

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**ИО директора физтех-школы
радиотехники и компьютерных
технологий**

Д.А. Гаврилов

Рабочая программа дисциплины (модуля)

по дисциплине:	Методы сжатия данных
по направлению:	Прикладные математика и физика
профиль подготовки:	Радиотехника и компьютерные технологии Физтех-школа Радиотехники и Компьютерных Технологий кафедра инфокоммуникационных систем и сетей
курс:	4
квалификация:	бакалавр

Семестр, формы промежуточной аттестации: 7 (осенний) - Дифференцированный зачет

Аудиторных часов: 30 всего, в том числе:

лекции: 30 час.

семинары: 0 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 15 час.

Всего часов: 45, всего зач. ед.: 1

Программу составил: К.В. Семенков, канд. физ.-мат. наук, старший преподаватель

Программа обсуждена на заседании кафедры инфокоммуникационных систем и сетей 15.02.2022

Аннотация

Данный курс предназначен для получения студентами знаний к научно-исследовательской деятельности, практических навыков и компетенции в области фундаментальных и прикладных проблем кодирования и сжатия информации, таких как сложность и энтропия информации; методы и алгоритмы кодирования и сжатия данных, изображений, звука и видео.

Изучение дисциплины позволит студентам овладеть необходимыми компетенциями и знаниями по современным подходам к сжатию данных и решать практические задачи, связанные со сжатием данных.

Курс проводится в формате лекционных занятий и выполнения самостоятельных работ с дополнительной литературой.

1. Цели и задачи

Цель дисциплины

Целью освоения дисциплины является формирование у студентов системы знаний в области сжатия информации для дальнейшего применения на практике методов и алгоритмов сжатия.

Задачи дисциплины

- формирование базовых знаний в области теории кодирования источников как части теории информации;
- обучение студентов основным методам сжатия данных без потерь и с потерями;
- ознакомление студентов с современными методами сжатия изображения, видео и звука.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1 Анализирует задачу, выделяя этапы ее решения, действия по решению задачи
	УК-1.2 Находит, критически анализирует и выбирает информацию, необходимую для решения поставленной задачи
	УК-1.3 Рассматривает различные варианты решения задачи, оценивает их преимущества и недостатки
	УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки
	УК-1.5 Определяет и оценивает практические последствия возможных вариантов решения задачи
ОПК-1 Способен применять фундаментальные знания, полученные в области физико-математических и (или) естественных наук, и использовать их в профессиональной деятельности	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.2 Способен строить математические модели, производить количественные расчеты и оценки
	ОПК-1.3 Способен определять границы применимости полученных результатов
ОПК-4 Способен осуществлять сбор и обработку научно-технической и (или) технологической информации для решения фундаментальных и прикладных задач	ОПК-4.1 Владеет методами научного поиска и интеллектуального анализа информации при решении задач профессиональной деятельности
	ОПК-4.4 Владеет навыками работы с компьютером и компьютерными сетями с целью получения, хранения и обработки научной (технической, технологической) информации
ПК-4 Способен критически оценивать применимость используемых методик и методов	ПК-4.1 Знает численные порядки величин, характерных для соответствующей профессиональной области
	ПК-4.3 Способен обосновать причинно-следственные отношения используемых понятий и моделей

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные понятия теории информации, применяющиеся для кодирования дискретных источников;
- математические методы построения оптимальных кодов;
- математические методы, лежащие в основе сжатия данных;
- основные методы сжатия данных;
- основные методы сжатия изображений;
- основные методы сжатия звука;
- основные методы сжатия видео.

уметь:

- эффективно использовать на практике теоретические компоненты науки: понятия, суждения, умозаключения, законы;
- понимать принципы работы современных средств сжатия данных;
- работать на современном экспериментальном оборудовании;
- применять математические методы сжатия данных в самостоятельной работе и исследованиях.

владеть:

- современными методами сжатия данных;
- навыками самостоятельной работы в лаборатории на современном экспериментальном оборудовании;
- математическими методами сжатия данных.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Введение в теорию информации.	2			1
2	Виды методов сжатия. Базовые методы сжатия. Статистические методы сжатия.	4			2
3	Методы контекстного моделирования.	4			2
4	Словарные методы.	4			2
5	Сжатие изображений.	8			4
6	Сжатие видео.	4			2
7	Сжатие звука.	4			2
Итого часов		30			15
Подготовка к экзамену		0 час.			
Общая трудоёмкость		45 час., 1 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 7 (Осенний)

1. Введение в теорию информации.

Основные понятия. Шенноновская энтропия и ее основные свойства. Энтропия пары и условная энтропия. Колмогоровская сложность и ее основные свойства. Теорема Колмогорова-Соломонова. Связь сложности и энтропии. Кодирование информации.

2. Виды методов сжатия. Базовые методы сжатия. Статистические методы сжатия.

