UNIVERSIDAD DE ALCALÁ

ESCUELA POLITÉCNICA SUPERIOR

DEPARTAMENTO DE ELECTRÓNICA

Tesis Doctoral

Title of the pfc/tfm/tfg/tesis

Full name of author

Year of the examination

UNIVERSIDAD DE ALCALÁ

ESCUELA POLITÉCNICA SUPERIOR

DEPARTAMENTO DE ELECTRÓNICA

Title of the pfc/tfm/tfg/tesis

Autor

Full name of author

Director

Full name of advisors

Año

Tesis Doctoral

A ...

"Empieza haciendo lo necesario, luego haz lo posible y de pronto empezarás a hacer lo imposible."

Francisco de Asís

Agradecimientos

“Más vale un minuto de ilusión que mil horas de razonamiento”... (cortesía de Roberto Barra)

Resumen

El trabajo realizado...

Palabras clave: Primera, segunda, tercera, cuarta y quinta (máximo de cinco).

Abstract

The work carried out...

Keywords: First, second, third, fourth and fifth (up to a maximum of five).

Índice general

Resumen **ix**

Abstract **xi**

Índice general **xiii**

Índice de figuras **xv**

Índice de tablas **xvii**

Lista de acrónimos **xix**

Lista de símbolos **xxi**

1 Introducción **1**

1.1 Presentación 1

1.2 Motivación y objetivos 1

1.3 Diagrama 1

1.4 Uso de acrónimos 1

1.5 Símbolos 2

1.6 Organización de la memoria 3

2 Estudio teórico **5**

2.1 Introducción 5

2.2 Estado del Arte 5

2.3 Técnicas utilizadas 5

2.4 Conclusiones 5

3 Desarrollo **7**

3.1 Introducción 7

3.2 Desarrollo del sistema de experimentación 7

3.3 Librerías desarrolladas 7

3.4 Conclusiones 7

4 Resultados **9**

4.1 Introducción 9

4.2 Entorno experimental 9

4.2.1 Bases de datos utilizadas 9

4.2.2 Métricas de calidad 9

4.2.3 Estrategia y metodología de experimentación 9

4.3 Resultados experimentales 9

4.4 Conclusiones 12

5 Conclusiones y líneas futuras **13**

5.1 Introducción 13

5.2 Conclusiones 13

5.3 Líneas futuras 13

Bibliografía **15**

A Manual de usuario **17**

A.1 Introducción 17

A.2 Manual 17

B Herramientas y recursos **19**

Índice de figuras

1.1 Clasificación de los objetos para la gramática. 2

2.1 Departmento de Electrónica 6

4.1 Optimal trames number in the training data set 11

Índice de tablas

4.1 Comparativa 10
4.2 Resultados de la Correlacion cruzada 10

Lista de acrónimos

ANN	Artificial Neural Network.
DBN	Dynamic Bayesian Network.
EIR	Emotion Identification Rate.
EMODB	Berlin Database of Emotional Speech.
ES	Emotional Strength.
ETTS	Emotional Text To Speech.
SIR	Speaker Identification Rate.
SOC	System on a Chip.
SQ	Speech Quality.
STRAIGHT	Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrum.
TD-PSOLA	Time Domain Pitch Synchronous OverLap Add.
ÁÉÍÓÚÜÑáéíóúüñ	Long ÁÉÍÓÚÜÑáéíóúüñ.

Lista de símbolos

Å	non-SI unit of length.
Ω	unit of electrical resistance.
$x(t)$	Audio signal.
$x_i(t)$	Audio signal captured at microphone i.

Capítulo 1

Introducción

1.1 Presentación

Este capítulo presenta una introducción al Trabajo Fin de Carrera desarrollado. En este caso incluímos además secciones "tutoriales" en las que explicamos la inclusión de diagramas y el uso de acrónimos y símbolos.

Importante para las tesis: si necesitáis incluir ficheros pdf (los de autorización e informes de los tutores, por ejemplo), esta plantilla lo permite: mirad el book.tex

1.2 Motivación y objetivos

La motivación de este proyecto...

Los objetivos principales de este Trabajo Fin de Carrera son (ejemplo utilizando "enumerate"):

1. Primer objetivo .....
2. Segundo objetivo.....
   (a) Objetivo 2.1....
   (b) Objetivo 2.2....
3. Tercer objetivo....

1.3 Diagrama

En la figura 1.1 se muestra un ejemplo de gráfico generado automáticamente a partir de un fichero `.dia`.

