Las RNA han encontrado muchas aplicaciones con éxito en la visión artificial, en el procesamiento de señales e imágenes, reconocimiento del habla y de caracteres, sistemas expertos, análisis de imágenes médicas, control remoto, control de robots, inspección industrial y exploración científica. El dominio de aplicación de las RNA se puede clasificar de la siguiente forma: asociación y clasificación, regeneración de patrones, regresión y generalización, y optimización.

Las Redes Neuronales son útiles para:

- Clasificación de patrones
- Aproximación de funciones
- Mapeo
- En los que exista información imprecisa y con tolerancia a fallos
- Aprendizaje Adaptativo.
- Autoorganización.
- Tolerancia a Fallos.
- Operación en tiempo real.
- Fácil construcción en circuitos integrados.

TOPOLOGÍAS DE REDES NEURONALES

La topología es el número de elementos de procesado que forman la red y las interconexiones existentes entre ellos.

Los elementos de procesado de una red neuronal se distribuyen por capas (conjunto de elementos de procesado que se encuentran en el mismo nivel en la estructura). Existen tres tipos de capas:

1. Capa de entrada: Dependen del problema a abordar.
2. Capa de salida: Salidas que proporciona la estructura.
3. Capas ocultas: conjunto de elementos de procesado asociados en capas que se encuentran entre la capa de entrada y la capa de salida.

El número de elementos de procesado puede variar en unas capas respecto a otras. Las funciones (tanto de activación como de salida) de los elementos de procesado pueden ser diferentes. No existe limitación en el número de capas ocultas. En general, cuanto mayor sea la diferencia entre los elementos de entrada y los de salida, será necesario dotar a la estructura de más capas ocultas. Estas capas ocultas crean una representación interna de los patrones de entrada. La habilidad de la red para procesar información crece en proporción directa al número de capas ocultas.

En función de la dirección del movimiento de la información se puede realizar una clasificación en: redes feedforward, redes feedback, redes feedlateral y redes recurrentes, tal como puede verse en la siguiente figura:

También es posible que las redes sean de una capa con el modelo de pesos hacia atrás o bien el modelo multicapa hacia adelante. Es posible así mismo, el conectar varias redes de una sola capa para dar lugar a redes más grandes.

MECANISMOS DE APRENDIZAJE

Construir un sistema que aprenda ha sido tradicionalmente uno de los objetivos más escurridizos de la inteligencia artificial. El aprendizaje es un proceso de adaptación al entorno; durante el se crean y manipulan representaciones que sean capaces de explicar dicho entorno.

En las redes neuronales la esencia del aprendizaje se encuentra en la modificación colectiva de los pesos de los elementos de procesado. Una definición de aprendizaje en redes neuronales podría ser la siguiente: Proceso por el cual los parámetros libres de una red neuronal son ajustados a través de un proceso continuo de estimulación por parte del entorno en donde se sitúa el sistema. El tipo de aprendizaje viene determinado por la forma en la que tienen lugar dichos cambios. Esta definición implica tres hechos fundamentales:

La red neuronal se encuentra estimulada por el entorno, cambia como consecuencia de esos estímulos y responde de manera diferente al entorno a causa de los cambios que se han producido en sus estructura interna.

Existen tres paradigmas de aprendizaje: Supervisado, No supervisado (o auto supervisado) y Reforzamiento (se puede considerar como un caso especial del aprendizaje supervisado).

Supervisado

En el aprendizaje supervisado (aprender con un maestro), la adaptación sucede cuando el sistema compara directamente la salida que proporciona la red con la salida que se desearía obtener de dicha red. Existen tres tipos básicos: por corrección de error, por refuerzo y estocástico.

En el aprendizaje por corrección de error el entrenamiento consiste en presentar al sistema un conjunto de pares de datos, representando la entrada y la salida deseada para dicha entrada (este conjunto recibe el nombre de conjunto de entrenamiento). El objetivo es minimizar el error entre la salida deseada y la salida que se obtiene.

El aprendizaje por refuerzo es más lento que el anterior. No se dispone de un ejemplo completo del comportamiento deseado pues no se conoce la salida deseada exacta para cada entrada sino que se conoce el comportamiento de manera general para diferentes entradas. La relación entrada-salida se realiza a través de un proceso de éxito o fracaso, produciendo este una señal de refuerzo que mide el buen funcionamiento del sistema. La función del supervisor es más la de un crítico que la de un maestro.

El aprendizaje estocástico consiste básicamente en realizar cambios aleatorios de los valores de los pesos y evaluar su efecto a partir del objetivo deseado.

No supervisado

En el aprendizaje no supervisado (o auto supervisado), la red se sintoniza a las regularidades estadísticas de datos de entrada de forma tal que se formen categorías que optimizan, respecto

a los parámetros libres de la red, una medida de calidad de la tarea que se quiere realizar con la red. El funcionamiento de estas redes se basa en la búsqueda de características, regularidades, correlaciones y categorías del conjunto de datos de entrada.

Existen diferentes interpretaciones que se le pueden dar a las salidas generadas por una red que utilice este tipo de aprendizaje: similitud (semejanzas entre la información actual y la información pasada), clusterización (establecimiento de clases), codificación (versión codificada de las entradas) o mapeo (representación topográfica de los datos de entrada).

Reforzamiento

En el aprendizaje con reforzamiento, la red aprende de relaciones entrada-salida. Sin embargo, al contrario que en el aprendizaje supervisado, el sistema conoce si la salida es correcta o no, pero no conoce la salida correcta.

RED NEURONAL DE RETROPROPAGACION

Además del perceptrón, existe otra estructura de las redes neuronales: la red de retropropagación, en la cual se interconectan varias unidades de procesamiento en capas; las neuronas de cada capa no se interconectan entre sí. Sin embargo, cada neurona de una capa proporciona una entrada a cada una de las neuronas de la siguiente capa, esto es, cada neurona transmitirá su señal de salida a cada neurona de la capa siguiente. La Figura I.4.4 muestra un ejemplo esquemático de la arquitectura de este tipo de redes neuronales.

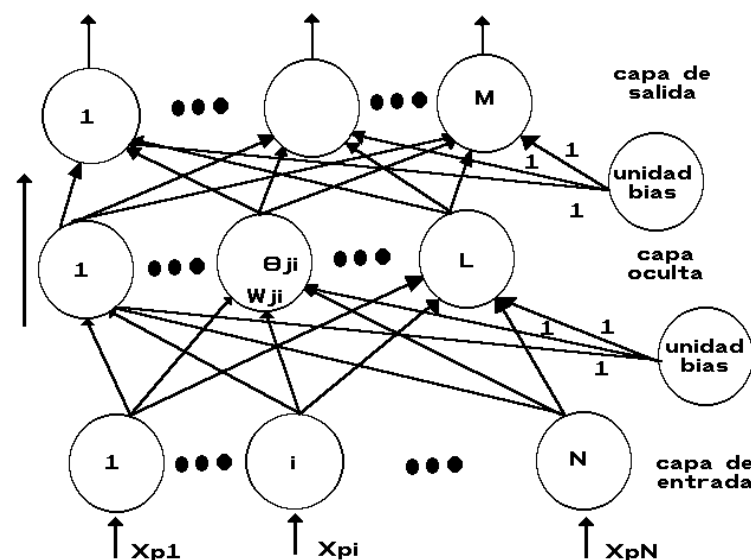


Figura I.4.4. Red neuronal artificial de retropropagación

El perceptrón solo es el ejemplo más elemental de una red neuronal, de hecho, no puede siquiera ser considerado una "red", puesto que no intervienen otros elementos. Si se combinan varios perceptrones en una "capa", y los estímulos de entrada después se suman, tendremos ya

una red neuronal. Una red neuronal muy eficaz para resolver fundamentalmente problemas de reconocimiento de patrones es la red neuronal de propagación hacia atrás (retropropagación), en inglés back propagation network.