Кодирование длин серий. Алгоритм «стопка книг» (Move-to-Front). Модель источника без памяти. Префиксные коды. Неравенство Крафта. Коды Хаффмана. Арифметическое кодирование. Адаптивные коды Хаффмана и арифметические коды.

3. Методы контекстного моделирования.

Методы контекстного моделирования. Марковские источники. Алгоритм PPM.

4. Словарные методы.

Введение в словарные методы. Алгоритмы Зива-Лемпела. (LZ77, LZ78, LZW). Формат deflate. Примеры других словарных алгоритмов: LZMA, RAR, PNG.

5. Сжатие изображений.

Подходы к сжатию изображений. Алгоритмы сжатия без потерь. Сжатие с потерями. Методы преобразования, используемые при сжатии изображений (дискретное косинусное преобразование, преобразование Уолша-Адамара, преобразование Хаара). Прогрессирующее сжатие изображений. Алгоритм JPEG. Вейвлетные методы. Стандарт JPEG 2000.

6. Сжатие видео.

Основные стандарты видео. Подходы к сжатию видео. Субоптимальные алгоритмы поиска. Семейство стандартов MPEG. Принципы кодирования видео в стандартах MPEG-1, MPEG-2, MPEG-4. Принципы кодирования в стандарте H.261.

7. Сжатие звука.

Особенности оцифровки звука и восприятия звука человеком. Подходы к сжатию звука с потерями. Сжатие звука в стандарте MPEG-1.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Аудиторная доска. Компьютер, проектор, экран.

6. Перечень рекомендуемой литературы

Основная литература

1. Сжатие данных, изображений и звука [Текст] : учеб. пособие для вузов / Д. Сэломон ; пер. с англ. В. В. Чепыжова .— М. : Техносфера, 2004 .— 368 с.
2. Универсальное кодирование. Теория и алгоритмы [Текст]/Ю. М. Штарьков , -М., Физматлит, 2013
3. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: Диалог-МИФИ, 2003 – 384 с
4. Кудряшов Б.Д. Теория информации: Учебник для вузов. – СПб.: Питер, 2009 – 320 с.

Дополнительная литература

1. Работы по теории информации и кибернетике [Текст] : [учеб. пособие для вузов] / К. Э. Шеннон ; пер. с англ. под ред. Р. Л. Добрушина, О. Б. Лупанова ; с предисл. А. Н. Колмогорова .— М. : Изд-во иностр. лит., 1963 .— 829 с.

2. Курс теории информации [Текст] / В. Д. Колесник, Г. Ш. Полтырев .— М. : Наука, 1982 .— 416 с.
3. Колмогоровская сложность и алгоритмическая случайность [Текст]/Н. К. Верещагин, В. А. Успенский, А. Шень, -М., МЦНМО, 2013
4. Десять лекций по вейвлетам [Текст] / И. Добешигпер. с англ. Е. В. Мищенко ; под ред. А. П. Петухова - М. ; ИжевскРегулярная и хаотическая динамика,2004, 2001
5. Колмогоров А.Н. Три подхода к определению понятия «количество информации» — Проблемы передачи информации, 1965, том 1, выпуск 1, 3-11.

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

1. Журналы по теории информации, информационным технологиям, доступные через Internet.
2. Специализированные ресурсы, посвященные сжатию информации: <http://www.compression.ru>, <http://hutter1.net/>
3. Научные и научно-технические журналы: <http://scitation.aip.org/>, <http://www.sciencemag.org/>
4. Электронные конспекты лекций, учебные пособия и сборники задач, разработанные для данного курса.

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На лекционных занятиях используются мультимедийные технологии, включая демонстрацию презентаций (MS Office Power Point).

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике. В результате изучения дисциплины студент должен знать основные определения, понятия. Успешное освоение курса требует напряжённой самостоятельной работы студента. В программе курса приведено минимально необходимое время для работы студента над темой. Самостоятельная работа включает в себя:

- чтение и конспектирование рекомендованной литературы;
- проработку учебного материала (по конспектам лекций, учебной и научной литературе);
- подготовку к дифференцированному зачету.

