



GRADO

GUÍA DE ESTUDIO DE LA ASIGNATURA TÉCNICAS DE COMPRESIÓN DE DATOS

2ª PARTE | PLAN DE TRABAJO Y ORIENTACIONES PARA SU DESARROLLO



2015-2016

David Fernández Amorós
GRADO EN INGENIERÍA INFORMÁTICA y
GRADO EN INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN

1. PLAN DE TRABAJO

El plan de trabajo de esta asignatura se desarrollará según la tabla de contenidos/actividades que se expone más adelante. Pero antes de entrar en detalles es preciso realizar ciertas consideraciones previas.

La asignatura de Técnicas de compresión de datos requiere del alumno un estudio sistemático y continuado a lo largo del cuatrimestre, dado que se han de asimilar conceptos teóricos y prácticos simultáneamente. Para facilitar su superación es conveniente planificar las etapas de estudio desde el principio, teniendo en cuenta los plazos de entrega y dedicando semanalmente el tiempo necesario, ya que es difícil asimilar la asignatura si se deja el trabajo para el final del cuatrimestre. El objetivo de esta asignatura es que el estudiante adquiera los conocimientos y competencias reflejados en la primera parte de la Guía de Estudio en los 6 ETCS (créditos europeos) que tiene asignados. Un crédito equivale a 25 horas, lo que implica unas 150 horas de estudio y trabajo. Por otro lado, para un desarrollo razonable de este plan de trabajo se considera que el alumno dispone a lo largo del cuatrimestre de 15 semanas efectivas para el estudio y trabajo de la asignatura, lo que equivale a una distribución uniforme de la dedicación de 10 horas semanales a la misma. También es realizable el mismo plan en el supuesto de disponer solo de 13 semanas efectivas, en cuyo caso se requerirá una dedicación media de 11,5 horas semanales. En base a estas diferencias, el plan de trabajo que se presenta contiene la dedicación estimada para cada actividad y sub-actividad en horas, y las horas de dedicación acumuladas a lo largo del cuatrimestre.

Los créditos asignados a cada parte del temario están en consonancia con los contenidos, distribuidos en dos unidades didácticas de tres y cuatro temas respectivamente tal y como se detalla en el programa de la asignatura. Hay un total de cuatro temas de tipo principalmente *descriptivo (D)* y tres de tipo principalmente *práctico (P)*. Esta distinción está presente a lo largo de todo el plan de trabajo, como se aprecia en la mayor dedicación temporal a los temas prácticos que a los descriptivos, o en las diferencias en las actividades propuestas para unos y otros tipos de temas. Esta organización de temas se ilustra en la siguiente tabla.

UNIDAD DIDÁCTICA	TEMA		ACTIVIDAD EVALUABLE
Denominación	Título	Tipo	
UD-I: Conceptos básicos	1. Aproximaciones a la compresión	D	1ª TAREA EVALUABLE (Bloque 1)
	2 Codificación de Huffman	D	
	3. Métodos de diccionario	P	
UD-II: Técnicas avanzadas	4. Codificación aritmética	P	2ª TAREA EVALUABLE (Bloque 2)
	5. Compresión de imágenes	D	
	6. Compresión de audio	D	
	7. Otros métodos	P	

En esta asignatura el material didáctico gira en torno al texto base, que contiene tanto teoría como problemas. El material está orientado a la aplicación práctica de los conceptos teóricos adquiridos durante todo el proceso de estudio. Hay que resaltar que tanto las figuras como los ejemplos que acompañan a gran parte de los conceptos desarrollados en el texto base constituyen elementos fundamentales en el

aprendizaje, y se recomienda al lector que los estudie detalladamente. La correcta asimilación de todo lo anterior permitirá al alumno adquirir las destrezas básicas necesarias para abordar la realización de ejercicios prácticos de mayor complejidad. Es recomendable que intente resolverlos sin ver la solución propuesta. De esta forma ejercitará su capacidad de afrontar problemas y podrá comprobar que puede haber diferentes formas de resolver un mismo ejercicio.

En base a las anteriores premisas en el cuadro siguiente se muestra el cronograma que marca unas pautas adecuadas para que el alumno medio alcance los objetivos al final del curso. Este cronograma incluye los contenidos de cada tema y el plan de actividades, tanto de estudio teórico como de trabajo con ejercicios prácticos, y las actividades de evaluación a realizar.

