

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

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

	Рабочая программа дисциплины (модуля)
по дисциплине:	Технологии синтеза и распознавание речи
по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра когнитивных технологий
курс:	1
квалификация:	магистр

Семестр, формы промежуточной аттестации: 1 (осенний) - Дифференцированный зачет

Аудиторных часов: 60 всего, в том числе:

лекции: 0 час.

семинары: 60 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 75 час.

Всего часов: 135, всего зач. ед.: 3

Количество контрольных работ, заданий: 2

Программу составил: Д.С. Богданов, канд. техн. наук

Программа обсуждена на заседании кафедры когнитивных технологий 02.04.2024

Аннотация

Курс "Технологии синтеза и распознавания речи" предназначен для формирования у студента знаний и навыков в области обработки речевых сигналов. В ходе этого курса студенты изучают различные аспекты синтеза и распознавания речи, такие как:

1. Основы речевой обработки: Вводятся основные понятия и методы обработки речевых сигналов, включая представление речи, акустическую модель и лингвистическую модель.
2. Синтез речи: Изучаются методы синтеза речи, которые позволяют создавать искусственную речь на основе текстовых данных. В рамках курса студенты могут изучить различные подходы к синтезу речи, такие как конкатенативный синтез и синтез на основе скрытых моделей Маркова (НММ).
3. Распознавание речи: Рассматриваются методы распознавания речи, которые позволяют компьютеру преобразовывать аудиосигналы в текстовую форму. Студенты могут изучить различные алгоритмы распознавания речи, включая скрытые модели Маркова (НММ) и нейронные сети.
4. Обработка естественного языка: Изучаются методы обработки естественного языка, которые позволяют компьютеру понимать и генерировать текст на естественном языке. В рамках курса студенты могут изучить различные задачи обработки естественного языка, такие как морфологический анализ, синтаксический анализ и семантический анализ.
5. Практические задания: Курс может включать практические задания, которые помогут студентам применить изученные методы на практике. Это может включать разработку систем синтеза и распознавания речи, анализ речевых данных или эксперименты с алгоритмами обработки речи.

Курс "Технологии синтеза и распознавания речи" поможет студентам развить навыки работы с речевыми данными, а также понять основные принципы и методы в этой области.

1. Цели и задачи

Цель дисциплины

Дать достаточно полноценное представление о проблемах, методах, алгоритмах и современных инструментальных средствах, используемых при создании современных систем автоматического синтеза и распознавания речи.

Задачи дисциплины

- Выработка у студентов понимания важности постоянного внимания к эффективности алгоритмов, используемых при программировании и познакомить их с методами, которые могут использоваться для достижения эффективности;
- подготовка к участию в научных семинарах, научно-технических конференциях и симпозиумах, составлению научных обзоров, рефератов и библиографии по тематике исследований;
- подготовка к оказанию консалтинговых услуг по данной тематике;
- подготовка к участию в международных проектах по тематике дисциплины;
- подготовка к участию в разработке корпоративной политики и мероприятиях в области повышения социальной ответственности бизнеса перед обществом, включая разработку и реализацию решений, направленных на поддержку социально-значимых проектов;
- совершенствование и расширение общенаучной базы. Повышение уровня общекультурного и нравственного совершенствования своей личности.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
	ОПК-3.1 Способен анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения

ОПК-3 Способен выбирать и (или) разрабатывать подходы к решению типовых и новых задач в области информатики и вычислительной техники, учитывая особенности и ограничения различных методов решения	ОПК-3.7 Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач
	ОПК-3.2 Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
	ОПК-3.6 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК-5 Способен и готов к профессиональному росту и руководству коллективом в области информатики и вычислительной техники, толерантно воспринимая социальные, этнические, конфессиональные и культурные различия	ОПК-5.3 Стремится к получению новых знаний, профессиональному и личностному росту
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности основные законы естествознания, современный математический аппарат и алгоритмы, современные информационно-коммуникационные технологии	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационных технологий, владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности
	ПК-2.3 Имеет практический опыт научно-исследовательской деятельности в области информационно-коммуникационных технологий
	ПК-2.4 Владеет методами и алгоритмами решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического поиска, опыт работы с научными источниками
ПК-3 Владеет навыками участия в научных дискуссиях, выступления с сообщениями и докладами устного, письменного и виртуального (размещение в информационных сетях) характера, представления материалов собственных исследований	ПК-3.1 Знает основы ведения научной дискуссии и формы устного научного высказывания
	ПК-3.2 Умеет вести корректную дискуссию в области информационных технологий задавать вопросы и отвечать на поставленные вопросы по теме научной работы
	ПК-3.3 Имеет практический опыт участия в научных студенческих конференциях, очных, виртуальных, заочных обсуждениях научных проблем в области информационных технологий
ПК-1 Готов к включению в профессиональное сообщество; способен проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности	ПК-1.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет навыками подготовки научных обзоров, публикаций, рефератов и библиографий по тематике проводимых исследований на русском и английском языке
	ПК-1.2 Умеет решать научные задачи с пониманием существующих подходов к верификации моделей программного обеспечения в связи с поставленной целью и в соответствии с выбранной методикой
	ПК-1.3 Имеет практический опыт выступлений и научной аргументации при анализе объекта научной и профессиональной деятельности

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- фундаментальные концепции и профессиональные результаты, системные методологии в профессиональной области;
- современное состояние и принципиальные возможности языков и систем программирования;
- основные методы быстрой сортировки, методы организации динамически изменяемых справочных систем, методы решения оптимизационных задач на графах.

уметь:

- использовать новые знания и применять их в профессиональной деятельности; использовать современные теории, методы, системы и средства прикладной математики и информационных технологий для решения научно-исследовательских и прикладных задач.

владеть:

- основами методологии научного познания и системного подхода при изучении различных уровней организации материи, информации, пространства и времени.

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.		20		15
2	Автоматическая обработка и интерпретация речевого сигнала.		15		10
3	Современный инструментарий для разработки речевых технологий.		5		10
4	Речевые базы данных.		5		10
5	Методы синтеза речи по тексту.		5		10
6	Автоматическое распознавание речи.		5		10
7	Речевой диалог с компьютером.		5		10
Итого часов			60		75
Подготовка к экзамену		0 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

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

Семестр: 1 (Осенний)

1. Введение.

Комплекс целевых задач речевых технологий:

Озвучивание текста, т.е. генерация речи по тексту.

Идентификация диктора.

Сжатие речи.

Речевое управление - голосовой ввод запросов/команд.

Система диктовки - голосовой ввод произвольного текста.

История развития научных подходов к проблемам распознавания и синтеза речи.

Интеллектуальные речевые системы.

2. Автоматическая обработка и интерпретация речевого сигнала.

Формы представления речевого сигнала: осциллограмма, спектр, кепстр, фазовые характеристики.

Отображение спектра речевого сигнала в нелинейных шкалах.

Акустические характеристики, используемые для описания речевого сигнала.

Определение основного тона и гладкое спектральное представление.

Вычисление формант гласных.

Фонетические характеристики.

Методы выделения акустических событий.

3. Современный инструментарий для разработки речевых технологий.

Инструментарий для визуализации различных представлений сигнала и результатов его обработки.

Инструментарий для проведения пакетной обработки большого количества речевых сигналов.

Инструменты создания речевых баз данных.

Системы фонетического обеспечения речевых технологий.

4. Речевые базы данных.

Определения и обоснование необходимости накопления речевых баз данных.

Классификация речевых баз данных.

Технология создания речевых баз данных.

Примеры существующих речевых баз данных.

Использование речевых баз данных в речевых технологиях.

5. Методы синтеза речи по тексту.

Использование фонетических знаний в системах синтеза речи по тексту.

Классификация методов автоматического синтеза речи по тексту.

Блок лингвистической обработки - подготовка текста к озвучиванию.

Нормализация текста.

Синтаксический и морфемный анализ.

Акцентно-интонационный транскриптор.

Фонемный транскриптор.

Блок управления.

Блок генерации сигнала.

6. Автоматическое распознавание речи.

Общая информация о системах автоматического распознавания речи (АРР).

