

Характеристика дисциплины «Цифровая обработка аудио/видео данных»

Цель преподавания дисциплины. Целью преподавания дисциплины является обучение магистров алгоритмическим основам кодирования аудио-речевых сообщений, изображений; методологии проектирования и применения цифровых кодеров в системах мультимедиа.

Задачи изучения дисциплины. Содержание курса включает как теоретические так и практические основы цифровой обработки аудио-видео данных на базе многоступенчатой дискретизации сигнала, а также обучение основным приложениям цифровой обработки сигналов в системах мультимедиа. Дисциплина базируется на разделах курсов "Теория и применение цифровой обработки сигналов", "Системы обработки медиаданных". Содержание курса должно найти отражение в диссертационной работе.

Перечень дисциплин, усвоение которых необходимо для изучения данной дисциплины.

<i>№ пп</i>	<i>Название дисциплины</i>	<i>Раздел, тема</i>
1.	Теория и применение цифровой обработки сигналов	<i>Все разделы курса</i>
2.	Системы обработки медиаданных	<i>Все разделы курса</i>

СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Дается наименование разделов и тем, их содержание, объемы в часах лекционных, практических (семинарских) и лабораторных занятий, курсовой проект (работа) и его характеристика. Указанные сведения раскрывают содержание типовой (базовой) учебной программы с учетом результатов развития науки, техники, культуры и производства за период, который прошел со времени ее утверждения.

Определяются цели и задачи каждой темы, осуществляется их распределение по неделям и семестрам с учетом дневной, вечерней и заочной форм обучения.

1. Название тем лекционных занятий, их содержание, объем в часах.

Пример:

<i>№ пп</i>	<i>Название темы</i>	<i>Содержание</i>	<i>Объем в часах</i>
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Раздел 1. Системы обработки аудио/речевых сигналов			
1.	Цифровые модели	Голосовой тракт, излучение,	2

	речевых сигналов	возбуждение, полная модель. Цифровое представление речевых сигналов. Дискретизация и квантование речевого сигнала. Общая теория разностного квантования: кодер, декодер. Оценка соотношения сигнал-шум. Тип предсказателя, синтез линейного предсказателя.	
2.	Кодирование речевого сигнала	Информационная структура речевого сигнала. Кодирование речевого сигнала. Цель проектирования речевого кодера. Общая схема речевого кодера. Элементы системы кодирования речи	2
3.	Устранение статистической и перцептуальной избыточности	Минимизация дисперсии внесенных искажений. Мера пологости спектра сигнала. Соотношение сигнал-шум и сигнал-порог маскирования. Оптимальное распределение бит по каналам банка фильтров	2
4.	Перцептуальное кодирование аудио- и речевых сигналов	Обобщенная структура перцептуального кодера; анализ и синтез; элементы психоакустики: абсолютный порог восприятия, критическая шкала частот, расчет порогов маскирования. Функция Джонстона – перцептуальная энтропия. Частотно-временной анализ	2
5.	Кодеры речевого сигнала на основе синусоидальной модели	Анализ и синтез речевого сигнала. Кодеры речевого сигнала на базе метода декомпозиции речевого сигнала на периодическую и шумовую компоненты: структура кодера: выделение гармонических составляющих частоты основного тона	2

		(периодической компоненты), анализ шумовой компоненты речевого сигнала; структура декодера: синтез речевого сигнала	
6.	Выбор банка цифровых фильтров перцептуального аудио кодера	Полифазная структура банка ДПФ фильтров. Расчет фильтра-прототипа. Кратковременное скачущее преобразование Фурье (STFT – short time Fourier transform). Пакет дискретного вэйвлет преобразования (ПДВП). Дерево ПДВП, согласованное с критической шкалой частот	6
7.	Структуры аудио кодеров	Стандарт MPEG. Субполосные аудио кодеры: декомпозиция сигнала на основе вэйвлет преобразования, адаптивный пакет вэйвлет преобразования. Аудио кодеры на базе гибридных моделей: синусоидальной + вэйвлет преобразование. Измерение качества реконструированного сигнала аудио кодеров	4
<i>Итого: за семестр 1</i>			20
Раздел 2. Системы обработки изображений			
8.	Формирование цифровых изображений	Методы получения цифровых изображений. Типы изображений. Растр. Связность. Пакеты прикладных программ обработки изображений (Матлаб, Фотошоп и др.)	2
9.	Оценка качества цифровых изображений	Критерии визуального качества изображений. Аналитические критерии анализа качества изображений. Среднеквадратичный критерий. Отношение сигнал-шум. Критерий резкости изображения. Оценка погрешности цифрового	2