CRONOGRAMA

CONTENIDOS: CONOCIMIENTOS, ACTITUDES HABILIDADES Y DESTREZAS	ACTIVIDADES ASOCIADAS	HORAS	HORAS ACUMULADAS
Tema 1.- Conceptos básicos. Aproximaciones a la compresión. Códigos de longitud variables. Códigos de rachas. Métodos de diccionario. Transformadas. Cuantización.	Lectura y estudio de los apartados 1.1, 2.2, 1.2, 1.3, 1.4, 3 y 4.4 del texto base o capítulo 1 del texto alternativo.	20	20
Tema 2.- Codificación de Huffman. Decodificación de Huffman. Codificación adaptativa.	Lectura y estudio de los apartados 2.1, 2.2, 2.3, 2.6, 2.8 y 2.9 del texto base o capítulo 2 del texto alternativo.	20	40
Tema 3.- Métodos de diccionario. LZ78, LZW, Zip y Gzip.	Lectura y estudio de los apartados 3.8, 3.12 y 3.23 del libro de texto base o capítulo 3 del texto alternativo.	25	60
Tema 4.- Codificación aritmética. La idea básica. Detalles de implementación. Underflow. Codificación aritmética adaptativa. Codificación de rangos.	Lectura y estudio de los apartados 2.14 y 2.15 del libro de texto base o capítulo 4 del texto alternativo.	25	80
Tema 5.- Compresión de imágenes. Introducción. Aproximación a la compresión de imágenes. Transformadas ortogonales. La transformada discreta del coseno. JPEG.	Lectura y estudio de los apartados 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 y 4.8 libro de texto base o bien el capítulo 5 del texto alternativo.	20	105
Tema 6.- Compresión de audio. Companding. Predicción lineal. Companding Mu-Law y A-Law	Lectura y estudio de los apartados 7.1, 7.2, 7.3, 7.5 y 7.9 del libro de texto base o capítulo 6 del texto alternativo.	20	130
Tema 7.- Otros métodos. El método de Burrows-Wheeler. Clasificación de símbolos. El método SCSU (compresión en Unicode).	Lectura y estudio de los apartados 8.1, 8.2 y 8.12 del libro de texto base o capítulo 7 del texto alternativo.	20	150

2. ORIENTACIONES PARA EL ESTUDIO DE LOS CONTENIDOS

2.1. Introducción

En este apartado se analizan las competencias genéricas y específicas que se deben lograr con esta asignatura, los conocimientos previos precisos para la misma y los materiales necesarios para el estudio. Además, se incluyen algunos consejos sobre la forma de preparar satisfactoriamente la asignatura.

2.2. Contextualización

Dentro del contexto general del Plan de Estudios del Grado en Ingeniería Informática y del Grado en Ingeniería en Tecnologías de la Información, esta asignatura se ubica en la materia denominada "Ingeniería del Software" y ha de contribuir a la consecución de las siguientes competencias genéricas:

- Competencias de gestión y planificación: Iniciativa y motivación. Planificación y organización (establecimiento de objetivos y prioridades, secuenciación y organización del tiempo de realización, etc.). Manejo adecuado del tiempo.
- Competencias cognitivas superiores: selección y manejo adecuado de conocimientos, recursos y estrategias cognitivas de nivel superior apropiados para el afrontamiento y resolución de diversos tipos de tareas/problemas con distinto nivel de complejidad y novedad: Análisis y Síntesis. Aplicación de los conocimientos a la práctica. Resolución de problemas en entornos nuevos o poco conocidos. Pensamiento creativo. Razonamiento crítico. Toma de decisiones.

También ha de contribuir a la consecución de las siguientes competencias específicas:

- Capacidad para la resolución de los problemas matemáticos que pueden plantearse en la ingeniería, en particular de algorítmica, estadística y optimización.
- Capacidad para diseñar, desarrollar, seleccionar y evaluar aplicaciones y sistemas informáticos.
- Conocimiento, diseño y utilización de las estructuras de datos más adecuadas a la resolución de un problema.
- Capacidad para identificar y analizar problemas y diseñar y desarrollar soluciones software sobre la base de un conocimiento adecuado de las teorías, modelos y técnicas actuales.