1.4 Uso de acrónimos

El uso del paquete `glossaries` permite definir los acrónimos y el sistema automáticamente gestiona su inclusión completa la primera vez que se usa. Así, si nos referimos a *Emotional Text To Speech (ETTS)*

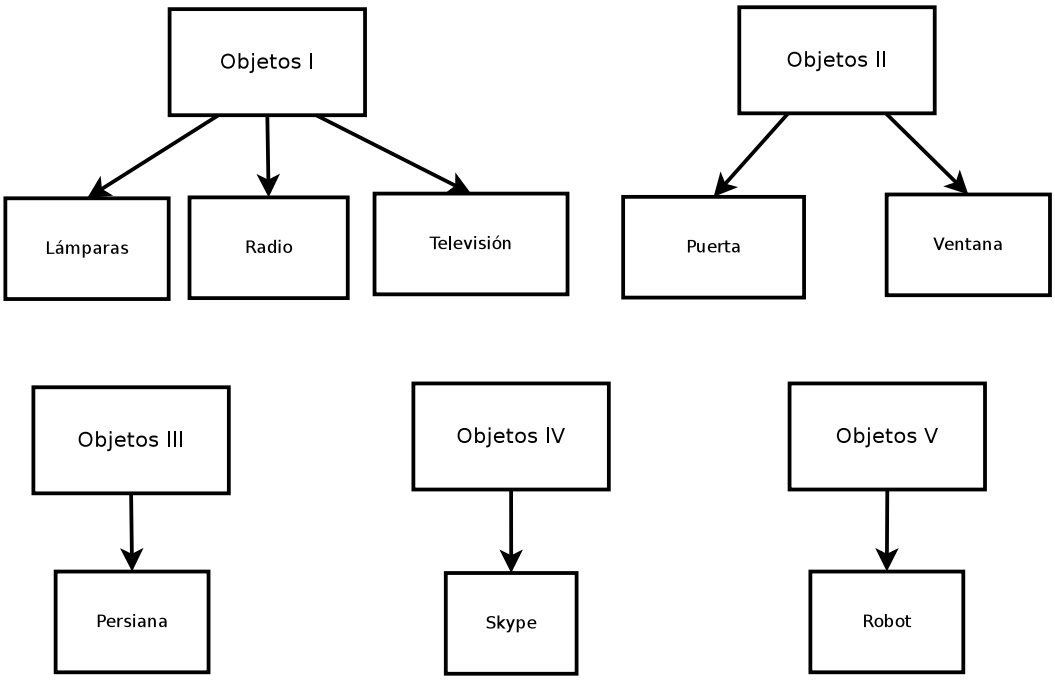


Figura 1.1: Clasificación de los objetos para la gramática.

o bien a *Berlin Database of Emotional Speech (EMODB)*, veremos como aparecen expandidas la primera vez. A partir de ahí, sólo se usará el acrónimo como puede verse al volver a hablar de ETTS y EMODB.

Tiene también soporte para resetear todos los acrónimos como si no estuvieran usados. Vuelvo a incluir el párrafo anterior tras un reset:

El uso del paquete acronym permite definir los acrónimos y el sistema automáticamente gestiona su inclusión completa la primera vez que se usa. Así, si nos referimos a *Emotional Text To Speech (ETTS)* o bien a *Berlin Database of Emotional Speech (EMODB)*, veremos como aparecen expandidas la primera vez. A partir de ahí, sólo se usará el acrónimo como puede verse al volver a hablar de ETTS y EMODB.

Y permite también forzar que se vuelva a citar completo aunque ya se haya utilizado (con el acrónimo entre paréntesis), como puede verse en Emotional Text To Speech (equivalente a Emotional Text To Speech que vale para cualquier glosario), y también a usar forzosamente el acrónimo. Primero reseteamos de nuevo.

Y ahora forzamos el acrónimo: EMODB (eqivalente a EMODB que vale para cualquier glosario). También podemos forzar a que lo ponga todo, con Berlin Database of Emotional Speech (EMODB).

Podemos seguir definiendo entradas de acrónimos, referirnos a *Dynamic Bayesian Network (DBN)* por primera vez, y las siguientes aparecerá como DBN. Pongo ahora el resto de acrónimos *Speech Quality (SQ)*, *Emotion Identification Rate (EIR)*, *Speaker Identification Rate (SIR)* y *Emotional Strength (ES)*. Finalmente los repito para que se vea el efecto: SQ, EIR, SIR y ES.

Y gestiona bien los plurales, ponemos el plural como *Systems on a Chip (SOCs)* la primera vez, y luego la segunda como SOCs. Y podemos volver al singular con SOC.

