

2.3. Выводы по главе 2 38

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)						
Изм.	Лист	№ докум.	Подпись	Дата							
Разраб.		Лукьянчикова			Модель и алгоритмы поиска сигналов в звуковых базах данных	Лит.	Лист	Листов			
Провер.		Гай В.Е.						3	52		
Реценз.						НГТУ им. Р.Е. Алексеева					
Н. Контр.											
Утверд.											

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
<div> <div> <div>Изм.</div> <div>Лист</div> <div>№ докум.</div> <div>Подпись</div> <div>Дата</div> </div> <div> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </div> </div>					<div>Лист</div> <div>4</div>

Глава 3. Вычислительный эксперимент.....39

3.1. Описание эксперимента39

3.2. Описание программного продукта41

3.3. Анализ результатов эксперимента.....44

3.4. Выводы по главе 349

Заключение50

Список литературы.....51

Перв. примен.	<h2 style="text-align: center;">Введение</h2> <h3>Актуальность темы</h3> <p>В настоящее время область распознавания звуковых сигналов получила широкое распространение. С тех пор, как в повседневной жизни плотно укоренились смартфоны с функцией интерпретации голосовых команд и других звуковых сигналов, стали появляться все новые и новые алгоритмы распознавания. Самыми известными приложениями по распознаванию звукового сигнала стали Siri, Google Поиск, Shazam и другие.</p> <p>При разработке такого алгоритма, следует решить ряд задач:</p> <ul style="list-style-type: none"> - проблема формирования исходного описания; - проблема формирования системы признаков; - проблема принятия решения в условиях априорной неопределенности; - проблема поиска сигнала в звуковой базе данных. <p>В данном исследовании я хочу предложить свой метод решения этих проблем и подробнее остановится на проблеме поиска сигнала в звуковой базе.</p> <p>Данный алгоритм распознавания можно будет использовать не только для распознавания звуковых композиций. В современном мире происходит регулярная кража индивидуальной собственности. С его помощью можно будет определять заимствования современными певцами и композиторами звуковых фрагментов из чужих песен. Еще одной применение - в криминалистике. Данный алгоритм может помочь при сопоставлении двух записей голосов и выявления сходства.</p> <h3>Цель работы и задачи исследования</h3> <p>Целью данной работы является разработка новых моделей и алгоритмов нахождения фрагментов сигналов в звуковых базах данных.</p> <p>Задачи, которые придется решить для достижения поставленной цели:</p> <ul style="list-style-type: none"> - анализ существующих методов распознавания звуковых сигналов; - анализ существующих методов поиска звуковых сигналов в базах данных; - разработка модели признакового описания звукового сигнала; - разработка алгоритма поиска звукового сигнала в базах данных; 				
Подпись и дата	Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.	<div style="display: flex; justify-content: space-between;"> <div> <div style="border: 1px solid black; padding: 2px;">Изм.</div> <div style="border: 1px solid black; padding: 2px;">Лист</div> </div> <div> <div style="border: 1px solid black; padding: 2px;">№ докум.</div> <div style="border: 1px solid black; padding: 2px;">Подпись</div> </div> <div> <div style="border: 1px solid black; padding: 2px;">Дата</div> </div> </div>
					<div style="display: flex; justify-content: space-between;"> <div></div> <div>Лист</div> </div> <div style="text-align: center; font-weight: bold;">5</div>

Перв. примен.	<p>- проведение вычислительного эксперимента для подтверждения работоспособности разработанного алгоритма.</p> <p>Объект исследования</p> <p>Объектом исследования в данной работе является сигнал, полученный для поиска схожих сигнатур в звуковой базе данных.</p> <p>Предмет исследования</p> <p>Предметом исследования в данной работе являются модели и методы реализации поиска сигнала в звуковой базе данных.</p> <p>Методы исследования</p> <p>В ходе выполнения данного исследования были использованы методы формирования признакового описания звукового сигнала, основанные на теории активного восприятия, методы для сравнения цифровых последовательностей. В качестве инструмента для реализации представленного алгоритма был выбран язык программирования R.</p> <p>Положения, выносимые на защиту</p> <p>На защиту выносятся следующие положения:</p> <ul style="list-style-type: none"> - информационная модель признакового описания звукового сигнала; - разработанный алгоритм поиска сигнала в звуковой базе; - результаты эксперимента, подтверждающие работоспособность разработанного метода. <p>Научная новизна</p> <p>Научная новизна данного исследования заключается в применении новых подходов к реализации некоторых этапов распознавания звукового сигнала. На этапе признакового</p>																	
	Справ. №																	
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.													
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td rowspan="2">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</td> <td>Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td>6</td> </tr> </table>						Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист						6
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист												
						6												