2.3. Conocimientos previos

El seguimiento de la asignatura requiere la comprensión de algoritmos, principalmente descritos de manera informal. Del mismo modo, el alumno deberá recordar conceptos ya vistos en otras etapas de su formación como nociones básicas de probabilidad y programación.

2.4. Materiales de estudio

El material básico para preparar la asignatura, tanto desde el punto de vista teórico como práctico, es el libro que se indica a continuación.

David Salomon, *Data Compression - The complete reference*. 4th edition. Springer, 2007.

El texto base está en inglés, pero se puede descargar gratis de la página del autor un pdf con la traducción al español, aquí: <http://www.davidsalomon.name/DC4advertis/dataCompression4thesp.pdf>

Como bibliografía alternativa, se recomienda el siguiente libro:

David Salomon, *A concise introduction to data compression*. Springer, 2008.

Como bibliografía complementaria se recomienda el siguiente libro:
Khalid Sayood, *Introduction to data compression*. Fourth edition. Morgan-Kaufmann, 2012.

2.5. Orientaciones generales para preparar la materia

Como se ha explicado con anterioridad, la programación de la asignatura es un punto fundamental para el éxito de su estudio. La materia que configura el programa no es difícil aunque en un primer momento puede parecer demasiado extensa. Para la correcta preparación de la asignatura es conveniente seguir en cada bloque temático el desarrollo que se incluye en el cronograma de la asignatura. No obstante, para obtener los mejores resultados, es recomendable que el estudiante siga los siguientes consejos:

- Realizar una lectura comprensiva de cada tema, profundizando en los contenidos teóricos de los mismos.
- Subrayar aquellos conceptos teóricos más importantes y estudiarlos.
- Al finalizar el estudio de cada unidad, conviene hacer un repaso de aquellos apartados que hayan resultado de especial dificultad.
- Esquematizar las ideas básicas del tema estudiado.
- Resolver los ejercicios de cada capítulo del libro, sin consultar la solución y posteriormente comparar la solución obtenida con la propuesta en el libro.
- Se recomienda reservar un periodo al final del cuatrimestre (una o dos semanas) para hacer un repaso general del temario, encaminado especialmente a la preparación específica del examen.

2.6. Orientaciones generales para afrontar el examen

En el estudio de la asignatura enfocado a la **superación de un examen** es necesario hacer unas consideraciones adicionales, como son el tiempo limitado y escaso del que disponemos (2 horas), y el tipo de examen al que nos enfrentamos.

No obstante, el estudiante junto a la realización de dicho examen (que supondrá el 70% de la nota final) podrá realizar la evaluación continua (que supondrá el 30% de la nota final) y que consistirá en dos actividades más a lo largo del cuatrimestre (véase el punto 3).

El estudiante debe considerar que si bien todos los temas que forman parte del temario son igualmente importantes, hay que tener en cuenta que unos requieren un estudio más profundo que otros.

En cuanto al tipo de examen, su modalidad hace que el estudio se afronte desde una perspectiva diferente que si se tratara de un examen de tipo test. Al ser un **examen tipo desarrollo** debemos tener claros los conceptos fundamentales, pero además es necesario ejercitar la memoria.

Técnicas de compresión de datos

Cuando nos enfrentamos al estudio de cualquier materia, es importante la realización de lecturas rápidas y globales de todo su contenido para hacernos una idea del conjunto y de su contexto. Una vez hecho esto es aconsejable la redacción de resúmenes y esquemas que resalten las ideas básicas y más importantes.

En cuanto a la preparación de la parte práctica del *examen* recomendamos lo siguiente:

- Una vez estudiados y asimilados los conceptos teóricos, siga con detenimiento el desarrollo de los ejemplos que figuran en el libro recomendado.
- Aclarados los conceptos y entendidos los pasos realizados para la resolución de ejemplos, es conveniente que intente resolver los ejercicios resueltos del libro recomendado. No consulte la solución hasta haberlos resuelto por su cuenta.

2.7. Orientaciones concretas para el estudio de los contenidos

Una vez aportadas las orientaciones generales para el estudio de la asignatura es interesante conocer con mayor profundidad cada tema, de forma que su estudio, basado en unas pautas concretas, sea más provechoso para el estudiante.