1.5 Símbolos

El Å se usa en biología estructural, mientras que el Ω se usa en electrónica.

También podemos poner $x(t)$.

$$x(t) \tag{1.1}$$

Y usar la que nos falta $x_i(t)$, también en fórmula:

$$x_i(t) = \sqrt{i} \tag{1.2}$$

Acabamos con un par de acrónimos: *Time Domain Pitch Synchronous OverLap Add (TD-PSOLA)* y *Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrum (STRAIGHT)*. Y con el de acentos *Long ÁÉÍÓÚÜÑáéíóúüñ (ÁÉÍÓÚÜÑáéíóúüñ)* la primera vez, ÁÉÍÓÚÜÑáéíóúüñ la segunda, y luego acrónimo, descripción y completo: ÁÉÍÓÚÜÑáéíóúüñ, Long ÁÉÍÓÚÜÑáéíóúüñ y Long ÁÉÍÓÚÜÑáéíóúüñ (ÁÉÍÓÚÜÑáéíóúüñ).

1.6 Organización de la memoria

Esta memoria se organiza en cuatro grandes capítulos. El primero ...

Capítulo 2

Estudio teórico

2.1 Introducción

En este capítulo se cuenta tal y tal.

El capítulo se estructura en n apartados...

2.2 Estado del Arte

En el estado del arte se enumeran los trabajos más relevantes de otros grupos de investigación. A continuación se muestra un ejemplo del uso de viñetas que nos proporciona ”itemize”:

- En el trabajo .....
- En el siguiente trabajo.....

O citas en un párrafo real: Sin embargo, hay entornos acústicos donde las tasas de error conseguidas son todavía demasiado altas. En concreto, las aplicaciones en las que la captura de la señal de habla se hace usando micrófonos alejados del locutor (típicamente para distancias superiores a un metro) muestran una fuerte sensibilidad a los problemas de reverberación, ruido aditivo y baja relación señal a ruido ([1],[2]). En estos entornos, se ha propuesto el uso de arrays de micrófonos como un método para mejorar la calidad del habla capturada [3][4].

Existen múltiples formas de insertar figuras en Latex. A continuación, se muestra un ejemplo del uso de ”figure”. Como se puede ver en la Figura 2.1 también se pueden poner referencias a las figuras por medio de ”ref” y la etiqueta ”label” de la figura en particular.

2.3 Técnicas utilizadas

Blah, blah, blah.

2.4 Conclusiones

Blah, blah, blah.

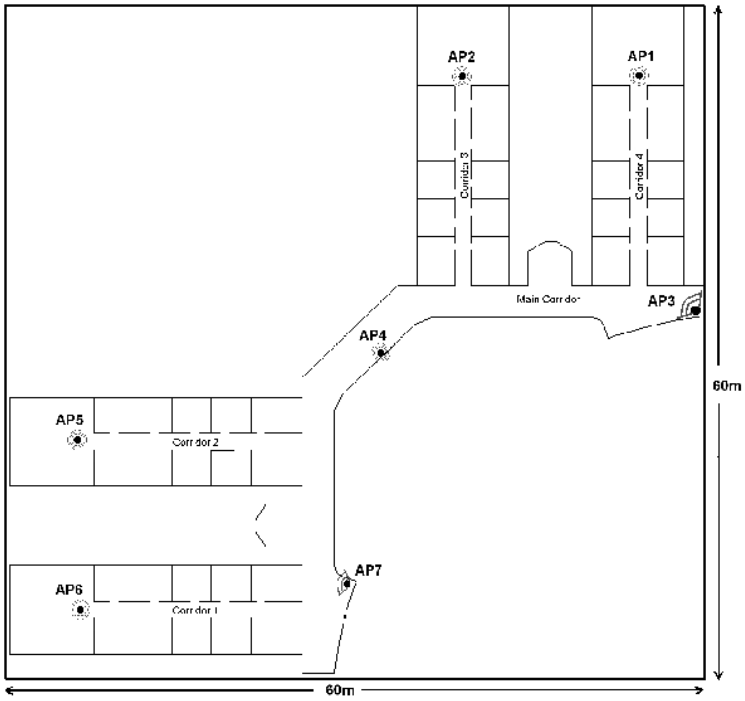


Figura 2.1: Departmento de Electrónica

Capítulo 3

Desarrollo

3.1 Introducción

En este capítulo se incluirá la descripción del desarrollo del trabajo.

El capítulo se estructura en n apartados:...