Перв. примен.	описания применяются методы описанные в теории активного восприятия. База данных представляет из себя бинарное дерево, которое содержит в себе сегменты звуковых композиций, пронумерованные специальными индикаторами и содержащие признаковые описания.													
Справ. №	<p>Теоретическая и практическая ценность работы</p> <p>Теоретическая ценность работы заключается в разработке подхода распознавания и поиска звукового сигнала с применением новых методов решения проблем, возникающих на этапах системы распознавания.</p> <p>Практическая ценность данного исследования - это создание по описанному алгоритму программного обеспечения, которое впоследствии может применяться в различных исследованиях, криминалистических экспертизах или для защиты интеллектуальной собственности.</p>													
Подпись и дата	<p>Апробация работы</p> <p>Апробация данного исследования была проведена в ходе защиты отчета по распределенной научно-исследовательской работе, включавший себя презентацию разработанного алгоритма и демонстрацию результатов вычислительного эксперимента.</p>													
Инв. № дубл.	<p>Обоснованность и достоверность</p> <p>Обоснованность и достоверность данного исследования доказывается анализом результатов вычислительного эксперимента, проводимого для разработанного алгоритма, и сравнения их с результатами других известных алгоритмов.</p>													
Взам. инв. №	<p>Публикации</p> <p>По данному исследованию планируется публикация в виде доклада или тезиса на одной из крупных научных конференций.</p>													
Подпись и дата														
Инв. № подл.														
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата					
Изм.	Лист	№ докум.	Подпись	Дата										
<div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</div> <div>Лист 7</div>														

Перв. примен.	<h2 style="text-align: center;">Объем и структура работы</h2> <p>Выпускная квалификационная работа состоит из введения, трех глав основной части, заключения и списка литературы. Общий объем работы 52 страниц. Диссертация содержит 19 рисунков, 13 таблиц, 11 формул. Список литературы включает в себя 19 наименований.</p>					Справ. №
<div>Подпись и дата</div> <div>Инв. № дубл.</div> <div>Взам. инв. №</div> <div>Подпись и дата</div> <div>Инв. № подл.</div>						
	Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)
						<div>Лист</div> <div>8</div>

Глава 1. Обзор существующих методов

1.1 Структура системы распознавания сигналов

В большинстве случаев система распознавания сигналов состоит из 3 этапов (рисунок 1.1):

- предварительная обработка сигнала;
- формирование системы признаков;
- классификация сигнала на основе системы признаков.

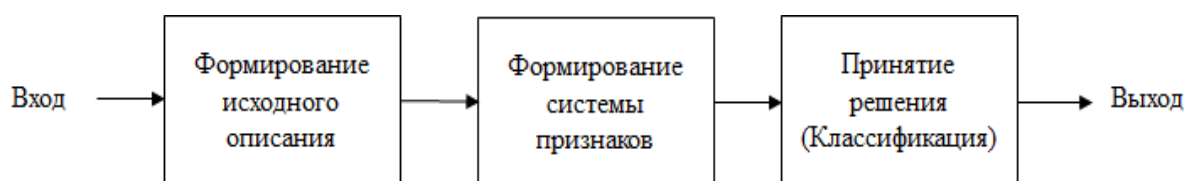


Рисунок 1.1 Схема системы распознавания звукового сигнала

На этапе предварительной обработки сигнал подвергается преобразованиям, которые помогают представить его в удобной форме для последующего анализа.

На этапе формирования системы признаков решается одна из главных задач распознавания - однозначная идентификация сигнала.

На этапе классификации происходит формирование сегментов сигналов в группы исходя из признакового описания. Чаще всего для этого используются уже имеющиеся алгоритмы кластеризации.

Поиск интересующего фрагмента сигнала в уже сформированной базе данных обычно происходит в 2 этапа:

- составление признакового описания входного сигнала;
- перебор имеющихся записей и сравнение их признакового описания с описателем входного сигнала.

На этапе составления признакового описания к сигналу применяются методы, в результате которых получается уникальный идентификатор этого сигнала.

На этапе перебора полученный идентификатор сравнивается с идентификаторами, имеющимися в базе данных и выявляется сходство. Известно много способов поиска схожих фрагментов сигналов.

Известны методы поиска повторяющихся фрагментов в речевых сигналах и музыкальных базах.

Перв. примен.	Справ. №									
Подпись и дата	Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.						
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)					Лист
										10

1.2 Обзор существующих методов поиска звукового фрагмента в речевых сигналах

В последнее время актуальны алгоритмы распознавания сигналов в условиях нулевых ресурсов.

Главными преимуществами таких алгоритмов является отсутствие заготовленной базы данных, вследствие чего это освобождает их от длительного перебора, сравнения признаков описателей и выявления схожих сегментов сигнала. На входе такого процесса поступает звуковой сигнал, который анализируется определенным образом, и результатом этих манипуляций являются временные промежутки, фонемы или признаковые описатели идентифицирующие найденные повторяющиеся звуковые фрагменты.