2.7.1. Unidad Didáctica I: Conceptos Básicos

Tema 1: Aproximaciones a la compresión

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 1.1, 2.2, 1.2, 1.3, 1.4, 3 y 4.4 del texto base. Alternativamente se corresponde con el capítulo 1 de la bibliografía alternativa.

- Resumen

Este tema hace un repaso descriptivo de las principales técnicas de compresión de datos. El término matemático para la compresión de datos es del "codificación de fuentes". El proceso de compresión, es decir, de codificación, se suele realizar en diversas etapas, de forma que se puede considerar que hay varias capas de compresión. En este tema se describen técnicas simples para realizar dichas codificaciones que se combinarán en los siguientes temas para dar paso a algoritmos concretos basados en formatos estandarizados (JPEG, ZIP, etc...).

En la forma más sencilla de codificación, cada elemento del mensaje original es sustituido por un código correspondiente a dicho elemento. Los códigos de longitud fija no son apropiados para la compresión de datos, por lo que uno de los primeros conceptos introducidos es el de códigos de longitud variable. Si la codificación se realiza de forma que a los elementos que ocurren con más frecuencia se les asignan los códigos más cortos, entonces la información codificada ocupará menos espacio que la información original.

Otra forma sencilla de codificación es la codificación por rachas. En muchos contextos es habitual que se repita exactamente el mismo elemento muchas veces, por ejemplo en imágenes fotográficas. En dichos casos, la codificación puede consistir en parejas del tipo "elemento, número de repeticiones".

Los “métodos de diccionario” consisten en la utilización de una estructura de diccionario que almacena habitualmente cadenas de texto que se utiliza para producir una codificación y decodificación del mensaje. Su ventaja es que es un método sencillo pero que produce en general buenos resultados, frente a otros tipos de codificación que requieren un modelado estadístico que puede resultar complejo y/o excesivamente lento.

El concepto de transformadas resulta sencillo conceptualmente: El problema se transforma a otro espacio en el que la compresión resulta más sencilla y posteriormente se deshace la transformación, de forma que la compresión permanezca.

- Esquema

1. Aproximaciones a la compresión
 - 1.1. Códigos de longitud variable
 - 1.2. Códigos de rachas
 - 1.3. Métodos de diccionario
 - 1.4. Transformadas

- Resultados del aprendizaje

Al final del estudio de este tema, el estudiante deberá ser capaz de:

- Entender intuitivamente el concepto de redundancia.
- Entender los conceptos de codificación y decodificación.
- Familiarizarse con los conceptos básicos de la compresión de datos.

- Materiales

Al ser un tema fundamentalmente teórico, para su preparación únicamente será necesaria la consulta de los apartados 1.1, 2.2, 1.2, 1.3, 1.4, 3 y 4.4 del libro recomendado, o bien el capítulo 1 de la bibliografía alternativa.

- Orientaciones para el estudio

Este tema es eminentemente teórico, por lo tanto, la actividad básica prevista para este tema es la lectura y asimilación de los contenidos descritos en el texto.

Tema 2: Codificación de Hufmann

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 2.1, 2.2, 2.3, 2.6, 2.8 y 2.9. Alternativamente, se corresponde con el capítulo 2 de la bibliografía alternativa.

- Resumen

Técnicas de compresión de datos

El análisis probabilístico de los mensajes con el objetivo de maximizar la compresión de datos se corresponde con la disciplina conocida como "Teoría de la información". En este tema se describen algunos conceptos y resultados básicos de la teoría de la información. El concepto central de esta teoría es la entropía, una cantidad relacionada con la cantidad de información presente en un mensaje y también con el máximo grado de compresión que puede alcanzarse. Un método de compresión cuya razón de compresión coincida con la entropía es un método óptimo. El concepto de entropía facilita la comparación entre diversos métodos de codificación de longitud variable. Pese a ser un método teórico (requiere de la construcción de un modelo estadístico, que por fuerza suele ser inexacto) es robusto, en el sentido de que pequeños errores de estimación van a producir pequeñas ineficiencias en la compresión.