Origen

Al parecer el algoritmo fue ideado a principios de los 70's por Werbos, y redescubierto a principios de los 80's por Parker y Rumelhart independientemente; sin embargo, no se hizo popular hasta 1986, cuando Rumerlhart, Hinton y Williams presentaron una descripción clara y concisa del mismo. Y es que en un primer momento no se valoró como se merecía. El hecho de que permaneciera en el olvido tanto tiempo también debe ser una consecuencia de la condición interdisciplinaria del campo, repartido entre las matemáticas y ciencias de la computación, las neurociencias y la psicología.

Desde esa fecha, 1986, han surgido nuevas versiones que han tratado de aumentar la velocidad de convergencia del algoritmo y han tratado de superar algunos de sus inconvenientes, como la tendencia a alcanzar mínimos locales y no globales.

Características

Al hablar de redes de retropropagación o redes de propagación hacia atrás hacemos referencia a un algoritmo de aprendizaje más que a una arquitectura determinada. La retropropagación consiste en propagar el error hacia atrás, es decir, de la capa de salida hacia la capa de entrada, pasando por las capas ocultas intermedias y ajustando los pesos de las conexiones con el fin de reducir dicho error. Hay distintas versiones o reglas del algoritmo de retropropagación y distintas arquitecturas de conexión a las que pueden ser aplicados.

Durante mucho tiempo no se dispuso de algoritmos para entrenar redes multicapa, y como las redes de una capa estaban muy limitadas en cuanto a lo que eran capaces de representar, el campo de las redes neuronales artificiales estaba estancado. La invención y perfeccionamiento del algoritmo de retropropagación dio un gran impulso al desarrollo de este campo. Tiene un buen fundamento matemático y a pesar de sus limitaciones ha expandido enormemente el rango de problemas donde se aplican las redes neuronales artificiales.

Descripción matemática

Se explica una versión del algoritmo (Hinton, 1992) para redes con las siguientes características:

- No recurrentes
- Función de activación sigmoideal
- Capas totalmente interconectadas
- Operación totalmente síncrona

Algoritmo de Aprendizaje

Aleatorizamos los pesos de las conexiones.

Presentamos un patrón de entrada y calculamos la salida.

Dada una unidad j-ésima de la capa de salida y unidades i-ésimas de la capa oculta inmediatamente anterior, calculamos la entrada total ponderada y la salida o activación de la misma.

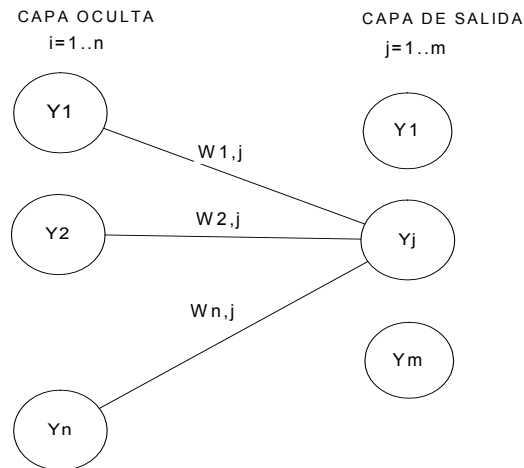


Figura I.4.5. Ilustra el cálculo de la capa de salida

$$x_j = \sum_{i=1}^n (y_i \cdot w_{ij})$$

$$y_j = \frac{1}{1 + e^{-x_j}}$$

Una vez computadas las actividades de todas las unidades de salida se calcula una estimación del error, generalmente una función cuadrática de los errores individuales cometidos por cada unidad, siendo cada error individual la diferencia entre la salida deseada y la obtenida.

$$E = \frac{1}{2} \sum_j (d_j - y_j)^2 \quad \text{siendo } d_j \text{ la salida deseada para la unidad } j\text{-ésima}$$

Nota: Se van a indicar por un lado las expresiones matemáticas y por otro lado la explicación intuitiva de cada paso. Conviene recordar que nuestro objetivo es calcular como varía el error al variar el peso de cada conexión (tasa de variación del error respecto al peso de una conexión, EP).

El cómputo de la rapidez de variación del error al cambiar la actividad de cada unidad de salida (EA, error respecto a la actividad):

$$EA_j = \frac{\partial E}{\partial y_j} = d_j - y_j$$

Es justamente la diferencia entre la salida deseada y la salida real obtenida, es decir, la diferencia entre la actividad deseada y la actividad real.

El cómputo de la rapidez de variación del error al cambiar la entrada total que recibe cada unidad de salida:

$$\delta_j = \frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial x_j} = EA_j y_j (1 - y_j) = y_j (1 - y_j) (d_j - y_j)$$

Es igual a la tasa de variación del error al variar su activación multiplicado por la tasa de variación de la activación al cambiar su entrada (que es justamente la derivada de la función sigmoïdal).

El cómputo de la rapidez de variación del error al ser modificado un peso de la conexión aferente a una unidad de salida:

$$EP_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial w_{ij}} = \delta_j y_i$$

Es igual a la tasa de variación del error al variar su entrada, por la tasa de variación de la entrada al variar ese peso.

Hasta ahora sabemos calcular el EA sólo para las unidades de salida, ¿qué pasa con las unidades ocultas? En este caso no tenemos una estimación directa del error aportado por cada unidad oculta; aquí es donde interviene la retropropagación o propagación hacia atrás del error:

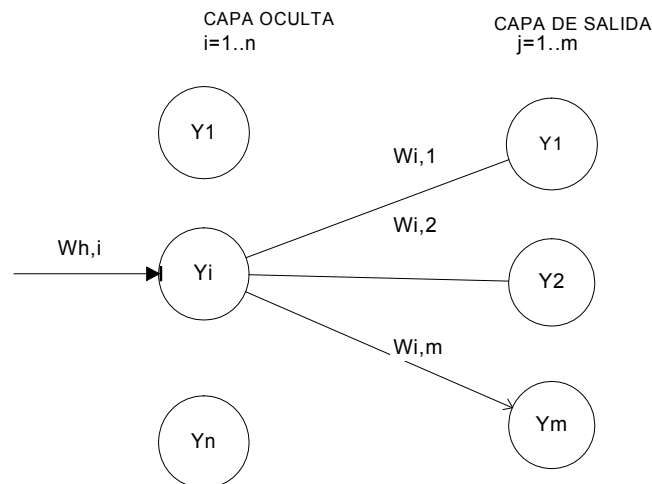


Figura I.4.6. Ilustra la propagación hacia atrás del error.

La unidad i -ésima de la capa oculta afecta a todas las unidades de salida, por lo tanto, para estimar como varía el error al variar la actividad de esa unidad oculta, habrá que sumar los efectos individuales de su actividad sobre todas las neuronas de salida. Cada efecto individual sobre la variación del error, será igual a la tasa de variación del error de la unidad de salida al cambiar su entrada total, multiplicado por la tasa de variación de su entrada al variar la actividad de la unidad oculta.

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_{j=1}^m \left(\frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial y_i} \right) = \sum_{j=1}^m (\delta_j \cdot w_{ij})$$

Conociendo EA para las unidades de cualquier capa podemos calcular δ y EP con las expresiones ya conocidas.

$$\delta_i = EA_i y_i (1 - y_i) = y_i (1 - y_i) \sum_{j=1}^m (\delta_j w_{ij})$$

$$EP_{hi} = \delta_i \cdot y_h$$

Disponiendo de la tasa de variación del error respecto al peso de una conexión (EP), podemos usar distintas reglas para modificar ese peso en aras a reducir dicho de error. Una de las primeras reglas que aprovechó este algoritmo es la regla delta generalizada, que calcula el incremento a aplicar a un peso como una proporción directa de la tasa de variación del error.

$$\Delta w_{ij} = \eta \delta_j y_i \quad \text{siendo } \eta \text{ el coeficiente de aprendizaje, típicamente con valores comprendidos entre 0.01 y 1.0}$$

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}$$

El algoritmo de retropropagación presenta ciertos problemas, algunos referentes a su dudosa plausibilidad neurofisiológica, y otros referentes a ciertos aspectos computacionales.