История развития данной области.

Типология систем АРР.

Области применения АРР.

Основные подходы к построению систем АРР.

Архитектура системы АРР.

Модуль обработки речевого сигнала.

Модуль акустического моделирования.

Модуль языкового моделирования.

Модуль поиска оптимального решения.

Статистические модели, используемые в АРР.

Вероятностные распределения характеристик.

Стохастические модели.
Скрытые Марковские модели.
Языковое моделирование в АРР, N-граммы.
Поиск правильного решения.
Алгоритмы поиска правильного решения.
Динамическое программирование.
Реализация систем АРР (на примере систем, разработанных в ведущих лабораториях).

7. Речевой диалог с компьютером.

Техническая необходимость использования речевых каналов общения человека с компьютерными системами.
Техническая подготовленность к применению речевых интерфейсов с автоматическими системами.
Классификация систем речевого интерфейса.
Современное состояние речевых технологий в мире и в России.

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

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6.Перечень рекомендуемой литературы

Основная литература

1. Цифровая обработка сигналов [Текст] = Discrete-Time Signal Processing : [учеб. пособие для вузов] / А. Оппенгейм, Р. Шафер ; пер. с англ. под ред. С. Ф. Боева .— 3-е изд., испр. — М. : Техносфера, 2012 .— 1048 с.
1. Запоминание и адаптивная обработка информации цифровых изображений [Текст] / М. В. Харинов ; под ред. Р. М. Юсупова - СПб.Изд-во СПб. ун-та,2006

Дополнительная литература

1. Фонетика немецкого языка [Текст] / О. Г. Козьмин, Г. А. Сулемова - М.Высшая школа,2004

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

- <http://www.omg.org/>
- <http://www.vdmttools.jp/en/>
- <http://www.ttcn-3.org/>

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

На семинарских занятиях используются мультимедийные технологии, включая демонстрации презентаций.

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

1. Ознакомление с основами: Понимание основных принципов синтеза и распознавания речи, включая акустические и лингвистические аспекты, поможет вам лучше понять процессы, лежащие в основе этих технологий.

2. Изучение алгоритмов: Изучение различных алгоритмов для синтеза и распознавания речи, таких как Hidden Markov Models (HMM), Deep Neural Networks (DNN) и Recurrent Neural Networks (RNN), поможет вам понять, как работают современные системы распознавания речи.
3. Практические упражнения: Проведение практических упражнений, например, использование открытых библиотек и инструментов для синтеза и распознавания речи, поможет вам лучше понять процессы на практике.
4. Чтение научных статей и книг: Изучение актуальных научных статей и книг по теме поможет вам быть в курсе последних достижений в области технологии синтеза и распознавания речи.
5. Участие в проектах: Участие в проектах, связанных с синтезом и распознаванием речи, поможет вам применить полученные знания на практике и улучшить свои навыки.
6. Консультация с экспертами: Общение с профессионалами в области технологии синтеза и распознавания речи может быть полезным для получения дополнительных советов и рекомендаций.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра когнитивных технологий
курс:	<u>1</u>
квалификация:	магистр

Семестр, формы промежуточной аттестации: 1 (осенний) - Дифференцированный зачет

Разработчик: Д.С. Богданов, канд. техн. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-3 Способен выбирать и (или) разрабатывать подходы к решению типовых и новых задач в области информатики и вычислительной техники, учитывая особенности и ограничения различных методов решения	ОПК-3.1 Способен анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-3.7 Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач
	ОПК-3.2 Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
	ОПК-3.6 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК-5 Способен и готов к профессиональному росту и руководству коллективом в области информатики и вычислительной техники, толерантно воспринимая социальные, этнические, конфессиональные и культурные различия	ОПК-5.3 Стремится к получению новых знаний, профессиональному и личностному росту
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности основные законы естествознания, современный математический аппарат и алгоритмы, современные информационно-коммуникационные технологии	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационных технологий, владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности
	ПК-2.3 Имеет практический опыт научно-исследовательской деятельности в области информационно-коммуникационных технологий
	ПК-2.4 Владеет методами и алгоритмами решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического поиска, опыт работы с научными источниками
ПК-3 Владеет навыками участия в научных дискуссиях, выступления с сообщениями и докладами устного, письменного и виртуального (размещение в информационных сетях) характера, представления материалов собственных исследований	ПК-3.1 Знает основы ведения научной дискуссии и формы устного научного высказывания
	ПК-3.2 Умеет вести корректную дискуссию в области информационных технологий задавать вопросы и отвечать на поставленные вопросы по теме научной работы
	ПК-3.3 Имеет практический опыт участия в научных студенческих конференциях, очных, виртуальных, заочных обсуждениях научных проблем в области информационных технологий
ПК-1 Готов к включению в профессиональное сообщество; способен	ПК-1.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет навыками подготовки научных обзоров, публикаций, рефератов и библиографий по тематике проводимых исследований на русском и английском языке

проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности

ПК-1.2 Умеет решать научные задачи с пониманием существующих подходов к верификации моделей программного обеспечения в связи с поставленной целью и в соответствии с выбранной методикой

ПК-1.3 Имеет практический опыт выступлений и научной аргументации при анализе объекта научной и профессиональной деятельности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Технологии синтеза и распознавание речи» обучающийся должен:

знать:

- фундаментальные концепции и профессиональные результаты, системные методологии в профессиональной области;
- современное состояние и принципиальные возможности языков и систем программирования;
- основные методы быстрой сортировки, методы организации динамически изменяемых справочных систем, методы решения оптимизационных задач на графах.

уметь:

- использовать новые знания и применять их в профессиональной деятельности; использовать современные теории, методы, системы и средства прикладной математики и информационных технологий для решения научно-исследовательских и прикладных задач.

владеть:

- основами методологии научного познания и системного подхода при изучении различных уровней организации материи, информации, пространства и времени.

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

Пример варианта контрольной работы/задания:

- Перечислите характеристики наборов инструкций AVX, AVX2, AVX512,
- Дайте определение ассоциативности сверхбыстродействующей памяти,
- Определить число операций умножения, сложения, доступа к памяти на каждом шаге и для алгоритма Штрассена быстрого умножения матриц.

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

1. Генерация речи по тексту.
2. Акустические характеристики, используемые для описания речевого сигнала.
3. Визуализация различных представлений сигнала и результатов его обработки.
4. Использование речевых баз данных в речевых технологиях.
5. Основные подходы к построению систем APP.
6. Статистические модели, используемые в APP.
7. Алгоритмы поиска правильного решения.
8. Классификация систем речевого интерфейса.
9. Технологии автоматического распознавания речи.
10. Технологии синтеза и распознавания речи в реальных приложениях.

Критерии оценивания

отлично (10) - 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.

отлично (9) - выставляется студенту, показавшему свободное оперирование знаниями учебной программы дисциплины, выполнение заданий творческого характера.

отлично (8) - выставляется студенту, показавшему владение программным учебным материалом с наличием несущественных ошибок в действиях, самостоятельно исправляемых учащимся.

хорошо (7) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускается в ответе или в решении задач некоторые неточности.

хорошо (6) - выставляется студенту если он осознает воспроизведение программного учебного материала, в том числе и различной степени сложности, с несущественными ошибками, затруднения в применении отдельных навыков.

хорошо (5) - выставляется студенту если теоретическое содержание освоено не полностью, некоторые практические навыки сформированы недостаточно, в некоторых случаях были допущены ошибки.

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 владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации.

удовлетворительно (3) - выставляется студенту в случае большого количества недочетов и неправильных ответов, а также пассивной работе в ходе занятий, многие учебные задания не выполнены.

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

неудовлетворительно (1) - выставляется студенту, который не освоил теоретическое и практическое содержание курса, все выполненные учебные задания содержат грубые ошибки.

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

При проведении дифференцированного зачета обучающемуся предоставляется 20 минут на подготовку. Опрос обучающегося не должен превышать двух астрономических часов.

Во время проведения дифференцированного зачета обучающиеся могут пользоваться программой дисциплины, а также литературой из списка рекомендуемой литературы.