El método de codificación de Huffman es un método óptimo de codificación de longitud variable. Conceptualmente es sencillo de explicar. El código es construido fusionando en cada paso los dos elementos restantes con menor probabilidad hasta que solo quede un elemento. Puesto que el método de Huffman requiere conocer la distribución de probabilidad de los elementos y esta estimación puede resultar problemática, existe también una versión adaptativa de la codificación de Huffman.

- Esquema

1. Codificación de Huffman.
 - 1.1. Conceptos de teoría de la información
 - 1.2. Códigos de longitud variable
 - 1.3. Códigos de prefijo.
 - 1.4. La desigualdad de Kraft-MacMillan
 - 1.5. Codificación de Huffman
 - 1.6. Codificación adaptativa de Huffman

- Resultados del aprendizaje

Al final del estudio de este tema, el estudiante deberá ser capaz de:

- Conocer los conceptos de código prefijo, código de decodificación única y código instantáneo.
- Conocer el concepto de entropía y sus propiedades básicas.
- Entender las implicaciones de la desigualdad de Kraft-MacMillan
- Realizar manualmente una codificación/decodificación de Huffman, en su variante simple o adaptativa.

- Materiales

Al ser un tema teórico, para su preparación será necesaria la consulta de los apartados 2.1, 2.2, 2.3, 2.6, 2.8 y 2.9 del libro recomendado y la realización de los ejercicios propuestos en el texto.

- Orientaciones para el estudio

Este tema contiene desarrollos teóricos e introduce conceptos importantes y necesarios para la resolución de ejercicios prácticos. Por tanto se precisa de una lectura profunda y detenida del texto que permita una correcta asimilación de sus contenidos. Para afianzar los conceptos aprendidos es preciso resolver algún ejercicio práctico sobre el tema. Para ello, una vez finalizado el estudio es conveniente que intente resolver los ejercicios. No consulte la solución hasta haberlos resuelto por su cuenta.

Tema 3: Métodos de diccionario

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 3.8, 3.12 y 3.23 del libro recomendado. Alternativamente, se pueden estudiar por el capítulo 3 de la bibliografía alternativa.

- Resumen

Los métodos de diccionario almacenan cadenas de símbolos y realizan su codificación/decodificación en función del contenido de dicho diccionario. Estas estructuras pueden ser estáticas o dinámicas. En los diccionarios estáticos, se puede permitir la inserción de nuevos elementos, pero no el borrado. En los dinámicos se permite tanto la inserción como el borrado. Los métodos de diccionario no necesitan hacer estimaciones de las probabilidades de los símbolos como ocurre con los códigos de longitud variable.

Como es habitual, algoritmos célebres de compresión que usan diccionarios mezclan estas técnicas con otras codificaciones genéricas para mejorar la compresión. Como casos de estudio se analizarán los algoritmos LZ78, LZW, Zip y Gzip.

- Esquema

1. Métodos de diccionario.
 - 1.1. LZ78
 - 1.2. LZW
 - 1.3. Zip
 - 1.4. Gzip

- Resultados del aprendizaje

Al final del estudio de este tema, el estudiante conocerá:

- La implementación de diccionarios, incluyendo el uso de dispersión y pilas.
- El concepto de encadenamiento del diccionario entre el codificador y el decodificador.
- Como combinar distintas técnicas de compresión de datos.

- Materiales

Al ser un tema teórico y práctico, para su preparación será necesaria la consulta de los apartados 3.8, 3.12 y 3.23 libro recomendado, o el capítulo 3 de la bibliografía alternativa y la realización de los ejercicios propuestos en el texto.

- Orientaciones para el estudio

Este tema introduce conceptos importantes y necesarios para la resolución de ejercicios prácticos. Por tanto se precisa de una lectura profunda y detenida del texto que permita una correcta asimilación de sus contenidos. Se recomienda una primera lectura de los principios básicos del mecanismo de codificación

Técnicas de compresión de datos

/decodificación para concentrarse en los detalles de implementación en una segunda lectura. Para afianzar los conceptos aprendidos es preciso resolver algún ejercicio práctico sobre el tema. Para ello, una vez finalizado el estudio es conveniente que intente resolver los ejercicios del libro recomendado. No consulte la solución hasta haberlos resuelto por su cuenta.

Tema 4: Codificación aritmética

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 2.14 y 2.15 del texto recomendado, o bien con el capítulo 5 de la bibliografía alternativa.