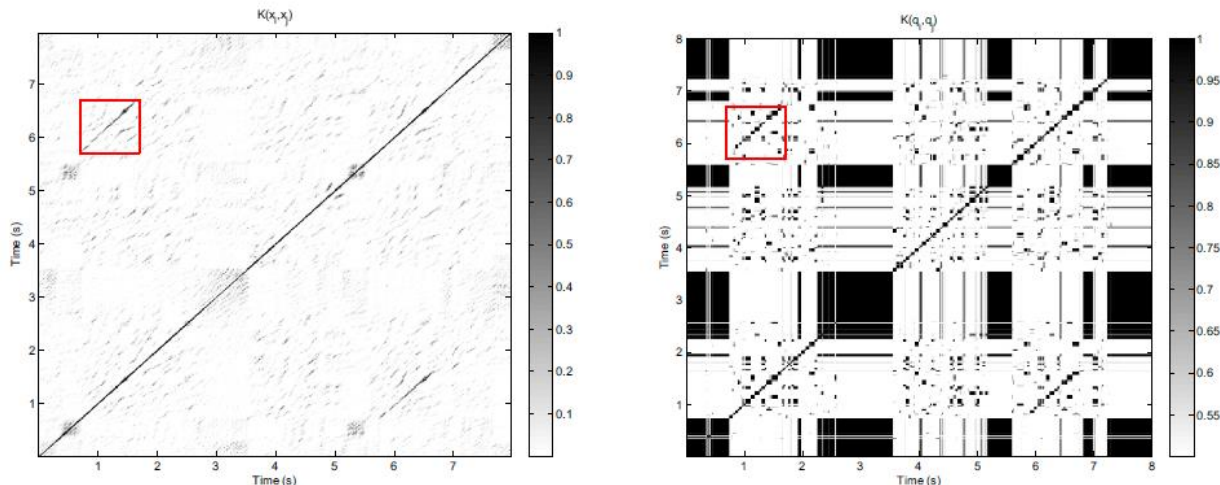
Очень популярным подходом в данных алгоритмах является метод построения матриц схожести или кодирование звуковых последовательностей, путем присвоения определенных признаков описателей единичным звуковым фонемам, так же широко применяется алгоритм динамической трансформации временной шкалы (*dynamic time warping*) [1][2].

Далее будут приведены обзоры двух алгоритмов распознавания звуковых фрагментов в условиях нулевых ресурсов.

1.2.1 Обзор метода поиска повторяющегося фрагмента сигнала с помощью постериограмм и одномерного преобразования Хафа

Решением задачи поиска терминов в повторяющемся промежутке речи основан на графическом способе «*dotplots*», применяющегося для сравнения последовательностей, и использующий алгоритм «*segmental dynamic time warping (S-DTW)*» [1].

Dotplots является графическим методом сравнения последовательностей, изначально разрабатывался для определения повторяющиеся последовательности белка в сообществе биоинформатики, а затем распространен до обработки больших масштабов текста в более общем виде. Берется строка символов $X = x_1, x_2 \dots x_n$, где каждый x_i это элемент какого-то алфавита, при этом *dotplots* будет логической матрицей M размера $n \times n$ и определяется как $M_{ij} = (x_i == y_j)$. Подстроки повторяются в разных точках X и проявляются как диагональные отрезки в визуализации M ; подстроки, состоящие из одного

Перв. примен.	<p>повторяющегося знака, будут проявляться в виде больших блоков. Данная работа применяет этот метод в акустической обработке с целью выявления повторяющихся интервалов в речи или аудио.</p> <p>Формально, $X = x_1, x_2 \dots x_n$, это векторные временные ряды, которые вычисляются по аудио сигналу, где каждый $x_i \in R^d$. Затем определяется акустический <i>dotplot</i> M как обобщенная матрица $M_{ij} = K(x_i, x_j)$, где $K: R^d \times R^d \rightarrow R$ является симметричной функцией сходства между векторами признаков. В этом исследовании, вычисляется косинусное сходство</p> $K(x, y) = \frac{1}{2} \left[1 + \frac{\langle x, y \rangle}{\ x\ \ y\ } \right] \quad (1.1)$ <p>которое принимает значение 1, когда точки x и y том же направлении, 0,5 - когда они ортогональны, и 0 - если они указывают в противоположных направлениях. Учитывая акустический <i>dotplot</i> вычисляется для пары высказываний, задача нахождения повторяющихся интервалов сводится к поиску диагональных отрезков.</p> <div></div> <p>Рисунок.1.2 Постериограммы звукового сигнала</p> <p>Далее задача сводится нахождения повторяющихся интервалов к поиску диагональных отрезков (рисунок 1.2). Для этого, после преобразования изображения, применяется классическое одномерное преобразование Хафа. Данное преобразование направлено на поиск простых изображений, типа прямых, кругов или эллипсов. В данном случае, будут искаться отрезки.</p>				
	Справ. №				
Подпись и дата					
	Инов. № дубл.				
	Взам. инв. №				
Подпись и дата					
Инов. № подл.					
	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)				
	Лист 12				
Изм.	Лист	№ докум.	Подпись	Дата	

