

Минобрнауки России

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)**

УТВЕРЖДАЮ



Заведующий кафедрой
Борисов Дмитрий Николаевич
Кафедра информационных систем

03.05.2023

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.В.06 Системы сжатия данных

1. Код и наименование направления подготовки/специальности:

09.03.02 Информационные системы и технологии

2. Профиль подготовки/специализация:

Информационные системы и сетевые технологии

3. Квалификация (степень) выпускника:

Бакалавриат

4. Форма обучения:

Очная

5. Кафедра, отвечающая за реализацию дисциплины:

Кафедра информационных систем

6. Составители программы:

Сычев Александр Васильевич, кандидат физико-математических наук, доцент кафедры информационных систем

7. Рекомендована:

НМС ФКН 03.05.2023, протокол № 7

8. Учебный год:

2026-2027

9. Цели и задачи учебной дисциплины:

Целью курса является формирование теоретических знаний о современных методах сжатия данных и формирование практических навыков применения полученных знаний при разработке и оценке информационных систем.

Задачи курса:

- изучение широкого спектра методов и алгоритмов сжатия с потерями и без потерь для текстов, двоичных данных, аудио- и видеоданных, образующих фундамент современных технологий мультимедиа и телекоммуникационных технологий;
- формирование практических навыков применения алгоритмов сжатия применительно к конкретным видам данным (тексты, аналоговые сигналы и др.).
- формирование практических навыков оценки алгоритмов сжатия для конкретных видов данных.

10. Место учебной дисциплины в структуре ООП:

Дисциплина относится к части блока дисциплин Б1, формируемой участниками образовательных отношений

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников) и индикаторами их достижения:

Код и название компетенции	Код и название индикатора компетенции	Знания, умения, навыки
ПК-4 Способен проводить анализ требований к программному обеспечению, выполнять работы по проектированию программного обеспечения	ПК-4.4 Описывает технологии обработки данных для возможности их использования в программной среде, включая вопросы параллельной обработки	Знает: современные методы и алгоритмы сжатия с потерями и без потерь для текстов, двоичных данных, аудио- и видеоданных, образующих фундамент современных технологий мультимедиа и телекоммуникационных технологий Умеет: применять алгоритмы сжатия применительно к конкретным видам данным (тексты, аналоговые сигналы и др.).
ПК-4 Способен проводить анализ требований к программному обеспечению, выполнять работы по проектированию программного обеспечения	ПК-4.5 Описывает применяемые математические методы и алгоритмы функционирования для компонентов программных средств	Умеет: обосновывать выбор метода кодирования входных и выходных данных для алгоритмов сжатия данных Владеет: методами оценки эффективности алгоритмов сжатия для конкретных видов данных

12. Объем дисциплины в зачетных единицах/час:

3/108

Форма промежуточной аттестации:

Зачет с оценкой, Контрольная работа

13. Трудоемкость по видам учебной работы

Вид учебной работы	Семестр 8	Всего
Аудиторные занятия	72	72
Лекционные занятия	36	36
Практические занятия		0
Лабораторные занятия	36	36
Самостоятельная работа	36	36
Курсовая работа		0
Промежуточная аттестация	0	0
Часы на контроль		0