- Resumen

La codificación de Huffman es un método óptimo a nivel de símbolo, pero lo normal es que la probabilidad de cada elemento no sea una potencia de n . En ese caso, nos alejamos un poco de compresión ideal con cada elemento que procesamos. Esto puede mejorarse haciendo el "símbolo" más grande (usando un bloque de varios símbolos), pero el problema de fondo sigue estando ahí.

La codificación aritmética sigue otro enfoque. Todo el mensaje se codifica como un número en el intervalo $[0,1)$. Cada símbolo tiene una probabilidad, de forma que cada símbolo está representado con un subintervalo cuya longitud es proporcional a su probabilidad. Cada vez que se añade un símbolo de entrada, el intervalo de trabajo se va reduciendo. El resultado final de la codificación es un número que pertenece al intervalo final.

- Esquema

1. Codificación aritmética
 - 1.1. La idea básica
 - 1.2. Detalles de implementación
 - 1.3. Underflow
 - 1.4. Codificación aritmética adaptativa
 - 1.5. Codificación de rangos

- Resultados del aprendizaje

Al final del estudio de este tema, el estudiante deberá conocer:

- El concepto de reescalado
- El concepto de underflow
- Cómo la codificación aritmética explota la redundancia para producir compresión

- Materiales

Al ser un tema teórico y práctico, para su preparación será necesaria la consulta de los apartados 2.14 y 2.15 del texto recomendado, o bien el capítulo 5 de la bibliografía alternativa y la realización de los ejercicios propuestos en el texto.

- Orientaciones para el estudio

Este tema introduce conceptos importantes y necesarios para la resolución de ejercicios prácticos. Como empieza a ser habitual, el capítulo comienza con una descripción sencilla de los mecanismos de compresión y descompresión, para pasar después a explicar detalles de implementación avanzados que mejoran la eficiencia del proceso. Es importante tener claros los principios básicos del método de compresión y entender que los detalles de implementación pueden ser farragosos y pueden dejarse para una segunda lectura del texto. Para afianzar los conceptos aprendidos es preciso resolver algún ejercicio práctico sobre el tema. Para ello, una vez finalizado el estudio es conveniente que intente resolver los ejercicios propuestos del libro recomendado. No consulte la solución hasta haberlos resuelto por su cuenta.

2.7.3 Unidad Didáctica II: Técnicas avanzadas

Tema 5: Compresión de imágenes

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 y 4.8 del libro recomendado, o alternativamente con el capítulo 5 de la bibliografía alternativa.

- Resumen

La compresión de imágenes requiere en primer lugar una compresión de las bases fisiológicas de la percepción visual humana, especialmente en lo relativo al color, para saber hasta que punto se puede perder información sin que el cerebro perciba una degradación excesiva. La pérdida de calidad en compresión de imagen se produce en un paso de cuantización, que puede ser parametrizado convenientemente.

En un primer momento las imágenes se dividen en cuadrículas, después, cada cuadrícula se procesa por separado. Es un hecho contrastado que un buen predictor del color de una imagen es la media de los valores de los puntos adyacentes. Esta observación lleva de forma natural al concepto de transformada. Puesto que los colores están correlacionados, es posible aplicar una transformación, como la Haar o la transformada discreta del coseno, que produce un primer número grande, y después unos cuantos números pequeños. Tras un paso de cuantización, muchos de estos números se repiten, por lo que pasa a ser útil considerar el uso de códigos de longitud variable o bien códigos de rachas, en función del tipo de imagen. La descompresión debe aplicar las mismas técnicas en orden inverso, terminando con la transformada inversa.

- Esquema

1. Compresión de imágenes
 - 1.1. Introducción
 - 1.2. Aproximación a la compresión de imágenes
 - 1.3. Transformadas ortogonales
 - 1.4. La transformada discreta del coseno
 - 1.5. JPEG

- Resultados del aprendizaje

Al final del estudio de este tema, el estudiante conocerá:

- Los principales tipos de imagen respecto a las posibilidades de compresión
- El concepto de transformada ortogonal
- La transformada discreta del coseno
- Los artefactos que puede introducir la compresión en una imagen

- Materiales

Al ser un tema fundamentalmente teórico, para su preparación únicamente será necesaria la consulta de los apartados 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 y 4.8 del libro recomendado, o alternativamente del capítulo 5 del texto alternativo.