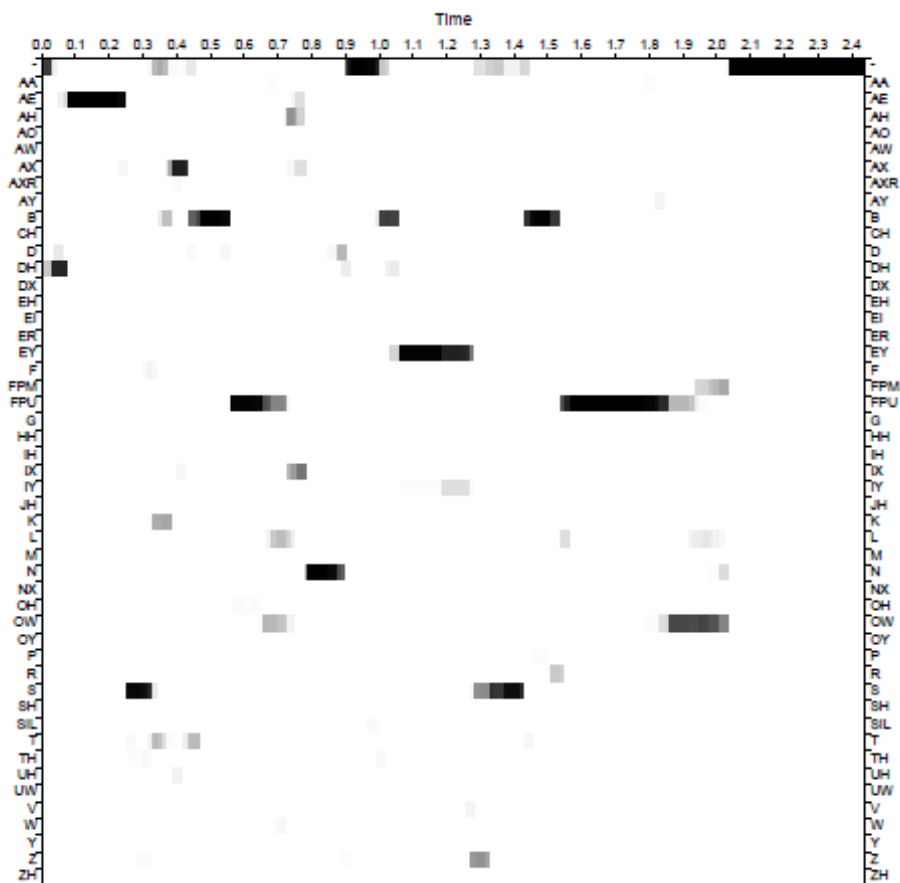
Перв. примен.																					
Справ. №																					
<h2 style="text-align: center;">Преобразование Хафа</h2> <p>Прямая может быть задана уравнением $y = mx + b$ и может быть вычислена по любой паре точек (x, y) на изображении. Главная идея преобразования Хафа [3] — учесть характеристики прямой не как уравнения, построенное по паре точек изображения, а в терминах её параметров, то есть m — коэффициента наклона и b — точки пересечения с осью ординат. Исходя из этого <i>прямая</i>, заданная уравнением $y = mx + b$, может быть представлена в виде <i>точки</i> с координатами (b, m) в пространстве параметров.</p> <p>Однако прямые, параллельные оси ординат, имеют бесконечные значения для параметра m. Поэтому лучше представить прямую с помощью других параметров, r — это длина радиус-вектор ближайшей к началу координат точки на прямой (т.е. нормали к прямой, проведенной из начала координат), а θ — это угол между этим вектором и осью абсцисс. При таком описании прямых не возникают бесконечные параметры.</p> <p>Уравнение прямой принимает вид:</p> $y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right) \quad (1.2)$ <p>После преобразования:</p> $r = x \cos\theta + y \sin\theta \quad (1.3)$ <p>Это соответствует синусоидальной линии в пространстве Хафа (r, θ), которая, в свою очередь, уникальна для данной точки и однозначно её определяет. Если эти линии (кривые), соответствующие двум точкам, накладываются друг на друга, то точка (в <i>пространстве Хафа</i>), где они пересекаются, соответствует прямой (в оригинальном месте изображения), которые проходят через обе точки. В общем случае, ряд точек, которые формируют прямую линию, определяют синусоиды, которые пересекаются в точке параметров для той линии. Таким образом, проблема обнаружения коллинеарных точек может быть сведена к проблеме обнаружения пересекающихся кривых.</p>																					
Инв. № подл.																					
Подпись и дата																					
Взам. инв. №																					
Инв. № дубл.																					
Подпись и дата																					
<table border="1" style="width: 100%;"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<table border="1" style="width: 100%;"> <tr> <td colspan="2">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</td> <td>Лист</td> </tr> <tr> <td colspan="2"></td> <td>13</td> </tr> </table>	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)		Лист			13
Изм.	Лист	№ докум.	Подпись	Дата																	
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)		Лист																			
		13																			

1.2.2 Обзор метода поиска фрагмента сигнала с применением фонетических постериограмм

Данный метод опирается на распознавание речевых терминов с запросом-по-образцу с использованием фонетических постериограммных шаблонов [2]. Этот метод нужен, когда пользователь нашел интересную для себя информацию и он желает найти еще больше аналогичных данных. Выбирается аудио фрагмент, содержащий ключевое слово, производится сравнение всего аудиосигнала с данным фрагментом и находятся соответствия.

Используются фонетические постериограммы размерностью класс ко времени, чтобы сопоставить сегменты примера запроса с фонетическими классами для каждого момента времени. На рисунок 1.3 представлена постериограмма для аудио сегмента включающего произнесенную фразу «*basketball and baseball*». По горизонтальной оси – время в секундах, по вертикальной – фонетические классы. Уровень затемнения соответствует вероятности встречи во временном промежутке данного класса.