Los resultados dependen de los valores iniciales, aleatorios, de las conexiones. Esto hace que sea conveniente entrenar varias redes con distintos valores iniciales y elegir la que mejor funcione.

A veces se requiere mucho tiempo para obtener soluciones sencillas. Este problema se reduce gracias al aumento de potencia de los procesadores y al uso de nuevas tecnologías, sin embargo, el tiempo de cómputo aumenta mucho al aumentar el tamaño de la red. Si bien el volumen de cálculo es proporcional al número total de conexiones. En la práctica, al aumentar el tamaño de la red, hacen falta más ejemplos de aprendizaje, y eso provoca un aumento mucho mayor del tiempo de aprendizaje. Para incrementar la velocidad de convergencia se han desarrollado diferentes modificaciones del algoritmo.

La “interferencia catastrófica” o empeoramiento en el rendimiento del sistema, como consecuencia de la incorporación de nuevos ejemplos de aprendizaje.

- La parálisis. Esto sucede cuando los pesos quedan ajustados a valores muy grandes, esto hace operar a las unidades de proceso con una activación muy próxima a 1, y por lo tanto, el gradiente del error tiende a 0, en consecuencia no se producen modificaciones en los pesos, el aprendizaje queda detenido. Por eso es conveniente aleatorizar los pesos de las conexiones con valores pequeños y usar la tasa de aprendizaje también pequeña, a pesar de que se alente el aprendizaje.
- Inestabilidad temporal. Si usamos un coeficiente de aprendizaje elevado, se van a producir incrementos grandes en los pesos, de manera que es fácil pasarse de incremento y tener que tratar de compensarlo en el siguiente ciclo, de manera que se producirían oscilaciones continuas. Esto se soluciona usando un coeficiente pequeño, o, para no tener un aprendizaje muy lento, modificar dicho coeficiente adaptativamente (aumentarlo si el error global disminuye, y disminuirlo en caso contrario).
- El problema de los mínimos locales. El algoritmo de retropropagación usa una técnica por gradiente descendiente, esto significa que sigue la “superficie del error” siempre hacia abajo, hasta alcanzar un mínimo local, pero no garantiza que se alcance una solución globalmente óptima. Sin embargo, se ha comprobado que el hecho de alcanzar mínimos locales no impide que se consigan resultados satisfactorios. Por otro lado, se han desarrollado métodos para solventar este problema, como el modo de operación asíncrono o probabilística y el uso de métodos estadísticos, como el equilibrio termodinámico simulado (ver siguiente apartado).

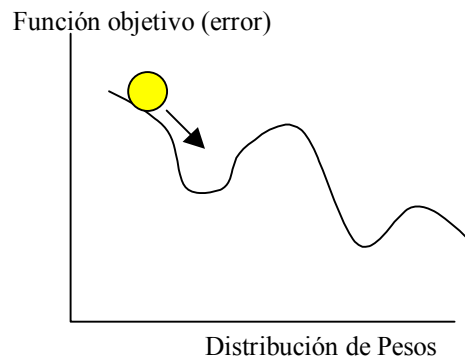


Figura I.4.7. Problema de los mínimos locales

Podemos considerar el error como una superficie llena de desniveles, si soltamos una pelota caerá en algún valle, pero no necesariamente en el más hondo, sino en el más cercano (un mínimo local). Una idea intuitiva para solucionar esto, sería aplicarle cierta energía a esa superficie agitándola o haciéndola vibrar, esto haría saltar a la pelota de valle en valle, como de los valles más profundos es más difícil salir, tendería a estar en valles cada vez más profundos. Si dejamos de agitar esa superficie poco a poco, al final tendremos la pelota en el valle más profundo de la superficie.

Otras técnicas que pueden ayudar a no caer en mínimos locales consisten en añadir cierto nivel de ruido a las modificaciones de los pesos de las conexiones. Otra medida propuesta es añadir ruido a las conexiones, pero esto es más útil para darle robustez y aumentar la

capacidad de generalización de la red. Estas medidas, por contra, aumentan el tiempo de aprendizaje.

Variantes del algoritmo de retropropagación

El algoritmo de retropropagación fue el mayor avance en el campo de la investigación de las redes neuronales. Sin embargo el algoritmo resulta demasiado lento para aplicaciones prácticas.

Cuando se aplica el algoritmo básico de retropropagación a problemas prácticos el tiempo de entrenamiento puede llevar días e incluso semanas de cómputo. Por lo que se ha investigado en la búsqueda de métodos para acelerar el algoritmo de convergencia.

La investigación de algoritmos más veloces se clasifica en dos categorías:

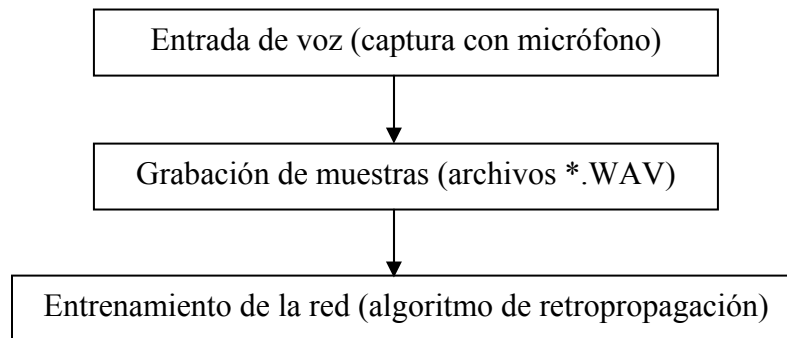
1. Técnicas Heurísticas
 - Momento
 - Razón de aprendizaje variable
2. Técnicas de optimización numérica
 - Algoritmo de gradiente conjugado
 - Algoritmo de Levenberg-Marquardt

CAPÍTULO II

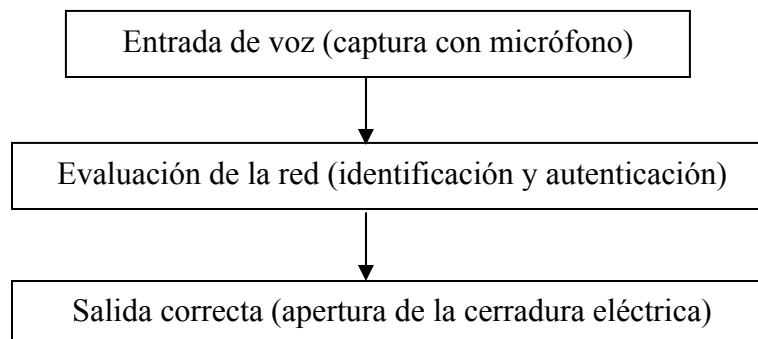
DESARROLLO

Para lograr el objetivo del presente trabajo, se requiere considerar las siguientes etapas:

Primer proceso



Segundo proceso



En el primer proceso se considera la captura de la voz mediante un micrófono conectado hacia la computadora en donde se cuenta con los programas de simulación de redes neuronales artificiales.

A continuación se muestra el programa utilizado en MATLAB para realizar la captura de la voz de dos personas tomando 5 muestras de cada una diciendo la palabra “abrir”, y a su vez eliminar el ruido no deseado y normalizar la señal obtenida.

Cabe señalar que la RNB se programó usando MATLAB R2009a en una computadora portátil marca Sony modelo VGN-C240FE con procesador Intel Core 2, 1.67 GHz.