		представления изображения	
10.	Точечные методы цифровой обработки изображений	Арифметические алгоритмы преобразования яркости. Гистограммные методы преобразования изображений. Изменение контраста изображений	2
11.	Локальные методы цифровой обработки изображений	Двумерная свертка. Методы сглаживания изображений. Линейная и нелинейная фильтрация изображений. Алгоритмы выделения краев. Изменение контраста изображений	2
12.	Глобальные методы цифровой обработки изображений	Дискретное преобразование Фурье (Быстрое преобразование Фурье). Его применение для обработки изображений. Основные принципы WDFТ. Обратное преобразование WDFТ. Оценка ошибки реконструкции WDFТ-синтеза. Вейвлет преобразование изображений	4
13.	Методы сжатия цифровых изображений	Цепной код. Статистическое кодирование изображений. Метод дифференциального кодирования. Методы кодирования с преобразованиями. Метод адаптивной выборки. Методы сжатия изображения JPEG и JPEG2000	4
14.	Практическое применение методов обработки изображений	Коррекция неравномерной засветки изображения. Устранение шумов на цифровом изображении. Детекция на изображении предметов заданной формы с помощью преобразования Хафа. Технология повышение контраста изображения. Аффинные преобразования изображений: поворот,	2

		масштабирование	
<i>Итого: за семестр 2</i>			<i>18</i>
<i>Всего за учебный год</i>			<i>38</i>

2. Перечень тем семинарских занятий, их содержание и объем в часах

<i>№ пп</i>	<i>Название темы</i>	<i>Содержание</i>	<i>Объем в часах</i>
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Раздел 1. Системы обработки аудио/речевых сигналов			
1.	Методы выделения частоты основного тона	Корреляционный метод Сондхи	2
2.	Методы классификации речи на вокализованную и невокализованную	Энергетические, корреляционные, вероятностные	2
3.	Методы построения перцептуальных вокодеров	Частотно-временной анализ, определение энтропии, энтропия Джонстона	2
4.	Векторное квантование параметров модели речеобразования вокодера на базе линейного предсказания	Определение LSF, построение кодовых книг, метод К-средних	2
5.	Синусоидальный кодер речи	Анализатор, синтезатор гармоник	2
6.	Декомпозиция речевого сигнала на периодическую и шумовую компоненты	Метод на основе ДПФ, синхронизированного с частотой основного тона	2
7.	Субполосный вокодер на основе пакета дискретного вэйвлет преобразования (ПДВП)	Аудио кодер на основе ПДВП, структура дерева которого согласована с критическими полосами частот. Аудио кодер, реализуемый на адаптивном ПДВП	4
8.	Оценка качества реконструированного речевого сигнала в вокодере	Субъективная оценка качества реконструированного речевого сигнала. Объективное измерение предсказанного субъективного качества речевых кодеров	4
<i>Итого: за семестр 1</i>			<i>20</i>

<i>Раздел 2. Системы обработки изображений</i>			
9.	Яркостная коррекция цифровых изображений	<i>Основные этапы коррекции яркости</i>	2
10.	Оптимизация палитры цветных изображений	<i>Основы теории цвета, определения, методы оптимизации</i>	2
11.	Линейные методы фильтрации изображений	<i>2Д-ДПФ, 2Д вейвлет преобразование</i>	4
12.	Сжатие статических изображений	Основные понятия: кодовая избыточность, межэлементная избыточность, психофизическая избыточность. Модели сжатия изображений	2
13.	Сжатие динамических изображений	Основные принципы кодирования. Описание методов и подходов.	4
14.	Стандарт кодирования видеоизображений MPEG-4 AVC/H.264	Многокадровое предсказание. Сжатие макроблоков без потерь. Квантование. Энтропийное кодирование квантованных коэффициентов трансформации.	4
<i>Итого: за семестр 2</i>			<i>18</i>
<i>Всего за учебный год</i>			<i>38</i>

3. Реферат

Целью подготовки реферата является проверка полученных знаний в области обработки аудио- и речевых сигналов, а также решения конкретных задач обработки изображений.

Темы :

1. Субполосное кодирование аудио сигнала на основе ДПФ-модулированного неравнополосного цифрового банка фильтров: схема перцептуального кодирования
2. Субполосное кодирование видео данных на основе ДПФ с неравномерным частотным разрешением
3. Методы детекции на цифровой фотографии изображения объекта заданного типа (лицо человека, самолет, дорожная сеть, буквы и т.п.)