Перв. примен.														
Справ. №														
Подпись и дата														
Инв. № дубл.														
Взам. инв. №														
Подпись и дата														
Инв. № подл.														
<table border="1"> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> </table>										Изм.	Лист	№ докум.	Подпись	Дата
Изм.	Лист	№ докум.	Подпись	Дата										
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)														
Лист 14														

Перв. примен.								
Справ. №								
Подпись и дата							<p>Рисунок 1.3 Постериограмма для сегмента сигнала изображающая отношения сказанной фонемы к временному интервалу</p> <p>Чтобы найти звуковые сегменты похожие на исходный запрос используя пастериограммы, авторы предлагают ввести определенное значение для индивидуальных апостериорных распределений. Постериограмма речевого сегмента Q представляется серией векторов, включающих фонетическую апостериорную вероятность для N кадров в речевом сегменте:</p> $Q = \{q_1, \dots, q_n\} \quad (1.4)$ <p>Для обозначения постериограмм сегмента запроса используется Q, а X – для постериограмм тестового высказывания, включающего M кадров. Цель – это выявить сходства между Q и некоторыми участками X.</p> <p>Учитывая два распределения q и x, вероятность того, что эти распределения</p>	
Инв. № дубл.								
Взам. инв. №								
Подпись и дата								
Инв. № подл.								
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)			Лист 15

Перв. примен.	<p>принадлежат к одному фонетическому событию легко представить, как их скалярное произведение:</p> $P(\text{phone}\{q\} = \text{phone}\{x\}) = q \cdot x \quad (1.5)$ <p>Мы можем интерпретировать эту вероятность:</p> $D(q, x) = -\log(q \cdot x) \quad (1.6)$ <p>Значения близкие к нулю означают больше сходство между q и x, а большие положительные значения – различие. Если q и x равны нулю, то $q \cdot x = 0$ и, следовательно $D(q, x) = \infty$. Для этого применяется сглаживание постериограмм:</p> $q = (1 - \lambda)q + \lambda u \quad (1.7)$ <p>Здесь u – вектор представляющий равномерное распределение вероятностей, и $\lambda > 0$ гарантирует ненулевую вероятность для фонетических постериоров в q.</p> <p>Для сравнения постериограмм запроса и тестового высказывания вычисляется сходство значений между индивидуальным апостериорным распределением для всех N кадров в запросе и индивидуальным апостериорным распределением для M кадров в тестовом высказывании. Результатом является матрица сходств размера $N \times M$.</p> <p>Далее, основной целью является нахождение временного интервала в тестовом высказывании, соответствующего запросу. На рисунок 1.4 показана постериограмма матрицы сходства между тестовым высказыванием (оно по оси x) и запрашиваемым термином (по оси y). Темные области представляют большое сходство между кадрами тестового высказывания и кадрами запроса. Белые – наоборот. В идеале сходство должно выглядеть как последовательность диагоналей блоков от левого верхнего угла до нижнего правого внутри матрицы сходств. Матрица на рисунке ниже показывает действительное сходство между запросом и тестовым высказыванием, зеленая линия представляет лучшее схожее выравнивание запроса со схожим участком в тестовом высказывании.</p>													
	Справ. №													
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.									
	<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата				
Изм.	Лист	№ докум.	Подпись	Дата										
<div style="display: flex; justify-content: space-between; align-items: center;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) <div style="border: 1px solid black; padding: 2px;"> Лист 16 </div> </div>														

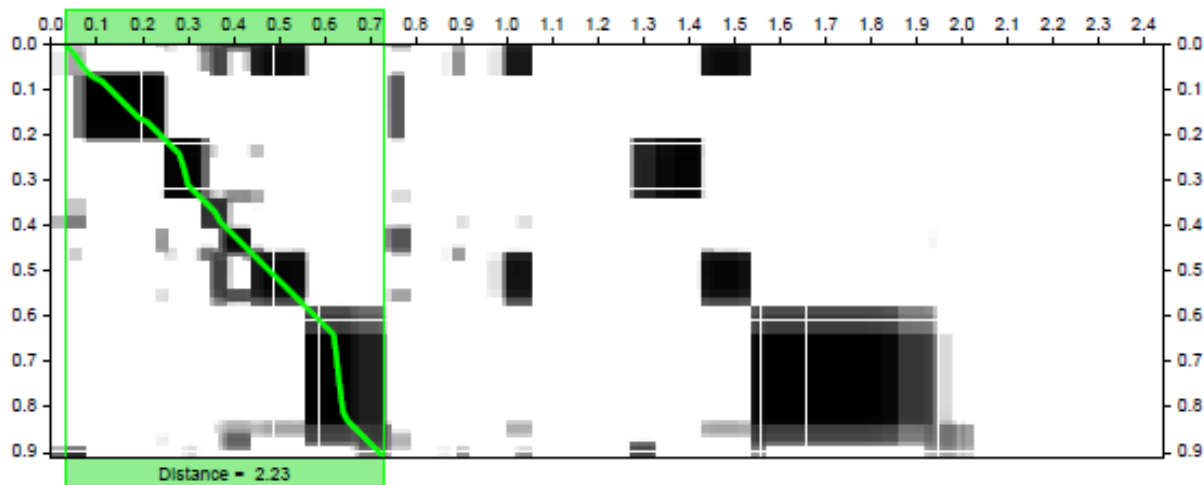


Рисунок 1.4 Постериограмма матрицы сходства между тестовым высказыванием и запрашиваемым термином

Для поиска лучшего пути используется модифицированный DTW-алгоритм. Путем накопления баллов выбирается наклон диагоналей отдельного блока, и как следствие строится оптимальный путь.

Также данный метод может использоваться, когда дано несколько шаблонов запрашиваемых терминов. Для решения этой проблемы выделяется 2 подхода. Первый – комбинирование свойств всех шаблонов в одном. Второй – использование всех шаблонов для получения оценки и объединения результатов. Здесь используется второй подход. Для каждого варианта запроса существуют различные баллы оценки для входного высказывания.