%%%

captura.m

```
clear all;
clc
fs=8000;
    pause
    y=wavrecord(2.5*fs,fs, 'double');
fr = 8000
    x=y;

% x2 es la señal normalizada %
    mwad=max(x);
    mwid=min(x);
x3=x/(max(mwad,mwid));

% Cálculo de potencia %
x2=x3;
gamma=1000/(fr*.2);
E(1)=0.0;
for i=2:length(x2)
E(i)=(1-gamma)*E(i-1)+gamma*x2(i)*x2(i)/0.00000000052998;
end

% Cálculo de inicio y final de la señal de voz %
maxp=max(E);
M=(maxp*.2);
N=(maxp*.9);
er=0;
for i=1: length(E)
    if E(i)>=M
        Id2=i;
        break;
    end
```

```
end

for j=Id2:-1:1
    if E(j)<=N
        Id1=j;
        er=er+1;
        break;
    end
    if er==0
        Id1=Id2;
    end
end

elz=0;
for i=length(E):-1:1
    if E(i)>=M
        Id22=i;
        break;
    end
end

for j=Id22:1:length(E)
    if E(j)<=N
        Id11=j;
        elz=elz+1;
        break;
    end

    if elz==0
        Id11=Id22;
    end
end

% PX inicio y final de la potencia %
PX=x2(Id1:Id11);

% Señal acotada %
xiyf=x(Id1:Id11);
wavwrite(xiyf,8000,'abrirA1.wav');
subplot(2,1,1);plot(x);
subplot(2,1,2);plot(PX);

%%%%%%%%%
```

Los archivos guardados se generaron como:

Persona no. 1

abrirA1.wav
abrirA2.wav
abrirA3.wav
abrirA4.wav
abrirA5.wav

Persona no. 2

abrirB1.wav
abrirB2.wav
abrirB3.wav
abrirB4.wav
abrirB5.wav

En el apartado de Resultados, se muestran las gráficas de cada una de estas muestras, así como su gráfica de señal acotada.

Como siguiente etapa se presenta la simulación de una red neuronal artificial del tipo retropropagación (RNB), que se entrena para reconocer la voz de dos personas distintas con las muestras ya grabadas en el paso anterior.

De manera general, la red desarrollada consta de 3 capas (entrada-oculta-salida), 11 neuronas para la capa de entrada y 7 para la capa oculta y 3 para la capa de salida.

A través de la aplicación del algoritmo de retropropagación se trató de encontrar un error cuadrático medio menor a 0.005%, con 100 000 de iteraciones (epochs), lo cual tomó pocos minutos.

El algoritmo de la red neuronal de retropropagación se muestra en las siguientes figuras:

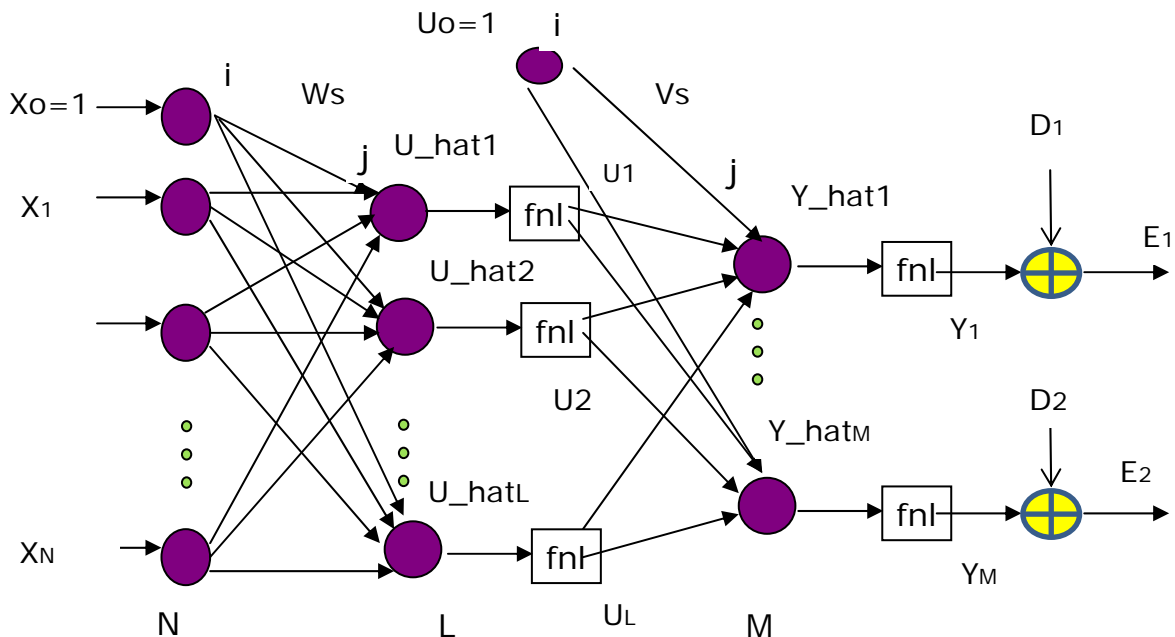


Figura II.1. Algoritmo de retropropagación

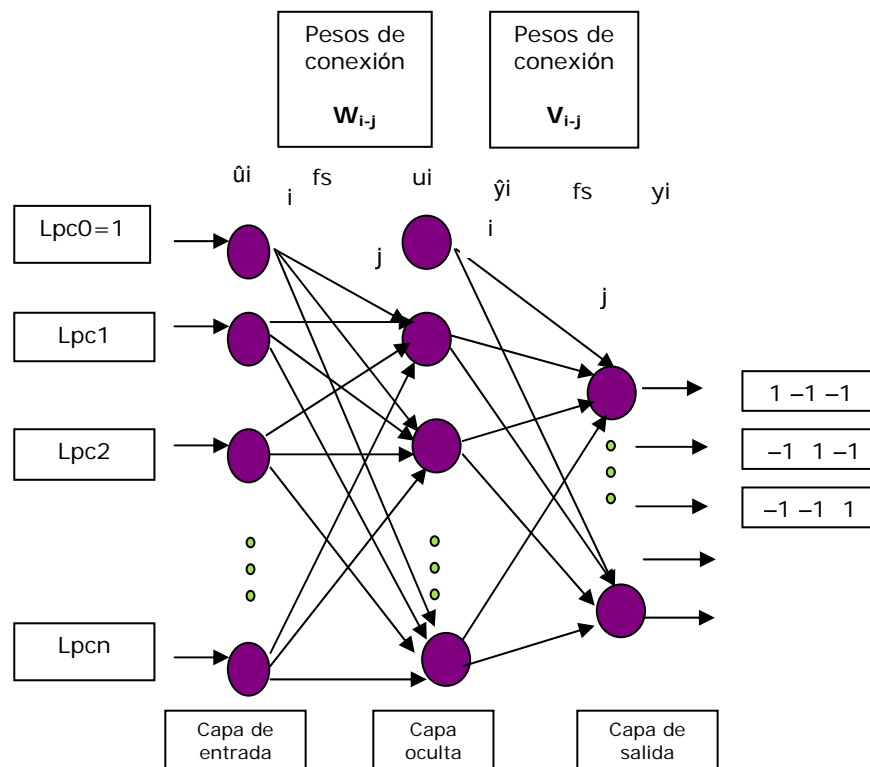


Figura II.2. Algoritmo de retropropagación

El programa en MATLAB que realiza esta función es el siguiente:

%%%

entrena.m

```
clear all;
clc
N=10;

[a1,fr1]=wavread ('abrirA1');
[b1,fr1]=wavread ('abrirB1');

[a2,fr2]=wavread ('abrirA2');
[b2,fr2]=wavread ('abrirB2');

[a3,fr3]=wavread ('abrirA3');
[b3,fr3]=wavread ('abrirB3');
```



```
[a4,fr4]=wavread ('abrirA4');
[b4,fr4]=wavread ('abrirB4');

[a5,fr5]=wavread ('abrirA5');
[b5,fr5]=wavread ('abrirB5');

% Coeficientes en frecuencia de comandos %
[lpca1,ga1]=lpc(a1,N);
[lpcb1,gb1]=lpc(b1,N);

[lpca2,ga2]=lpc(a2,N);
[lpcb2,gb2]=lpc(b2,N);

[lpca3,ga3]=lpc(a3,N);
[lpcb3,gb3]=lpc(b3,N);

[lpca4,ga4]=lpc(a4,N);
[lpcb4,gb4]=lpc(b4,N);

[lpca5,ga5]=lpc(a5,N);
[lpcb5,gb5]=lpc(b5,N);

t=2:11;

pa1=(abs(lpca1(:,t)));
pb1=(abs(lpcb1(:,t)));

pa2=(abs(lpca2(:,t)));
pb2=(abs(lpcb2(:,t)));

pa3=(abs(lpca3(:,t)));
pb3=(abs(lpcb3(:,t)));

pa4=(abs(lpca4(:,t)));
pb4=(abs(lpcb4(:,t)));

pa5=(abs(lpca5(:,t)));
pb5=(abs(lpcb5(:,t)));

N=11;          % neuronas de entrada (incluido el BIAS) %
L=7;           % neuronas de la capa oculta (incluido el BIAS) %
M=3;           % neuronas de salida %
NT=10;         % patrones de entrenamiento %
NI=100000;     % # maximo de iteraciones (epochs) %
epsilon=0.005; % error cuadratico medio requerido %
```

```
Ws=(2*rand(N,L-1))-1;
Vs=(2*rand(L,M))-1;

% Datos de entrenamiento %
PE=[pa1;pb1;
    pa2;pb2;
    pa3;pb3;
    pa4;pb4;
    pa5;pb5];

% Salida deseada de cada uno %
D=[ 1 -1 -1;
    -1 1 -1;
    -1 -1 1;
    1 -1 -1;
    -1 1 -1;
    -1 -1 1;
    1 -1 -1;
    -1 1 -1;
    -1 -1 1;
    1 -1 -1;
    -1 1 -1;
    -1 -1 1;
    1 -1 -1;
    -1 1 -1;
    -1 -1 1];

% for principal no. de iteraciones para entrenamiento %
for k=1:NI
    Err=0;
    for i=1:NT;
        if k<=70
            miu=0.3;
        end
        if k>70&k<=200 % Factor de convergencia %
            miu=0.1;
        end
        if k>200
            miu=0.05;
        end
        X=[1.0,PE(i,:)]; %BIAS y patrones de entrenamiento %

%Calculo de salida de la red %
        U_hat=X*Ws;
```

```

    U=[1.0,f_nl(1,U_hat)]; %agrega BIAS %
    Y_hat=U*Vs;
    Y=f_nl(1,Y_hat);
    E=D(i,:)-Y;

% Adaptación %
    delta1=E.*f_prime(1,Y_hat);
    Vs=Vs+miu*U'*delta1;
    S=Vs*delta1';
    delta2=S(2:L)'.*f_prime(1,U_hat);
    Ws=Ws+miu*X'*delta2;
    Err=Err+E*E';
end % i
Err=Err/(NT*M);
vec_err(1,k)=Err;
fprintf(1,'Err=%4.9f ==> %d\n',Err,k);
    if Err<epsilon
        break;
    end
end % k

for i=1:NT
    X=[1.0,PE(i,:)];
    U_hat=X*Ws;
    U=[1.0,f_nl(1,U_hat)];
    Y_hat=U*Vs;
    Y=hardlims(f_nl(1,Y_hat));
    fprintf(1,'Salida de la Red: %d %d %d \n\n',Y(1),Y(2),Y(3));
end

fidfun=fopen('Wij.dat','w');
fprintf(fidfun,'%6.4f %6.4f %6.4f %6.4f %6.4f %6.4f\n',Ws');
fclose(fidfun);

fidfun=fopen('Vij.dat','w');
fprintf(fidfun,'%6.4f %6.4f %6.4f %6.4f %6.4f %6.4f\n',Vs');
fclose(fidfun);

%%%%%%%%%%%%%%

```

Una vez entrenada la red se pasa al segundo proceso, el cuál va a comenzar nuevamente con la captura de la voz para su posterior evaluación y respuesta.

El programa en MATLAB que realiza esta función es el siguiente:

%%%

captura y evalua.m

```
clear all;
clc
Fs = 8000;
pause
    y = wavrecord(2.5*Fs, Fs, 'double');
    wavplay(y, Fs);

    fr=8000
    x=y;

% x2 es la señal normalizada %
    mwad=max(x);
    mwid=min(x);

x3=x/(max(mwad,mwid));

% Cálculo de potencia %
    x2=x3;
    gamma=1000/(fr*.2);
    E(1)=0.0;
    for i=2:length(x2)
        E(i)=(1-gamma)*E(i-1)+gamma*x2(i)*x2(i)/0.00000000052998;
    end

% Calculo de inicio y final de la señal de voz %
    maxp=max(E);
    M=(maxp*.2);
    N=(maxp*.9);

    er=0;
    for i=1: length(E)
        if E(i)>=M
            Id2=i;
            break;
        end
    end
end
```

```
for j=Id2:-1:1
    if E(j)<=N
        Id1=j;
    er=er+1;
    break;
end
    if er==0
        Id1=Id2;
    end
end

elz=0;
for i=length(E):-1:1
    if E(i)>=M
        Id22=i;
        break;
    end
end

for j=Id22:1:length(E)
    if E(j)<=N
        Id11=j;
    elz=elz+1;
    break;
    end

    if elz==0
        Id11=Id22;
    end
end

PX=x2(Id1:Id11); %PX inicio y final de la potencia %
xiyf=x(Id1:Id11); %xin0 señal acotada %

subplot(2,1,1);plot(x);
subplot(2,1,2);plot(PX);

N=10;
c1=xiyf;
[lpcc1,ga2]=lpc(c1,N);
t=2:11;
c11=(abs(lpcc1(:,t)));
```

```

N=11;           %neuronas de entrada (incluido el BIAS)
L=7;           %neuronas de la capa oculta (incluido el BIAS)
M=3;           %neuronas de salida
NT=1;          %patrones de entrenamiento
epsilon=0.005; %error cuadrático medio requerido

PE=[c11];
fidfun=fopen('Wij.dat','r');
Ws=fscanf(fidfun,'%f',[L-1,inf]);
Ws=Ws';
fclose(fidfun);

fidfun=fopen('Vij.dat','r');
Vs=fscanf(fidfun,'%f',[M,inf]);
Vs=Vs';
fclose(fidfun);

for i=1:1
    X=[1.0,PE(i,:)];
    U_hat=X*Ws;
    U=[1.0,f_nl(1,U_hat)];
    Y_hat=U*Vs;
    Y=hardlims(f_nl(1,Y_hat));
    fprintf(1,'Salida de la Red: %d %d %d \n\n',Y(1),Y(2),Y(3));
end

% Respuestas a la evaluación %
if (Y~= [1 -1 -1])|(Y~= [-1 1 -1]);
    fprintf('error acceso denegado');
end
if (Y== [1 -1 -1]);
    fprintf('abrir persona A');
end
if (Y== [-1 1 -1]);
    fprintf('abrir persona B');
end

%%%%%%%%%%

```

Una vez obtenida la respuesta, si es satisfactoria (abrir persona A o B), esto es, que coincide la voz grabada inicialmente con la voz presentada en el segundo proceso para ser evaluada, se envía una señal a un circuito electrónico, por medio de un puerto, por ejemplo el paralelo (DB25). En caso contrario, si no existe el reconocimiento de la voz, enviará un voltaje cero, indicando un error o acceso denegado.

La señal de salida de 5 Volts será enviada de la computadora, al ejecutar el programa de captura y evaluación, a través de un pin del puerto paralelo y alimentarán al circuito electrónico propuesto que se muestra a continuación.

CIRCUITO ELECTRÓNICO PARA ACTIVAR LA CERRADURA ELÉCTRICA

En el siguiente diagrama presentado se explica la manera de cómo operará el proyecto en combinación de un circuito electrónico que será alimentado una vez que se halla comprobado que la persona que emitió la voz tiene un acceso permitido.

Se puede observar que consta de un push botton, un diodo led, un flip flop SN7473 que operará con la salida de 5 volts del pin de salida del puerto seleccionado en la computadora. La configuración de cada pin del flip flop tiene una función específica, de acuerdo a lo siguiente:

- el pin 11 tierra
- el pin 12 salida digital
- el pin 3, 4 y 14 alimentados con 5 VCC
- el pin 1 entrada al cerrar el push botton dando la orden final

El pin 12 es la salida con señal digital que pasa por una resistencia para llegar al transistor de tipo NPN BC548 del cual su colector llega a un diodo de pico inverso que protege el circuito, mismo que se encuentra en paralelo con el relevador que será alimentado por 12 VCC.