Вид учебной работы	Семестр 8	Всего
Всего	108	108

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
1	Введение	Основные понятия и определения. Карта групп методов сжатия. Базовые стратегии сжатия. Сравнение алгоритмов по степени сжатия. Тестовые наборы CalgCC, CantCC и др.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
2	Кодирование источников данных без памяти	Разделение мантисс и экспонент. Канонический алгоритм Хаффмана. Нумерирующее кодирование. Векторное квантование.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
3	Кодирование аналоговых сигналов	Кодирование источников данных типа "аналоговый сигнал". Линейно-предсказывающее кодирование. Две модели поведения сигнала: эволюционная и шумовая. Одномерный и двумерный случаи. Характеристики алгоритма. Идея субполосного кодирования.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
4	Словарные методы сжатия	Словарные методы сжатия. Статистика распределения строк различной длины в тексте на естественном языке. Алгоритм Зива-Лемпела LZ77. Алгоритм Зива-Лемпела LZ78. Модификации алгоритмов LZ: LZSS, LZW. Характеристики алгоритмов LZ77 и LZ78. Архиваторы, использующие LZ.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
5	Контекстное моделирование	Методы контекстного моделирования. Основные понятия и концепции. Классификация стратегий контекстного моделирования. Алгоритм PPM. Достоинства и недостатки алгоритма. Характеристики алгоритма.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
6	Преобразование BWT	Преобразование Барроуза-Уиллера (BWT). Частичное сортирующее преобразование. Методы, используемые совместно с BWT: перемещение стопки книг, кодирование длин повторов (RLE), кодирование расстояний. Выбор метода сжатия.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
7	Выбор метода сжатия.	Особенности методов сжатия качественных данных. Основные типы данных. Критерии оценки методов сжатия. Рекомендации по выбору методу сжатия.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
8	Сжатие изображений	Особенности изображений как типа данных. Классы изображений и классы приложений, работающих с изображениями. Алгоритмы сжатия изображений. Критерии сравнения алгоритмов сжатия изображений.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
9	Методы сжатия изображений без потерь	Сжатие изображений без потерь: RLE, LZW.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
10	Методы сжатия изображений с потерями	Сжатие изображений с потерями. Алгоритм JPEG, фрактальный и вейвлетный алгоритмы. JPEG2000.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)
11	Сжатие видеоданных	Сжатие видеоданных. Основные понятия. Требования приложений к алгоритмам сжатия видео. Базовые технологии сжатия видеоданных. Основные стандарты сжатия видео: MPEG-1,2,4,7.	онлайн курс, размещенный в Электронном университете ВГУ (edu.vsu.ru)

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела)	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
1	Введение	2			1	3

№ п/п	Наименование темы (раздела)	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
2	Кодирование источников данных без памяти	4		6	3	13
3	Кодирование аналоговых сигналов	4			2	6
4	Словарные методы сжатия	5		6	6	17
5	Контекстное моделирование	4		6	6	16
6	Преобразование ВВТ	4		6	6	16
7	Выбор метода сжатия	1			2	3
8	Сжатие изображений	2			2	4
9	Методы сжатия изображений без потерь	2		2	2	6
10	Методы сжатия изображений с потерями	4		6	2	12
11	Сжатие видеоданных	4		4	4	12
		36	0	36	36	108

14. Методические указания для обучающихся по освоению дисциплины

1) При изучении дисциплины рекомендуется использовать следующие средства:

- рекомендуемую основную и дополнительную литературу;
- методические указания и пособия;
- контрольные задания для закрепления теоретического материала;
- электронные версии учебников и методических указаний для выполнения лабораторных работ.

2) Для лучшего усвоения дисциплины рекомендуется проведение письменного опроса (тестирование, решение задач) студентов по материалам лекций и практических работ. Подборка вопросов для тестирования осуществляется на основе изученного теоретического материала.

3) При проведении лабораторных занятий обеспечивается практическая демонстрация материалов лекционных занятий и осуществляется экспериментальная проверка методов, алгоритмов и технологий обработки данных, излагаемых в рамках лекций.

4) При переходе на дистанционный режим обучения для создания электронных курсов, чтения лекций онлайн и проведения лабораторно- практических занятий используются информационные ресурсы образовательного портала "Электронный университет ВГУ (<https://edu.vsu.ru>), базирующегося на системе дистанционного обучения Moodle, развернутой в университете

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины

№ п/п	Источник
1	Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин [идр.]. – М. : ДИАЛОГ-МИФИ, 2002. – 384 с.
2	Катунин, Г. П. Основы мультимедийных технологий : учебное пособие / Г. П. Катунин. — Санкт-Петербург : Лань, 2021. — 784 с. — ISBN 978-5-8114-2736-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/169093
3	Березкин, Е. Ф. Основы теории информации и кодирования : учебное пособие / Е. Ф. Березкин. — 3-е изд., стер. — Санкт-Петербург : Лань, 2019. — 320 с. — ISBN 978-5-8114-4119-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/115524

б) дополнительная литература:

№ п/п	Источник
1	Тропченко А.Ю. Методы сжатия изображений, аудиосигналов и видео [Электронный ресурс] : учебное пособие / А.Ю. Тропченко, А.А. Тропченко. — Электрон. дан. — Спб. : НИУ ИТМО (Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики), 2009. — 109 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=43541
2	Уэлстид С. Фракталы и вейвлеты для сжатия изображений в действии / С. Уэлстид. – М. : Триумф, 2003. — 319 с.
3	Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон. — М. : Техносфера, 2004. – 365 с.
4	Дворкович, В. П. Цифровые видеoinформационные системы. Теория и практика / В. П. Дворкович, А. В. Дворкович. — Москва : Техносфера, 2012. — 1008 с. — ISBN 978-5-94836-336-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/73516