4. Биометрические методы идентификации человека.
5. Реконструкция размытых изображений

4. ЛИТЕРАТУРА

4.1.ОСНОВНАЯ

- 4.1.1 Назаров М.В., Прохоров Ю.Н. Методы цифровой обработки и передачи речевых сигналов. М., Радио и связь, 1985.
- 4.1.2 Прохоров Ю.Н. Статистические модели и рекуррентное предсказание речевых сигналов. М., Радио и связь, 1984.
- 4.1.3 Сапожков М.А., Михайлов В.Г. Вокодерная связь. М., Радио и связь, 1983.
- 4.1.4 Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи. М., Радио и связь, 2000.
- 4.1.5 Kondoz A.M. Digital speech: coding for low bit rate communications systems. NJ, J.Wiley & Sons, 1996
- 4.1.6 Speech coding and synthesis// Edited by W.B.Kleijn, K.K.Paliwal. Amsterdam, Elsevier, 1998.
- 4.1.7 Advances in speech signal processing // edited by Sadaoki Furui, M.Mohan Sondhi. – Marcel Dekker, Inc., NY, 1992. – 871p.
- 4.1.8 E. Zwicker and H. Fastl, Psychoacoustics Facts and Models. Berlin, Germany: Springer-Verlag, 1990
- 4.1.9 P.P. Vaidyanathan, Multirate Systems and Filter Banks. Englewood Cliffs, NJ: Prentice-Hall, 1993
- 4.1.10 G. Strang, T. Nguyen. Wavelets and filter banks. Wellesley-Cambridge Press, 1996.- 491 p.
- 4.1.11 M. Vetterli, J. Kovacevic, Wavelets and Subband coding. Englewood Cliffs, NJ, Prentice-Hall, 1995. – 488 p.
- 4.1.12 J.G. Ackenhusen. Real-Time Signal Processing: Design and Implementation of Signal Processing Systems. – Prentice Hall, NJ, 1999. – 460 p.

- 4.1.13 Петровский А.А., Парфенюк М., Барович А., Лившиц М.З. Цифровые банки фильтров: анализ, синтез и применение в мультимедиа системах / Учебно-методическое пособие, Минск, БГУИР, 2006. 82с.
- 4.1.14 Рабинер Л., Шафер Р. Цифровая обработка речевых сигналов. М., Радио и связь. 1981.
- 4.1.15 Маркел Дж., Грэй А. Линейное предсказание речи. М., Связь, 1980.
- 4.1.16 Ярославский Л.П. Введение в цифровую обработку изображений. – М.: Сов. радио, 1979. – 312 с.
- 4.1.17 Прэтт У. Цифровая обработка изображений – М.: Мир, 1982. – 790с.
- 4.1.18 Методы компьютерной обработки изображений / Под ред. В.А. Соифера.- М.: ФИЗМАТЛИТ, 2003.- 784с.
- 4.1.19 A.Piotrowski, M.Parfieniuk. Cyfrowe banki filtrow: analiza, synteza I implementacja dla systemow multimedialnych. – Wydawnictwo Politechniki Bialostockiej, Bialystok, 2006. – 389s.
- 4.1.20 Анализаторы речевых и звуковых сигналов: методы, алгоритмы и практика (с МАТЛАВ примерами) // под редакцией д.т.н. профессора Петровского А.А. – Минск: Бестпринт, 2009 -454с.

4.2. ДОПОЛНИТЕЛЬНАЯ

- 4.2.1 С.Марпл-мл. Цифровой спектральный анализ и его приложения. М., Мир, 1990.
- 4.2.2 Barnwell T.P., Nayebi K., Richardson C.H. Speech coding: a computer laboratory textbook, NJ, J.Wiley & Sons, 1996
- 4.2.3 R.E. Crochiere and L.R. Rabiner, Multirate Digital Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1983
- 4.2.4 В.В. Витязев. Цифровая частотная селекция // М.: Радио и связь, 1993. – 240 с.
- 4.2.5 Петровский А.А., Петровский Ал.А., Лихачев Д.С. Речевые интерфейсы ЭВС. / Методическое пособие, Минск БГУИР, 2004. – 65с.

- 4.2.6 Применение цифровой обработки сигналов // Под редакцией Э.Оппенгейма. М., Мир, 1980, - 2.2.6. Психоакустические факторы, с.60-66; - 3. Цифровая обработка речевых сигналов, с. 137-191.
- 4.2.7 И.М.Журавель, И.М. Краткий курс теории обработки изображений [Электронный ресурс].– Режим доступа: <http://forest.akadem.ru/library/matlab/imageprocess/book2/index.html>
Дата доступа: 29.01.2008.
- 4.2.8 Бьемон Ж., Лагендейк Р.Л., Марсеро Р.М. Итерационные методы улучшения изображений // ТИИЭР. – 1990. – Т.78, № 5. – С. 58–84.
- 4.2.9 Старовойтов В.В. Локальные геометрические методы цифровой обработки и анализа изображений.- Минск: Ин-т техн. Кибернетики НАН Беларуси, 1997.- 284с.