Этот подход очень эффективно использовать, имея таблицу соответствий запрашиваемого термина и его балл. Таким образом можно быстро и легко находить нужные слова в речевом сигнале.

1.3 Обзор существующих методов поиска звукового сигнала в музыкальных базах данных

Алгоритмы распознавания звуковых фрагментов с использованием предварительно сформированной базой данных пока остаются самыми распространенными в повседневном использовании.

При разработке алгоритмов для уникальной идентификации сигнала крупные компании стараются использовать максимальные колебания амплитуды в песне, называемыми пиками. Набор таких пиков позволяет описать любую музыкальную композицию [6].

На этапе признакового описания сигнала в таких алгоритмах чаще всего используются частотные характеристики сигнала, полученные с помощью преобразования Фурье, или мелчастотные кепстральные коэффициенты [4][5].

Одной из важнейших проблем в этом типе алгоритмов является быстрый поиск схожего со входным сигналом экземпляра из базы данных. Тут на помощь приходят нейронные сети, алгоритмы фильтрации и простой перебор элементов в базе.

Задача, которую решают такие алгоритмы формулируется следующим образом: на сервер поступает фрагмент аудиосигнала, после чего происходит поиск среди треков, находящихся в базе данных того, из которого взят данный интересующий фрагмент.

Далее будут приведены обзоры двух наиболее известных алгоритмов распознавания музыкальных файлов в звуковых базах данных.

1.3.1 Обзор алгоритма поиска музыки в приложении Shazam

Принцип работы мобильного приложения *Shazam* заключается в записи фрагмента песни, анализа его и нахождения соответствия в звуковой базе данных [4].

На этапе определения признакового описания сигнала алгоритм использует быстрое преобразование Фурье.

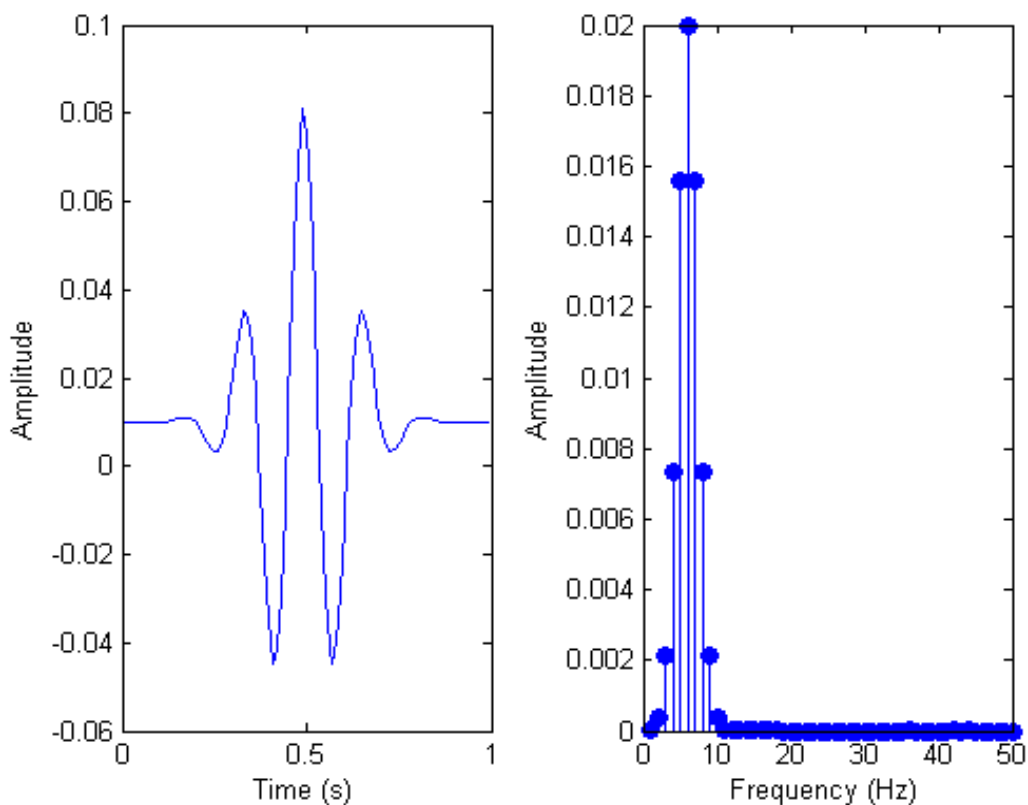


Рисунок 1.5 Сигнал и его преобразование Фурье.

Отрицательным моментом использования быстрого преобразования Фурье является потеря информации о времени после проведения анализа. Как результат - это звуковые частоты и их амплитуды без привязки к временному промежутку, в котором они появляются. Для этого алгоритм использует метод скользящего окна, который анализирует участки сигнала фиксированной длины и выдает признаковое описание для текущей части.