в) информационные электронно-образовательные ресурсы:

№ п/п	Источник
1	Электронный курс, размещенный на портале Электронный университет ВГУ (https://edu.vsu.ru/course/view.php?id=4151)
2	www.lib.vsu.ru - ЗНБ ВГУ

16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1	Электронный курс, размещенный на портале Электронный университет ВГУ (https://edu.vsu.ru/course/view.php?id=4151)

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение):

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств:

Промежуточная аттестация может включать в себя проверку теоретических вопросов, а также, при необходимости (в случае не выполнения в течение семестра), проверку выполнения практических заданий, позволяющих оценить уровень полученных знаний и/или практическое (ие) задание(я), позволяющее (ие) оценить степень сформированности умений и навыков.

Для оценки теоретических знаний используется перечень контрольно-измерительных материалов. Каждый контрольно-измерительный материал для проведения промежуточной аттестации включает два задания - вопросов для контроля знаний, умений и владений в рамках оценки уровня сформированности компетенции. При оценивании используется количественная шкала.

Для оценивания результатов обучения используется 4-балльная шкала: «отлично», «хорошо», «удовлетворительно», «незачтено». Соотношение показателей, критериев и шкалы оценивания результатов обучения представлено в следующей таблице.

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Обучающийся в полной мере владеет теоретическими основами методов сжатия данных, способен иллюстрировать ответ примерами, фактами, применять теоретические знания для решения практических задач, что соответствует полному освоению компетенций.	Повышенный уровень	Отлично
Обучающийся владеет теоретическими основами методов сжатия данных, способен иллюстрировать ответ примерами, допускает незначительные ошибки, неточности, которые исправлены после замечания преподавателя, что соответствует не достаточно полному освоению компетенций.	Базовый уровень	Хорошо

<p>Ответ на контрольно-измерительный материал неполный, без обоснований, объяснений. Демонстрирует частичные знания учебного материала, значительные затруднения в вопросах проведения анализа, что показывает недостаточное владение компетенциями. Ошибки устраняются по дополнительным вопросам преподавателя.</p>	<p>Пороговый уровень</p>	<p>Удовлетворительно</p>
<p>Ответ на контрольно-измерительный материал фрагментарный. Обучающийся демонстрирует несистематические, отрывочные знания, допускает грубые, принципиальные ошибки, которые не устранены после дополнительных наводящих вопросов, что соответствует не освоению компетенций.</p>	<p>-</p>	<p>Незачтено</p>

18. Материально-техническое обеспечение дисциплины:

Компьютерная лаборатория с локальной сетью из 15 персональных компьютеров с установленным системным и прикладным программным обеспечением и выходом в Интернет.

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Разделы дисциплины (модули)	Код компетенции	Код индикатора	Оценочные средства для текущей аттестации
1	Разделы 1-6, 8-11	ПК-4	ПК-4.4	Письменный опрос, контрольная работа
2	Раздел 7	ПК-4	ПК-4.5	Письменный опрос, контрольная работа
3				

Промежуточная аттестация

Форма контроля - Зачет с оценкой, Контрольная работа

Оценочные средства для промежуточной аттестации

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств:

Промежуточная аттестация включает в себя проверку теоретических вопросов, а также, при необходимости (в случае не выполнения в течение семестра), проверку выполнения практических заданий, позволяющих оценить уровень полученных знаний и/или практическое (ие) задание(я), позволяющее (ие) оценить степень сформированности умений и навыков.

Для оценки теоретических знаний используется перечень контрольно-измерительных материалов. Каждый контрольно-измерительный материал для проведения промежуточной аттестации включает два задания - вопросов для контроля знаний, умений и владений в рамках оценки уровня сформированности компетенции. При оценивании используется количественная шкала.

Для оценивания результатов обучения используется 4-балльная шкала: «отлично», «хорошо», «удовлетворительно», «незачтено». Соотношение показателей, критериев и шкалы оценивания

результатов обучения представлено в следующей таблице.