Руководство и контроль за самостоятельной работой студента осуществляется в форме индивидуальных консультаций.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к преподавателю.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Прикладные математика и физика
профиль подготовки:	Радиотехника и компьютерные технологии Физтех-школа Радиотехники и Компьютерных Технологий кафедра инфокоммуникационных систем и сетей
курс:	4
квалификация:	бакалавр
Семестр, формы промежуточной аттестации: 7 (осенний) - Дифференцированный зачет	
Разработчик:	К.В. Семенков, канд. физ.-мат. наук, старший преподаватель

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1 Анализирует задачу, выделяя этапы ее решения, действия по решению задачи
	УК-1.2 Находит, критически анализирует и выбирает информацию, необходимую для решения поставленной задачи
	УК-1.3 Рассматривает различные варианты решения задачи, оценивает их преимущества и недостатки
	УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки
	УК-1.5 Определяет и оценивает практические последствия возможных вариантов решения задачи
ОПК-1 Способен применять фундаментальные знания, полученные в области физико-математических и (или) естественных наук, и использовать их в профессиональной деятельности	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.2 Способен строить математические модели, производить количественные расчеты и оценки
	ОПК-1.3 Способен определять границы применимости полученных результатов
ОПК-4 Способен осуществлять сбор и обработку научно-технической и (или) технологической информации для решения фундаментальных и прикладных задач	ОПК-4.1 Владеет методами научного поиска и интеллектуального анализа информации при решении задач профессиональной деятельности
	ОПК-4.4 Владеет навыками работы с компьютером и компьютерными сетями с целью получения, хранения и обработки научной (технической, технологической) информации
ПК-4 Способен критически оценивать применимость используемых методик и методов	ПК-4.1 Знает численные порядки величин, характерных для соответствующей профессиональной области
	ПК-4.3 Способен обосновать причинно-следственные отношения используемых понятий и моделей

2. Показатели оценивания компетенций

В результате изучения дисциплины «Методы сжатия данных» обучающийся должен:

знать:

- основные понятия теории информации, применяющиеся для кодирования дискретных источников;
- математические методы построения оптимальных кодов;
- математические методы, лежащие в основе сжатия данных;
- основные методы сжатия данных;
- основные методы сжатия изображений;
- основные методы сжатия звука;
- основные методы сжатия видео.

уметь:

- эффективно использовать на практике теоретические компоненты науки: понятия, суждения, умозаключения, законы;
- понимать принципы работы современных средств сжатия данных;
- работать на современном экспериментальном оборудовании;
- применять математические методы сжатия данных в самостоятельной работе и исследованиях.

владеть:

- современными методами сжатия данных;
- навыками самостоятельной работы в лаборатории на современном экспериментальном оборудовании;
- математическими методами сжатия данных.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

С целью контроля освоения обучающимися учебного материала проводится устный опрос в начале занятия по теме прошлой лекции или в конце занятия по пройденной теме.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

Перечень контрольных вопросов для сдачи дифференцированного зачета:

1. Дайте определение собственной информации и шенноновской энтропии и сформулируйте основные свойства энтропии.
2. Дайте определение колмогоровской сложности.
3. Сформулируйте и докажите теорему Колмогорова-Соломонова.
4. Приведите и докажите неравенства, связывающие колмогоровскую сложность и энтропию.
5. Дайте определение кодирования информации. Приведите примеры кодов источника.
6. Дайте определение префиксных кодов. Приведите пример префиксного кода. Сформулируйте и докажите неравенство Крафта.
7. Сформулируйте и докажите прямую теорему побуквенного кодирования.
8. Сформулируйте и докажите обратную теорему побуквенного кодирования.
9. Дайте определение однозначно декодируемого кода, префиксного кода, кодов Хаффмана. Опишите алгоритм Хаффмана.
10. Опишите принципы работы алгоритма арифметического кодирования. Постройте арифметический код для простого примера.
11. Опишите принципы нормализации, применяемые при реализации алгоритма арифметического кодирования.
12. Опишите адаптивные алгоритмы Хаффмана и арифметического кодирования.
13. Опишите метод кодирования длин серий. Приведите пример.
14. Опишите метод «стопка книг» (Move-to-Front). Приведите пример.
15. Опишите принципы словарных методов кодирования. Приведите пример кодирования со статическим словарем.
16. Опишите алгоритм кодирования Зива-Лемпела LZ77.
17. Опишите алгоритм кодирования Зива-Лемпела LZW.
18. Опишите принципы работы алгоритма PPM.
19. Опишите один из механизмов вторичной оценки ухода (SEE).
20. Кратко опишите 2-3 метода субоптимального поиска, применяемых при сжатии видео.
21. Опишите методы подавления пауз и уплотнения, применяемые при сжатии звука с потерями.
22. Опишите методы сжатия звука в стандарте MPEG-1.

Критерии оценивания

Оценка отлично (10) выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (9) выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (8) выставляется студенту, показавшему систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо (7) выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо (6) выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо (5) выставляется студенту, если он в основном знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно (4) выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка удовлетворительно (3) выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже в стандартной ситуации.

Оценка неудовлетворительно (2) выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

Оценка неудовлетворительно (1) выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Дифференцированный зачет проводится в устной форме.

При проведении дифференцированного зачета обучающемуся предоставляется 20 минут на подготовку. Опрос обучающегося проводится в течение 30 минут.

Во время проведения дифференцированного зачета обучающиеся могут пользоваться программой дисциплины.