- Orientaciones para el estudio

Este tema es eminentemente teórico, por lo tanto la actividad básica prevista para este tema es la lectura y asimilación de los contenidos descritos en el texto.

Tema 6: Compresión de audio

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 7.1, 7.2, 7.3, 7.5 y 7.9 del libro recomendado, o bien, alternativamente, con el capítulo 6 de la bibliografía alternativa.

- Resumen

El audio es otro de los miembros importantes de la familia de los datos digitales multimedia, y su importancia no ha parado de crecer desde la aparición del mp3 y sus reproductores. Mp3 se convirtió en un estándar internacional en 1992 y desde entonces se han desarrollado e implementado muchos algoritmos diferentes, tanto con pérdida como sin ella.

El companding, también llamado cuantización no lineal, explota el hecho de que el oído humano es muy sensible al volumen bajo, pero pierde sensibilidad a medida que aumenta el volumen. El sistema oído-cerebro es muy sensible y puede percibir sonidos muy bajos, sin embargo, un sonido fuerte reduce la sensibilidad durante un corto periodo de tiempo. Estas propiedades pueden aprovecharse para crear un modelo psicoacústico para identificar las partes de sonido que no serán percibidas debido a dichas limitaciones. Esas partes pueden ser eliminadas o al menos groseramente cuantizadas.

El sonido suele estar fuertemente correlado, por ese motivo es posible calcular una combinación lineal de las últimas muestras y almacenar las diferencias con dichas muestras. Estas diferencias suele ser números pequeños que pueden ser comprimidos eficientemente.

- Esquema

1. Compresión de audio
 - 1.1. Companding
 - 1.2. Predicción lineal

- 1.3. Companding Mu-Law y A-Law
- 1.4. Acortamiento

- Resultados del aprendizaje

Una vez estudiado el contenido de este tema de corte descriptivo, debería:

- Tener una comprensión básica del funcionamiento del sistema oído-cerebro respecto a la percepción humana del sonido.
- Comprender el concepto de cuantización no lineal.
- Comprender los conceptos de predicción lineal, muestreo, tasa de muestreo y tamaño de la muestra.

- Materiales

Al ser un tema fundamentalmente teórico, para su preparación únicamente será necesaria la consulta de los apartados 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 y 4.8 del libro recomendado, o bien el capítulo 6 del texto alternativo.

- Orientaciones para el estudio

Este tema es eminentemente teórico, por lo tanto, al igual que en el tema anterior, la actividad básica prevista para este tema es la lectura y asimilación de los contenidos descritos en el texto.

Tema 7: Otros métodos

- Correspondencia con los capítulos del libro

Este tema se corresponde con los apartados 8.1, 8.2 y 8.12 del libro recomendado o bien con el capítulo 7 de la bibliografía alternativa.

- Resumen

Este tema está dedicado a métodos y técnicas de compresión que no se basan en ninguna de las aproximaciones discutidas previamente en el temario. Los siguientes algoritmos ilustran algunas de estas técnicas originales.

El algoritmo de Burrows-Wheeler (apartado 8.1) empieza con una cadena S de n símbolos y los baraja (e.d. permuta) en otra cadena L que satisface las dos condiciones siguientes:

1. Cualquiera área de L tenderá a tener una concentración de solo unos pocos símbolos.
2. Es posible reconstruir la cadena original S a partir del L . Desde su aparición a principios de los 1990, este inesperado método ha sido objeto de mucha investigación.

La técnica de ranking de símbolos (apartado 8.2) utiliza el contexto, más que la probabilidad, para clasificar símbolos.

El apartado 8.12 describe el algoritmo SCSU para la compresión de documentos basados en Unicode.

- Esquema

1. Otros métodos
 - 1.1. El método de Burrows-Wheeler (BW)
 - 1.2. Ranking de símbolos
 - 1.3. SCSU (compresión en Unicode)

- Resultados del aprendizaje

Una vez estudiado el contenido de este tema de corte descriptivo, debería:

- Conocer el funcionamiento de codificación y decodificación del método BW
- Comprender los conceptos subyacentes a la codificación y decodificación del método de ranking de símbolos
- Conocer el mecanismo de codificación y decodificación del método SCSU

- Materiales

Al ser un tema fundamentalmente teórico, para su preparación únicamente será necesaria la consulta de los apartados 8.1 8.2 y 8.12 del texto base o bien el capítulo 7 del texto alternativo.