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Обучающийся в полной мере владеет теоретическими основами методов сжатия данных, способен иллюстрировать ответ примерами, фактами, применять теоретические знания для решения практических задач, что соответствует полному освоению компетенций.	Повышенный уровень	Отлично
Обучающийся владеет теоретическими основами методов сжатия данных, способен иллюстрировать ответ примерами, допускает незначительные ошибки, неточности, которые исправлены после замечания преподавателя, что соответствует не достаточно полному освоению компетенций.	Базовый уровень	Хорошо
Ответ на контрольно-измерительный материал неполный, без обоснований, объяснений. Демонстрирует частичные знания учебного материала, значительные затруднения в вопросах проведения анализа, что показывает недостаточное владение компетенциями. Ошибки устраняются по дополнительным вопросам преподавателя.	Пороговый уровень	Удовлетворительно
Ответ на контрольно-измерительный материал фрагментарный. Обучающийся демонстрирует несистематические, отрывочные знания, допускает грубые, принципиальные ошибки, которые не устранены после дополнительных наводящих вопросов, что соответствует не освоению компетенций.	-	Незачтено

20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

20.1 Текущий контроль успеваемости

Тестовые задания - 1 балл за каждый правильный тест (максимум).

Компетенция ПК-4

1. Выберите правильное соответствие для базовых стратегий сжатия:

- a) Описание поступающих данных через уже обработанные. Никаких вероятностей не вычисляется.
 - b) Вычисление вероятностей для поступающих данных на основании статистики по уже обработанным данным.
 - c) Отдельно кодируется и добавляется к сжатому блоку его статистика.
 - d) Входящие данные разбиваются на блоки, которые затем трансформируются целиком.
- a) Адаптивная статистическая стратегия.
 - b) Блочная статистическая стратегия.
 - c) Преобразование блока.
 - d) Преобразование потока.

2. Выберите правильное соответствие для методов сжатия:

- a) Цель метода - сжатие потока R-битовых элементов в предположении, что значение каждого из

них отличается от значений соседних элементов незначительно. Основная идея состоит в том, чтобы формировать два потока: для каждой пары $S[2i]$, $S[2i+1]$ сохранять полусумму $(S[2i] + S[2i+1])/2$ и разность $(S[2i] - S[2i+1])$.

b) Цель - сжатие потока R-битовых элементов в предположении, что значение каждого из них является линейной комбинацией значений h предыдущих элементов. Основная идея состоит в том, чтобы в формируемый поток записывать ошибки предсказаний: разности между реальными и предсказанными значениями.

c) Цель преобразование потока R-битовых элементов, такое, чтобы в формируемом выходном потоке оставалось не более чем N значений.

d) Цель - сжатие блока R-битовых элементов в предположении, что у него есть одна важная характеристика C , которую выгодно сжимать отдельно от остальных. Основная идея состоит в том, чтобы формировать два блока: сохраняющий самую важную характеристику C и содержащий остальные данные D - так, чтобы все комбинации D были почти равновероятны. Далее обрабатывать эти блоки отдельно.

e) Идея состоит в замене строк символов на такие коды, что их можно рассматривать как индексы строк (фраз) некоторого словаря. При декодировании осуществляется обратная замена индекса на соответствующую ему фразу словаря.

- a) Векторное квантование
- b) Линейно-предсказывающее кодирование
- c) Нумерирующее кодирование
- d) Словарные методы сжатия
- e) Субполосное кодирование

Правильное соответствие:

3. Выберите правильное соответствие для методов сжатия изображений

- a) RLE.
- b) LZW.
- c) Алгоритм Хаффмана.
- d) JPEG.
- e) Фрактальный алгоритм.
- f) Вейвлетный алгоритм.

- a) Сжатие без потерь .
- b) Сжатие без потерь .
- c) Сжатие с потерями .
- d) Сжатие с потерями .
- e) Сжатие без потерь .
- f) Сжатие с потерями .

4. Выберите правильное соответствие :

- a) Информация в цифровом виде.
- b) Конечная последовательность цифровых данных.
- c) Последовательность с неизвестными границами.
- d) Последовательность данных с произвольным доступом.
- e) Последовательность данных с последовательным доступом.
- f) Описание при котором блок занимает меньше битов чем исходный с возможностью однозначного восстановления исходного блока (с точностью до бита).

- a) Блок.
- b) Блок.
- c) Данные.
- d) Поток.

- e) Поток.
- f) Сжатие блока.