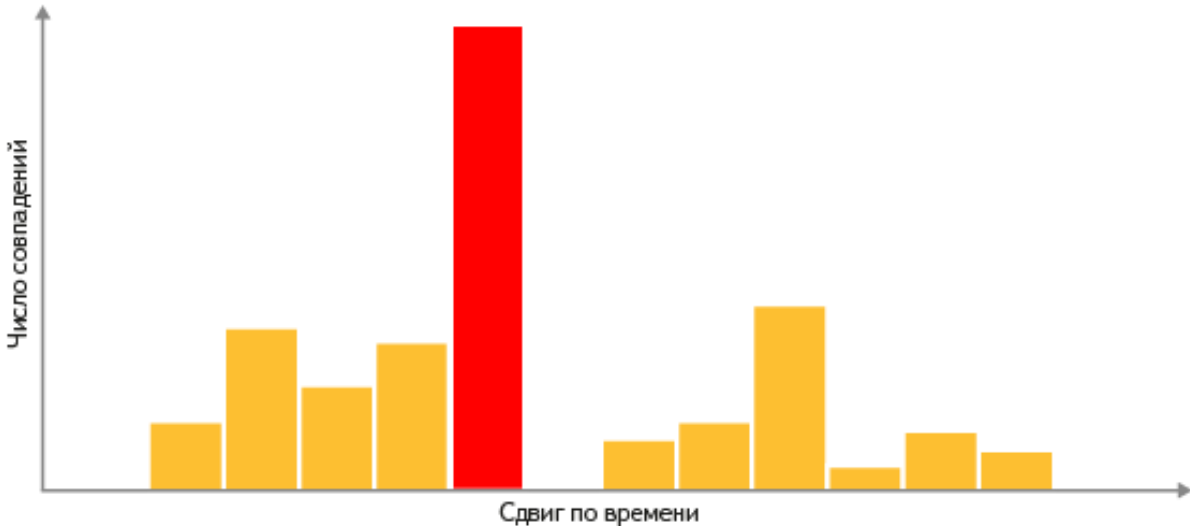
После этого, когда получены частотные характеристики сигнала, происходит формирование цифровой сигнатуры. Из огромного числа частот выбираются только те, которые важнее всего. Здесь отслеживаются частоты с максимальной амплитудой. Это объясняется тем, что частоты с максимальной амплитудой вероятнее всего выживут в условиях зашумленности звукового сигнала [6].

Для упрощения поиска музыкальных композиций их признаковых описаний используются как ключи в хэш-таблице. Каждому ключу соответствует значения времени, когда набор частот появился в песни и идентификатор музыкальной записи (название

Перв. примен.	<p>песни и её исполнитель). В таблице 1.1 представлен пример, как подобные записи могут храниться в базе.</p> <p>Таблица 1.1 Соответствие идентификатора сегмента ко времени и принадлежность к конкретной звуковой композиции</p> <table border="1" style="width: 100%; border-collapse: collapse; margin: 10px 0;"> <tr> <th style="width: 30%;">Хэш-тег</th> <th style="width: 20%;">Время, с</th> <th style="width: 50%;">Идентификатор</th> </tr> <tr> <td>30 51 99 121 195</td> <td>53,52</td> <td>Песня А Исполнитель А</td> </tr> <tr> <td>33 56 92 151 185</td> <td>12,32</td> <td>Песня В Исполнитель В</td> </tr> <tr> <td>34 57 95 111 200</td> <td>23,11</td> <td>Песня А Исполнитель А</td> </tr> <tr> <td>30 51 99 121 195</td> <td>10,89</td> <td>Песня С Исполнитель С</td> </tr> </table> <p>Для определения музыкальной композиции записывается звук и прогоняется через вышеописанный процесс определения сигнатур. Затем начинается поиск композиции в базе по вычисленному хэш-тэгу.</p> <p>При идентификации композиции может получиться так, что один хэш-тег полученного сигнала может соответствовать двум различным песням. Для того, чтобы избежать этого, исходный сигнал дробиться и проверяются совпадения хэш-тэгов нескольких сегментов.</p> <p style="text-align: center;">1.3.2 Обзор алгоритма поиска музыки в приложении Yandex</p> <p>В качестве базы данных алгоритм, предложенный компанией Яндекс, использует все треки приложения Яндекс.Музыка, которые были предоставлены правообладателями для прослушивания пользователями [5]. Также была создана подборка музыкальных композиций из интернета, которыми чаще всего интересуются пользователи.</p> <p>На этапе разработки алгоритма, разработчики отвергли ряд подходов к поиску поступившего фрагмента с треком, такие как:</p> <ul style="list-style-type: none"> - побитовое сравнение. Помехи, шумы, громкоговоритель источника звука, акустика помещения, неравномерная амплитудно-частотная характеристика микрофона - все это создает неточности в перекодировании записанного фрагмента и делает неприемлемым побитовое сравнение. 					Хэш-тег	Время, с	Идентификатор	30 51 99 121 195	53,52	Песня А Исполнитель А	33 56 92 151 185	12,32	Песня В Исполнитель В	34 57 95 111 200	23,11	Песня А Исполнитель А	30 51 99 121 195	10,89	Песня С Исполнитель С	Справ. №
	Хэш-тег	Время, с	Идентификатор																		
30 51 99 121 195	53,52	Песня А Исполнитель А																			
33 56 92 151 185	12,32	Песня В Исполнитель В																			
34 57 95 111 200	23,11	Песня А Исполнитель А																			
30 51 99 121 195	10,89	Песня С Исполнитель С																			
Подпись и дата						Инв. № дубл.															
Подпись и дата						Взам. инв. №															
Инв. № подл.						Подпись и дата															