Una vez que este relevador se cierre, el circuito se cerrará por completo con una cerradura eléctrica alimentada por 120 VCA que será la etapa final para permitir de manera mecánica el acceso al lugar permitido.

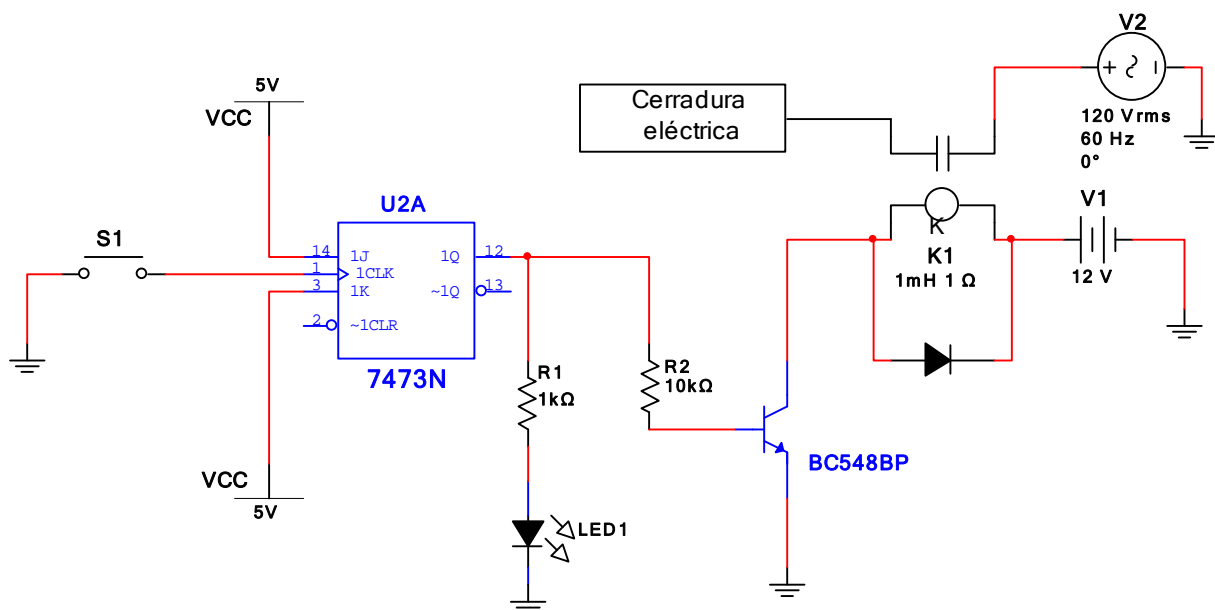
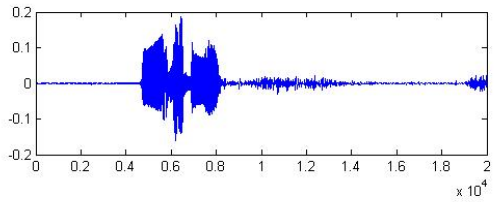


Figura II.1. Circuito electrónico para activar la cerradura eléctrica

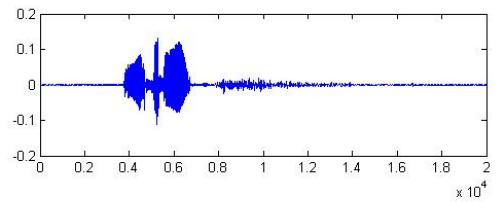
CAPÍTULO III

RESULTADOS

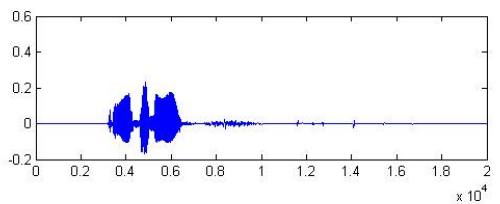
ESPECTROGRAMAS DE LAS MUESTRAS DE VOZ TOMADAS CON LA PALABRA “ABRIR” CON LA PERSONA NO. 1 (A)



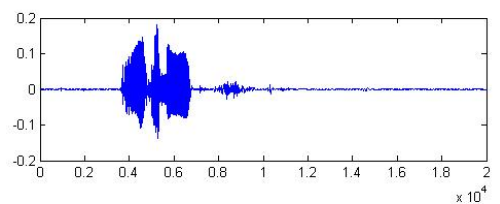
A1



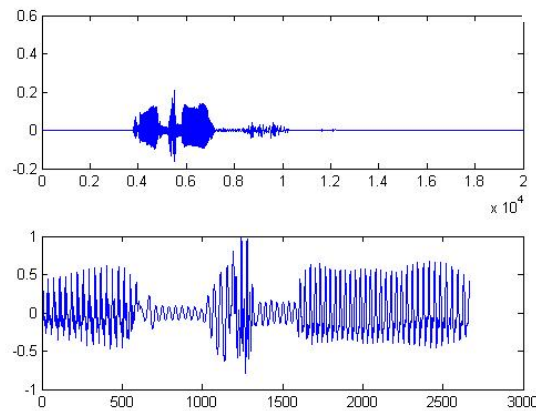
A2



A3

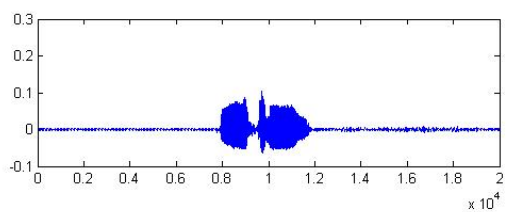


A4

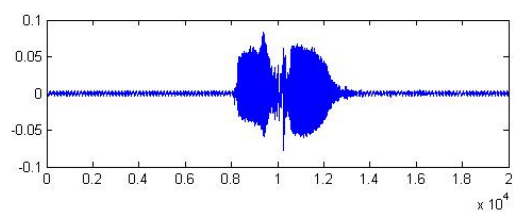


A5

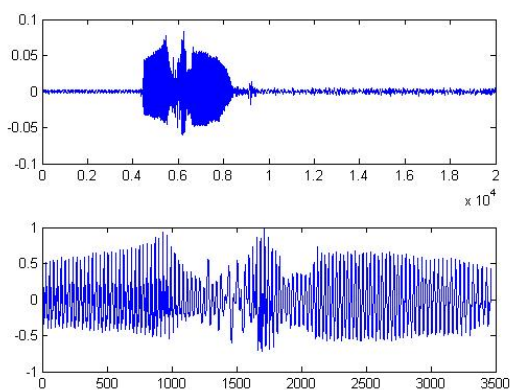
ESPECTROGRAMAS DE LAS MUESTRAS DE VOZ TOMADAS CON LA PALABRA “ABRIR” CON LA PERSONA NO. 2 (B)



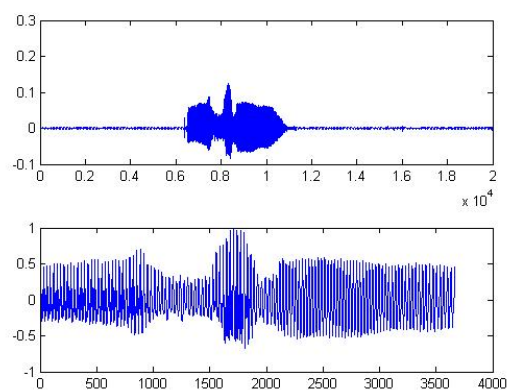
B1



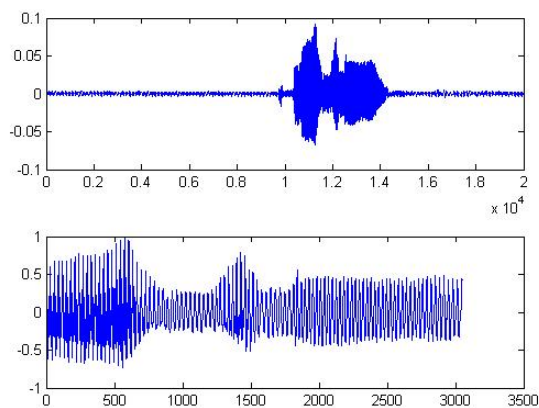
B2



B3



B4



B5

PRUEBA DE ENTRENAMIENTO NO. 1 DE LA RED NEURONAL

Err=0.121660105 ==> 4009

Err=0.121655277 ==> 4010

...