3.2 Desarrollo del sistema de experimentación

Blah, blah, blah.

3.3 Librerías desarrolladas

También resulta útil poder introducir ecuaciones que se encuentran tanto en línea con el texto $\sigma = 0{,}75$, como en un párrafo aparte (ver ecuación 3.1). Al igual que ocurre con las figuras, también se pueden referenciar las ecuaciones.

$$p[q_t = \sigma_t | q_{t-1} = \sigma_{t-1}] \tag{3.1}$$

3.4 Conclusiones

Capítulo 4

Resultados

4.1 Introducción

En este capítulo se introducirán los resultados más relevantes del trabajo.

La estructura del capítulo es...

4.2 Entorno experimental

Blah, blah, blah.

4.2.1 Bases de datos utilizadas

Blah, blah, blah.

4.2.2 Métricas de calidad

Blah, blah, blah.

4.2.3 Estrategia y metodología de experimentación

Blah, blah, blah.

4.3 Resultados experimentales

A continuación, se muestra un ejemplo de tabla simple (ver tabla 4.1).

Cuando las tablas ocupan más de un página se debe utilizar un tipo especial de tablas denominado ”longtbale”. A continuación, se muestra un ejemplo del mismo (ver tabla 4.2).

Tabla 4.1: Comparativa

Method	Training Time	Man-Work (%)
Propagation model	< 30 sec	5
Manual	9 h 30 min	24
Automatic	2 h	10 8

Tabla 4.2: Resultados de la Correlación cruzada

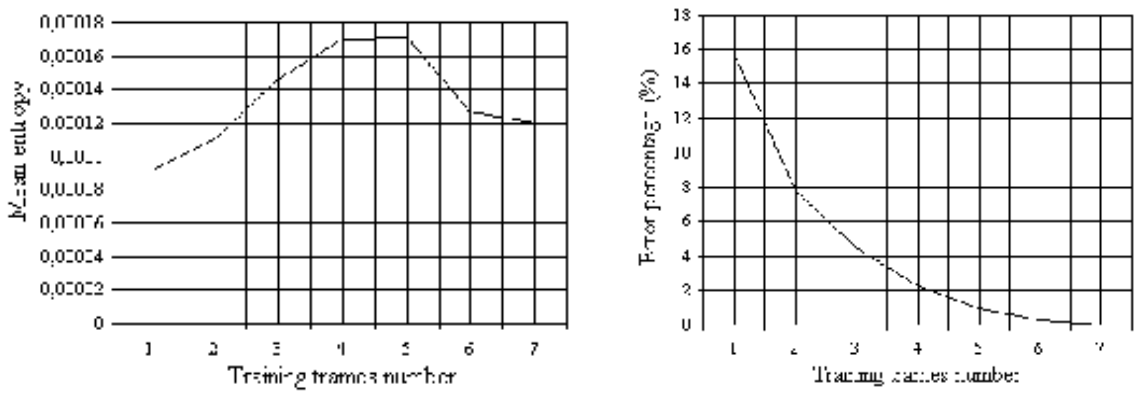
Posición Real	**Posición estimada**	**Coef. Correlación**	**Acierto/Fallo**
2P0	2P0	0,004954	A
2P1	2P4	0,005752	F
2P2	2P2	0,005461	A
2P3	2P0	0,004634	F
2P5	2P4	0,005991	F
2P6	2P16	0,004410	F
2P7	3P9	0,008038	F
2P8	3P9	0,003753	F
2P9	2P7	0,004908	F
2P10	2P10	0,007273	A
2P14	2P16	0,006485	F
2P15	2P15	0,004932	A
2P16	2P16	0,006237	A
2P17	2P15	0,005110	F
2P18	3P18	0,006235	F
2P19	3P18	0,004827	F
2P20	2P20	0,006877	A
2P22	3P18	0,003048	F
2P24	2P24	0,006833	A
2P25	2P25	0,004875	A
2P26	2P31	0,005511	F
2P27	2P28	0,004590	F
2P30	2P31	0,005576	F
2P31	2P31	0,007213	A
2P32	2P35	0,003340	F
2P34	2P34	0,004128	A
2P36	2P35	0,003329	F
2P37	2P37	0,003468	A
2P39	2P38	0,002577	F
2P40	2P43	0,004303	F
2P41	2P41	0,001573	A
2P42	2P41	0,000846	F
2P44	2P44	0,002732	A