- Orientaciones para el estudio

Este tema es eminentemente teórico, por lo tanto, al igual que en el tema anterior, la actividad básica prevista para este tema es la lectura y asimilación de los contenidos descritos en el texto.

3. ORIENTACIONES PARA LA REALIZACIÓN DEL PLAN DE ACTIVIDADES

En el plan de trabajo se han incluido las actividades evaluables que deberá presentar el estudiante de acuerdo a la evaluación continua. Ésta consistirá en la realización de dos Pruebas además del Examen Presencial. Las Pruebas de Evaluación Continua consistirán en dos cuestionarios de evaluación en línea. Las fechas concretas para estas pruebas se publicarán en el curso virtual (una a mediados del cuatrimestre y otra a finales del cuatrimestre) y se gestionarán con las herramientas que esta misma plataforma virtual ofrece para ello.

- Primera Prueba: Tres primeros temas.
- Segunda Prueba: Cuatro últimos temas.

Ambas pruebas tendrán una duración de 2 horas y consistirán en 10 preguntas tipo *test* repartidas entre conceptos teóricos y ejercicios prácticos. El estudiante solo tendrá un intento para realizar la prueba, por tanto no podrá abrirla de su panel de "actividades" del curso virtual, hasta que no esté seguro de que va a realizarla (y siempre dentro del periodo establecido). Una vez que empiece la actividad dispondrá de dos horas para completarla. Las pruebas se califican automáticamente, por tanto, una vez finalizadas el estudiante podrá conocer rápidamente el resultado de la misma. El resultado de estas pruebas (la media de ambas) ponderarán un 30% de la calificación final (siendo el 70% restante el resultado de la prueba presencial).

La prueba presencial constará de dos partes, una de naturaleza teórica y otra de naturaleza práctica. En la parte teórica el alumno deberá contestar a diversas cuestiones sobre la materia objeto de estudio. La parte

práctica consistirá en la resolución de ejercicios prácticos, y se valorará el planteamiento y desarrollo que el alumno haga de los mismos. Todas las preguntas pueden contestarse con las explicaciones del libro recomendado como bibliografía básica. Como se ha indicado, para superar la asignatura no será necesario alcanzar una calificación mínima en esta prueba. En la prueba presencial el alumno no podrá consultar ningún material.

En el cálculo de la Nota Final de la Asignatura (NFA) a partir de la Nota de las Pruebas de Evaluación Continua (NPEC) y de la Nota de la Prueba Presencial (NPP) se tendrán en cuenta los siguientes criterios:

- El peso de la Prueba Presencial en la Nota Final será del 70%.
- El peso de las Pruebas de Evaluación Continua en la Nota Final será del 30%.
- No será necesario obtener una Nota mínima ni en las Pruebas de Evaluación Continua ni en la Prueba Presencial para poder aprobar la asignatura.
- La Nota Final de la Asignatura requerida para aprobarla deberá ser mayor o igual a 5 puntos.
- No será necesaria la presencia del alumno en el Centro Asociado para realizar ninguna de las Pruebas de Evaluación Continua.
- Las Pruebas de Evaluación Continua sólo se realizarán y evaluarán a lo largo del primer cuatrimestre en el que se imparte la asignatura. En la convocatoria de septiembre, se mantendrá la nota obtenida en dichas actividades.

En definitiva, la expresión para el cálculo de la Nota Final de la asignatura tanto para la convocatoria de junio como para la de septiembre, es la siguiente:

$$NFA = 0,7 \cdot NPP + 0,3 \cdot NPEC$$

y para aprobar la asignatura el alumno deberá obtener una NFA ≥ 5 .

4.- GLOSARIO E ÍNDICE ANALÍTICO

El texto base de esta asignatura cuenta con un glosario de términos habituales en compresión de datos y un índice analítico de términos organizado por orden alfabético. En él encontrará la referencia a dónde se encuentra definido cada término y en qué tema o temas se hace uso del mismo.