Err=0.119203976 ==> 5000

Err=0.119202424 ==> 5001

...

Err=0.117853192 ==> 6000

Err=0.117851997 ==> 6001

...

Err=0.116880824 ==> 7000

Err=0.116880089 ==> 7001

...

Err=0.116317768 ==> 8000

Err=0.116317332 ==> 8001

...

Err=0.115424104 ==> 9000

Err=0.115410016 ==> 9001

...

Err=0.005074838 ==> 9201

Err=0.005044677 ==> 9202

Err=0.005014963 ==> 9203

Err=0.004985685 ==> 9204

Salida de la Red: 1 -1 -1

Salida de la Red: -1 1 -1

Salida de la Red: -1 -1 1

Salida de la Red: 1 -1 -1

Salida de la Red: -1 1 -1

Salida de la Red: -1 -1 1

Salida de la Red: 1 -1 -1

Salida de la Red: -1 1 -1

Salida de la Red: -1 -1 1

Salida de la Red: 1 -1 -1

PRUEBA DE ENTRENAMIENTO NO. 2 DE LA RED NEURONAL

Err=1.270276096 ==> 1
Err=1.303900886 ==> 2
Err=1.276871025 ==> 3
Err=1.488912066 ==> 4
Err=1.334766151 ==> 5
...
Err=0.842922868 ==> 200
Err=0.640058052 ==> 201
...
Err=0.284104059 ==> 500
Err=0.269905391 ==> 501
...
Err=0.128060166 ==> 700
Err=0.128021936 ==> 701
...
Err=0.123128918 ==> 1000
Err=0.123121140 ==> 1001
...
Err=0.121914329 ==> 1200
Err=0.121909510 ==> 1201
...
Err=0.120730212 ==> 1500
Err=0.120726668 ==> 1501
...
Err=0.119240739 ==> 1700
Err=0.119194884 ==> 1701
...
Err=0.005117221 ==> 1992
Err=0.005091964 ==> 1993
Err=0.005066964 ==> 1994
Err=0.005042217 ==> 1995
Err=0.005017719 ==> 1996
Err=0.004993466 ==> 1997

Salida de la Red: 1 -1 -1
Salida de la Red: -1 1 -1
Salida de la Red: -1 -1 1
Salida de la Red: 1 -1 -1
Salida de la Red: -1 1 -1
Salida de la Red: -1 -1 1
Salida de la Red: 1 -1 -1
Salida de la Red: -1 1 -1
Salida de la Red: -1 -1 1
Salida de la Red: 1 -1 -1

CONCLUSIONES

En la actualidad el gran avance de la tecnología nos permite un mayor control y administración de accesos a lugares que así lo requieren, como es el caso de bancos, centros de cómputo, edificios inteligentes, entre muchos otros.

Aunque la inversión en sistemas de control de acceso no debe ser necesariamente onerosa como la implantación de vidrios a pruebas de balas, guardias armados las 24 horas del día o cámaras de video; las empresas sí deben contemplar controles adecuadamente razonables para evitar el acceso de individuos e incluso de personal "no autorizado" a los centros de trabajo o a las áreas de manejo de datos o información crítica.

Este trabajo es una propuesta muy básica para lograr ese propósito, contar con un sistema de control de acceso aplicando la tecnología biométrica fisiológica, como lo es el reconocimiento de la voz humana logrando una identificación y autenticación.

Algunas razones llevan a concluir que una aproximación por redes neuronales es una solución apropiada para el problema de reconocimiento de voz, en primer lugar debido a que la voz tiene patrones perfectamente delimitados por los fonemas registrados; en segundo lugar, dado que las redes neuronales son adaptativas y entrenables, con muestras de voz grabadas en archivos se pueden tener las características necesarias para permitir una salida de la red idónea.

Aplicaciones exitosas de RNA en ingeniería han demostrado que pueden ser utilizadas como un método alternativo de resolución de problemas por su eficiencia y exactitud. Sus características inherentes, tales como su capacidad de aprender, generalizar y procesar en paralelo, las han transformado en la herramienta ideal para muchas aplicaciones.

ANEXOS

Simulador de Redes Neuronales Artificiales

MATLAB: Neural Network Toolbox.

MATLAB maneja una gama de tareas computacionales que se usan en la ingeniería y la ciencia, de la adquisición de datos y del análisis, al desarrollo. El ambiente de MATLAB integra procesador matemático, la visualización, y un lenguaje técnico de gran alcance. Las interfaces incorporadas le permiten tener acceso rápidamente e importar datos a los instrumentos, de los archivos, y de las bases de datos y de los programas externos. Además, MATLAB le deja integrar las rutinas externas escritas en C, C++, el FORTRAN, y Java con sus usos de MATLAB.

MATLAB es un sistema interactivo con un elemento de datos básico, es un arsenal que no requiere el dimensioning. Esto permite que usted solucione muchos problemas técnicos, especialmente éstos con formulaciones de matriz y de vector, en una fracción del tiempo que tomaría para escribir un programa en una lengua no interactivo escalar tal como C o FORTRAN.

MATLAB ofrece una familia de las soluciones de aplicación específica (application-specific) llamadas las cajas de herramientas. Muy importante para la mayoría de los usuarios de MATLAB, las cajas de herramientas permite que usted aprenda y que aplique tecnología especializada. Las cajas de herramientas son colecciones comprensivas de las funciones de MATLAB (M-archivos) que extienden el ambiente de MATLAB para solucionar clases particulares de problemas. Las áreas en las cuales las cajas de herramientas están disponibles incluyen el proceso de señal, sistemas de control, redes neuronales, lógica difusa, la simulación, y muchos otras.

Descripción de la caja de herramientas de redes neuronales (Neural Network Toolbox)

La caja de herramientas de redes neuronales extiende el ambiente de aplicación de MATLAB para proporcionar las herramientas para el diseño, la puesta en práctica, la visualización, y la simulación de redes neuronales. Las redes neuronales son herramientas únicamente de gran alcance en usos donde estaría difícil o imposible el análisis formal, por ejemplo el reconocimiento de patrones y la identificación y el control no lineales del sistema. La caja de herramientas de redes neuronales proporciona la ayuda comprensiva para muchos paradigmas probados de la red, así como una interfaz gráfica que permita que usted diseñe y que maneje sus redes. El diseño modular, abierto, y extensible de la caja de herramientas simplifica la creación de funciones y de redes modificadas para requisitos particulares.

Características Dominantes

- Interfaz gráfica (GUI) para crear, entrenar, y simular a sus redes neuronales
- La ayuda para usuario lo más comúnmente posible arquitecturas de red supervisada y no supervisada.

- Un sistema comprensivo de entrenamiento y de funciones de aprendizaje
- Una habitación de los bloques del Simulink, así como la documentación y demostraciones de los usos del sistema de control
- La generación automática de Simulink modelado de los objetos de la red neuronal
- Representación modular de la red, permitiendo un número ilimitado de la entrada que fija capas, e interconexiones de la red
- Funciones pre y de proceso del bias para mejorar el entrenamiento de la red y determinar el funcionamiento de la red
- Rutinas para mejorar la generalización
- Funciones de la visualización para el funcionamiento de la red neuronal

Lista de funciones (Versión 4.0.1)

Funciones del análisis

errsurf - superficie del error de la neurona de entrada

maxlinlr – máxima tarifa del aprendizaje para una capa lineal

Funciones de inicialización de la capa

initnw - función de la inicialización de la capa de Nguyen-Widrow

initwb - función de inicialización de los pesos y del bias.