Правильное соответствие:

5. Укажите, какие ключевые требования предъявляются приложениями к алгоритмам сжатия видео:

- a) Аудиовизуальная синхронизация
- b) Быстрый поиск вперед/назад
- c) Возможность длительного хранения на цифровых носителях
- d) Время кодирования/декодирования
- e) Защита от несанкционированного доступа
- f) Масштабируемость
- g) Небольшая стоимость аппаратной реализации
- h) Поддержка нескольких языковых дорожек
- i) Поддержка субтитров
- j) Показ кадров фильма в обратном направлении
- k) Произвольный доступ
- l) Редактируемость
- m) Устойчивость к ошибкам

6. Укажите, какие ключевые требования предъявляются приложениями к алгоритмам сжатия видео:

- a) Произвольный доступ.
- b) Быстрый поиск вперед/назад.
- c) Показ кадров фильма в обратном направлении.
- d) Аудиовизуальная синхронизация.
- e) Устойчивость к ошибкам.
- f) Время кодирования/декодирования.
- g) Редактируемость.
- h) Масштабируемость.
- i) Небольшая стоимость аппаратной реализации.
- j) Поддержка субтитров.
- k) Поддержка нескольких языковых дорожек.
- l) Защита от несанкционированного доступа.
- m) Возможность длительного хранения на цифровых носителях.

7. Выберите правильные утверждения, относящиеся к линейно-предсказывающему кодированию.

- a) Целью является сжатие потока R-битовых элементов в предположении, что значение каждого из них является линейной комбинацией значений h предыдущих элементов.
- b) Основная идея метода состоит в том, чтоб в формируемый поток записывать ошибки предсказаний: разность между реальными и предсказанными значениями.
- c) Методы этой группы являются трансформирующими и поточными.
- d) Идея метода состоит в замене строк символов на такие коды, что их можно трактовать как индекс строки некоторого словаря.
- e) В рамках данного метода решается задача оценки вероятностей появления символов в каждой позиции обрабатываемой последовательности в зависимости о непосредственно ему предшествующих.

8. Выберите правильное соответствие для методов сжатия изображений.

- a) RLE.
- b) LZW.
- c) Алгоритм Хаффмана.

- d) JPEG.
- e) Фрактальный алгоритм.
- f) Вейвлетный алгоритм.
- a) Сжатие с потерями.
- b) Сжатие без потерь.
- c) Сжатие без потерь.
- d) Сжатие с потерями.
- e) Сжатие с потерями.
- f) Сжатие без потерь.

Вопросы с кратким (вычисляемым) ответом - 1 балл за каждый правильный тест (максимум)

Компетенция ПК-4

1. Источник сообщений порождает сообщения на основе алфавита символов s_1, s_2, s_3, s_4, s_5 с соответствующими вероятностями $p_1=0.4; p_2=0.2; p_3=0.2; p_4=0.1; p_5=0.1$.

С использованием алгоритма Хаффмана был построен следующий префиксный двоичный код для указанных выше символов источника:

00
10
11
010
011

Рассчитайте среднюю длину для построенного двоичного кода в расчете на один символ алфавита источника.

Ответ следует привести с точностью до одной цифры после запятой.

2. В результате прямого преобразования Барроуза-Уиллера (BWT) над блоком символов был получен следующий код:

(ммзехиаен; 6). Требуется выполнить обратное преобразование BWT и записать исходный блок символов.

3. Выполните декодирование блока данных методом LZ78:

1 C
1 O
1 N
1 D
1 I
1 T
6 O
1 _
1 S
6 N
1 E
9 Q
1 U
1 A
9 N
3 N

4. Восстановите исходный блок символов на основе его арифметического кода:

н 0.3333333
е 0.3333333

с 0.1666667

д 0.1666667

с = 0.892490

Вопросы с развернутым ответом.

Критерии оценивания развернутого ответа:

В зависимости от степени полноты ответа можно получить до 3 баллов. За первую часть вопроса можно получить до 1,5 баллов, за вторую - до 1,5 баллов.

Компетенция ПК-4

1. Объясните, в чем состоит принципиальная разница между методами сжатия с потерями и методами сжатия без потерь? Приведите примеры таких методов.

20.2 Промежуточная аттестация

Промежуточная аттестация включает в себя проверку теоретических вопросов, а также, при необходимости (в случае не выполнения в течение семестра), проверку выполнения практических заданий, позволяющих оценить уровень полученных знаний и/или практическое (ие) задание(я), позволяющее (ие) оценить степень сформированности умений и навыков.