Continua en la página siguiente

Tabla 4.2 – continua en la página anterior

Posición Real	Posición estimada	Coef. Correlación	Acierto/Fallo
2P45	23P45	0,001958	F
2P47	2P34	0,002869	F
2P48	2P43	0,004569	F
2P49	3P51	0,001374	F
2P50	2P34	0,002274	F
2P51	2P63	0,003931	F
2P52	2P55	0,003537	F
2P53	3P56	0,003126	F
2P54	2P67	0,005560	F
2P56	2P55	0,002817	F
2P57	2P67	0,006168	F
2P58	2P58	0,005278	A
2P60	3P66	0,004966	F
2P61	3P61	0,004748	A
2P64	2P67	0,005342	F
2P66	2P4	0,004172	F
2P67	2P67	0,005706	A
3P0	3P0	0,003674	A
3P61	2P61	0,003263	F
3P64	2P67	0,003484	F
3P65	2P67	0,002975	F
3P66	2P58	0,005029	F
3P67	3P67	0,003714	A

En algunas ocasiones, también resulta útil emplear el entorno ”subfigure” para añadir múltiples imágenes dentro de la misma figura. A continuación, se muestra un ejemplo del uso en la figura 4.1. También se pueden referenciar las sub-figuras de forma individual, por ejemplo la sub-figura 4.1(b).



(a) Mean entropy

(b) Error Percentage

Figura 4.1: Optimal trames number in the training data set

4.4 Conclusiones

Blah, blah, blah.

Capítulo 5

Conclusiones y líneas futuras

5.1 Introducción

En este apartado se resumen las conclusiones obtenidas y se proponen futuras líneas de investigación que se deriven del trabajo.

La estructura del capítulo es...

5.2 Conclusiones

Para añadir una referencia a un autor, se puede utilizar el paquete ”cite”. En el trabajo [5], se muestra un trabajo...

Y podemos usar de nuevo algún acrónimo, como por ejemplo TD-PSOLA, o uno ya referenciado como *Artificial Neural Network (ANN)*.

5.3 Líneas futuras

Pues eso.

Bibliografía

[1] D. Gelbart and N. Morgan, "Double the trouble: Handling noise and reverberation in far-field automatic speech recognition," in *International Conference on Spoken Language Processing (ICSLP)*, 2002.

[2] S. Kochkin and T. Wickstrom, "Headsets, far field and handheld microphones: Their impact on continuous speech recognition," EMKAY, a division of Knowles Electronics, Tech. Rep., 2002.

[3] M. L. Seltzer, "Microphone array processing for robust speech recognition," Ph.D. dissertation, Carnegie Mellon University, 2003.

[4] W. Herbordt, *Sound capture for human/machine interfaces - Practical aspects of microphone array signal processing.* Springer, Heidelberg, Germany, March 2005.

[5] L. Armani, M. Matassoni, M. Omologo, and P. Svaizer, "Use of a csp-based voice activity detector for distant-talking asr," in *European Conference on Speech Communication and Technology*, 2003, pp. 501–504.

[6] "Información sobre gnu/linux en wikipedia," http://es.wikipedia.org/wiki/GNU/Linux [Último acceso 1/noviembre/2013].

[7] "Página de la aplicación emacs," http://savannah.gnu.org/projects/emacs/ [Último acceso 1/noviembre/2013].

[8] "Página de la aplicación kdevelop," http://www.kdevelop.org [Último acceso 1/noviembre/2013].

[9] L. Lamport, *LaTeX: A Document Preparation System, 2nd edition.* Addison Wesley Professional, 1994.

[10] "Página de la aplicación octave," http://www.octave.org [Último acceso 1/noviembre/2013].

[11] "Página de la aplicación cvs," http://savannah.nongnu.org/projects/cvs/ [Último acceso 1/noviembre/2013].

[12] "Página de la aplicación gcc," http://savannah.gnu.org/projects/gcc/ [Último acceso 1/noviembre/2013].

[13] "Página de la aplicación make," http://savannah.gnu.org/projects/make/ [Último acceso 1/noviembre/2013].

Apéndice A

Manual de usuario

A.1 Introducción

Introducción.

A.2 Manual

Pues eso.

Apéndice B

Herramientas y recursos

Las herramientas necesarias para la elaboración del proyecto han sido:

- PC compatible
- Sistema operativo GNU/Linux [6]
- Entorno de desarrollo Emacs [7]
- Entorno de desarrollo KDevelop [8]
- Procesador de textos LaTeX[9]
- Lenguaje de procesamiento matemático Octave [10]
- Control de versiones CVS [11]
- Compilador C/C++ gcc [12]
- Gestor de compilaciones make [13]