Funciones de aprendizaje

learncon - función de aprendizaje diagonal de la conciencia

learnkd - función de aprendizaje del gradiente descendente

learnkdm - función de aprendizaje del gradiente con momentum

learnh - función del aprendizaje de Hebb

learnhd - Hebb con la función de aprendizaje descendente

learnis - función de aprendizaje del peso de Instar

learnk - función de aprendizaje Kohonen

learnlv1 - función de aprendizaje LVQ1

learnlv2 - función de aprendizaje LVQ2

learnos - función de aprendizaje de Outstar

learnp - función de aprendizaje del Perceptron

learnpn - función normalizada de aprendizaje del perceptron

learnsom - función de aprendizaje del mapa de auto-organización

learnwh - regla de aprendizaje de Widrow-Hoff

Funciones de la búsqueda de la línea

srchbac - búsqueda de la vuelta hacia atrás

srchbre - interpolación de la combinación de la sección de oro y cuadrática de Brent

srchcha - interpolación cúbica de Charalambous

srchgol - búsqueda de la sección oro

srchhyb - búsqueda del híbrido bisection/cubic

Creación redes

red - crea una red neuronal
newc - crea una capa competitiva
newcf - crea una red conexión hacia adelante backpropagation
newelm - crea una red de Elman backpropagation
newff - crea una red feed-forward backpropagation
newfftd - crea un red feed-forward de entrada retrasada backpropagation
newgrnn - diseña una red neuronal generalizada de regresión
newhop - crea una red recurrente de Hopfield
newlin - crea una capa lineal
newlind - diseña una capa lineal
newlvq - crea una red de cuantización del vector de aprendizaje
newp - crea un perceptron
newpnn - diseña una red probabilística
newrb - diseña una red de base radial
newrbe - diseña una red de base radial exacta
newsom - crea un mapa de auto-organización

Funciones de inicialización de la red

initlay - función de inicialización de la red capa por capa

Funciones del error

mae - función absoluta del error
mse - función cuadrática media del error
msereg - función de regularización de ajuste del error medio
sse - función que ajusta la suma del error

Funciones para graficar

hintonw - gráfico de Hinton de la matriz del peso
hintonwb - gráfico de Hinton de la matriz del peso y del vector bias
plotbr - grafica el funcionamiento de la red del diagrama para el entrenamiento bayesiano de la regularización
plotes - traza una superficie del error de una sola neurona de entrada
plotpc - traza la línea en diagrama del vector de clasificación del perceptron
plotpv - diagrama de los vectores input/target del perceptron
plotep - traza una posición del peso-bias respecto a una superficie de error
plotperf - diagrama del funcionamiento de la red
plotsom - diagrama del mapa de auto-organización
plotv - diagrama de los vectores de origen como líneas
plotvec - diagrama de los vectores con diversos colores

Funciones de entrenamiento

trainb - entrenamiento con regla de aprendizaje de la hornada del peso y del bias
trainbfg - backpropagation quasi-Newton de BFGS
trainbr - regularización bayesiana
trainc - funciones ascendentes de entrenamiento cíclico

traincgb - backpropagation conyugal del gradiente de Powell-Beale
traincgf - backpropagation conyugal del gradiente de Fletcher-Powell
traincgp - backpropagation conyugal del gradiente de Polak-Ribiere
traingd - backpropagation de la pendiente del gradiente
traingdm - pendiente del gradiente con el backpropagation del ímpetu
traingda - pendiente del gradiente con el backpropagation adaptativo del lr
traingdx - pendiente w/momentum y backpropagation adaptativo del gradiente del lr
trainlm - backpropagation de Levenberg-Marquardt
trainoss - backpropagation secante de un solo paso
trainr - funciones ascendentes del entrenamiento w/learning al azar
trainrp - backpropagation resistente (Rprop)
trains - funciones incrementales del entrenamiento w/learning secuencial
trainscg - backpropagation conyugal del gradiente descendente

Funciones de transferencia

compet - función de transferencia competitiva
hardlim - función de transferencia escalon
hardlims - función de transferencia de limite simétrica
logsig - función de transferencia sigmoideal
poslin - función de transferencia lineal positiva
purelin - función de transferencia lineal
radbas - función de transferencia de base radial
satlin - saturación de la función de transferencia lineal
satlins - función de transferencia lineal de saturación simétrica
softmax - función de transferencia máxima suave
tansig - función de transferencia sigmoidea de la tangente hiperbólica
tribas - función de transferencia de base triangular

Simulación redes

sim - simula una red neuronal
init - inicializa una red neuronal
adap - permite que una red neuronal se adapte
train - entrena a una red neuronal
disp - exhibe las características de una red neuronal
exhibit - exhibe el nombre y las características de una variable de la red neuronal

Funciones del peso

dist - función euclidiana de la distancia del peso
dotprod - función del producto de punto del peso
mandist - función de la distancia de Maniatan del peso
negdist - función del producto de punto del peso
normprod - función normalizada del producto de punto del peso

Funciones de inicialización de los pesos y bias

initcon - función de inicialización diagonal de la conciencia
initzero - función de inicialización a cero de weight/bias

randnc - función de inicialización normalizada de la columna del peso
randnr - función de inicialización normalizada de la fila del peso
rands - función de inicialización simétrica al azar de weight/bias

Programas adicionales de funciones aplicadas para el entrenamiento de la red

f_nl.m

```
function Y=f_nl(F,X)

%Función no lineal
%   F: tipo de función    0--> logsig
%                           1--> tansig
%   X: datos de entrada
%   Y: salida de la función

if F== 0
    Y=logsig(X);
elseif F==1
    Y=tansig(X);
else
    Y=[];
end
```

f_prime.m

```
function y=f_prime(F,x)

% Cálculo de derivada de la función no lineal
%   y: derivada de la función con x
%   F: tipo de de la función
%       0 --> logsig
%       1 --> tansig

if F==0
    y=logsig(x).*(1-logsig(x));
elseif F==1
    y=(tansig(x)+1).*(1-tansig(x));
else
    y=[];
end
```

funv.m

```
x=-10:1/10:10
y1=f_prime(1,x)
y=f_prime(0,x)
subplot(2,1,1);plot(y1);
subplot(2,1,2);plot(y);
```

BIBLIOGRAFÍA

- [1] Faúndez Zanuy, Marcos, “*Tratamiento digital de voz e imagen y aplicación a la multimedia*”, 2001, México, D.F., Alfaomega, (p.p.)
- [2] Martín del Brío, Bonifacio, “*Redes neuronales y sistemas difusos*”, 2a. edición, 2002, México, D.F., Alfaomega, (p.p.)
- [3] Sánchez Azuara, Gilberto, “*Notas de fonética y fonología*”, 2a. edición, 2006, México, D.F., Trillas, (p.p.)
- [4] Aceituno Canal, Vicente, “*Seguridad de la información: Expectativas, riesgos y técnicas de protección*”, 2006, México, Limusa, (p.p.)
- [5] Notas de módulo V, *Diplomado de Seguridad Informática*, año 2007, CEM Polanco, DGSCA UNAM.
- [6] Calderón Osorno, José Luis, “*Notas de redes neuronales artificiales, teoría y diseño de redes neuronales artificiales supervisadas*”, Academia de señales y sistemas, Escuela Superior de Cómputo, Instituto Politécnico Nacional.
- [7] http://www.emagister.com/cursos-gratis/frame.cfm?id_centro=43204110021466565570676950524550&id_curso=54902040030768544853546855574554&id_user=29906391011200901664965706957666&id_segmento=3&id_categ=59&id_busqueda=314918
[Fecha de consulta: Oct/2009]
- [8] Implementación de controles posterior a un proceso de reingeniería
Curso gratis creado por Jose María López Hernández. Extraído de:
http://es.tldp.org/Manuales-LuCAS/doc-como-seguridad-fisica/html_out/
[Fecha de consulta: Oct/2009]
- [9] <http://roble.pntic.mec.es/~msanto1/lengua/1sofolet.htm>
[Fecha de consulta: Oct/2009]