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист
Изм.	Лист	№ докум.	Подпись	Дата		20

Перв. примен.	<p>- водяные знаки. Если бы компания Яндекс сама выпускала музыку, то могла бы встроить в музыкальные композиции звуковой аналог “водяных знаков”, которые легко распознают алгоритмы, но недоступны человеческому уху.</p> <p>- нестрогое сравнение спектрограмм. На рисунке 1.6 изображены спектрограммы оригинального трека и записанного фрагмента. Их можно рассматривать как изображения и использовать для поиска интересующей композиции.</p>				
	<div> <div> <div>Записанный фрагмент</div> </div> <div> <div>Искомый фрагмент оригинального трека</div> </div> </div> <p>Рисунок 1.6 Спектограммы записанного входного фрагмента и искомого фрагмента трека из базы данных</p> <p>Однако, у данного метода есть ряд проблем:</p> <p>- сравнение огромного числа картинок (6 миллионов треков в базе данных Яндекса) -</p>				
Справ. №					
Инв. № подл.					
Взам. инв. №					
Инв. № дубл.					
Подпись и дата					
Подпись и дата					
<div> <div>Изм.</div> <div>Лист</div> <div>№ докум.</div> <div>Подпись</div> <div>Дата</div> </div> <div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</div>					Лист
					21

Перв. примен.	это очень ресурсозатратная и долгая процедура				
	- оказывается, что одни различия более показательны, чем другие.				
Справ. №	В итоге, для каждого трека нужно хранить минимальное количество характерных признаков.				
	Устойчивыми к помехам и шумам являются пики спектрограмм, выделенные различными способами - например, как точки локального максимума амплитуды. Положение пиков в системе “частота-время” мало меняется. На один трек в среднем приходится 300 тыс. пиков и такая характеристика композиции лучше сопоставляется с миллионной базой треков, чем картинки.				
	Однако, шум при записи фрагмента добавляет новые пики и глушит другие. Поэтому в данном алгоритме используется двухэтапное сравнение:				
	- сначала для каждого трека ищется смещение, при котором с запросом совпадает максимальное количество пиков;				
	- затем из выбранных треков ищется трек с наибольшим числом совпадений.				
	Для этого строится гистограмма, в которой для каждой частоты пика из запроса и трека откладывается +1 по оси Y в том смещении, где нашлось совпадение (рисунок 1.7):				
Подпись и дата					
	Рисунок 1.7 Гистограмма отношения числа совпадений к сдвигу по времени				
Подпись и дата	Выбирается композиция с самым высоким столбцом, который является мерой близости данного трека с запросом.				
Инв. № подл.					
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)					Лист
					22

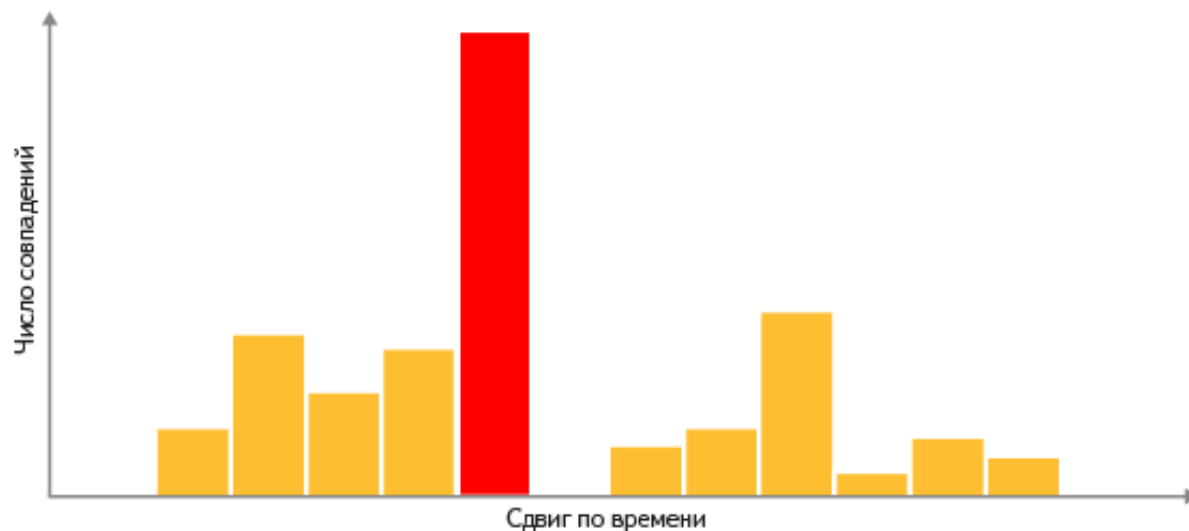


Рисунок 1.7 Гистограмма отношения числа совпадений к сдвигу по времени

Выбирается композиция с самым высоким столбцом, который является мерой близости данного трека с запросом.

Из всех пиков сигнала выбираются не все, а только устойчивые к искажениям. Такой отбор происходит следующим образом:

- внутри одной частоты, по оси времени по всей записи вводится порог, который при обнаружении пика выше текущего положения, срезает разницу между значением порога и верхней точкой (рисунок 1.8). Потом порог поднимается на первоначальную высоту этого пика.

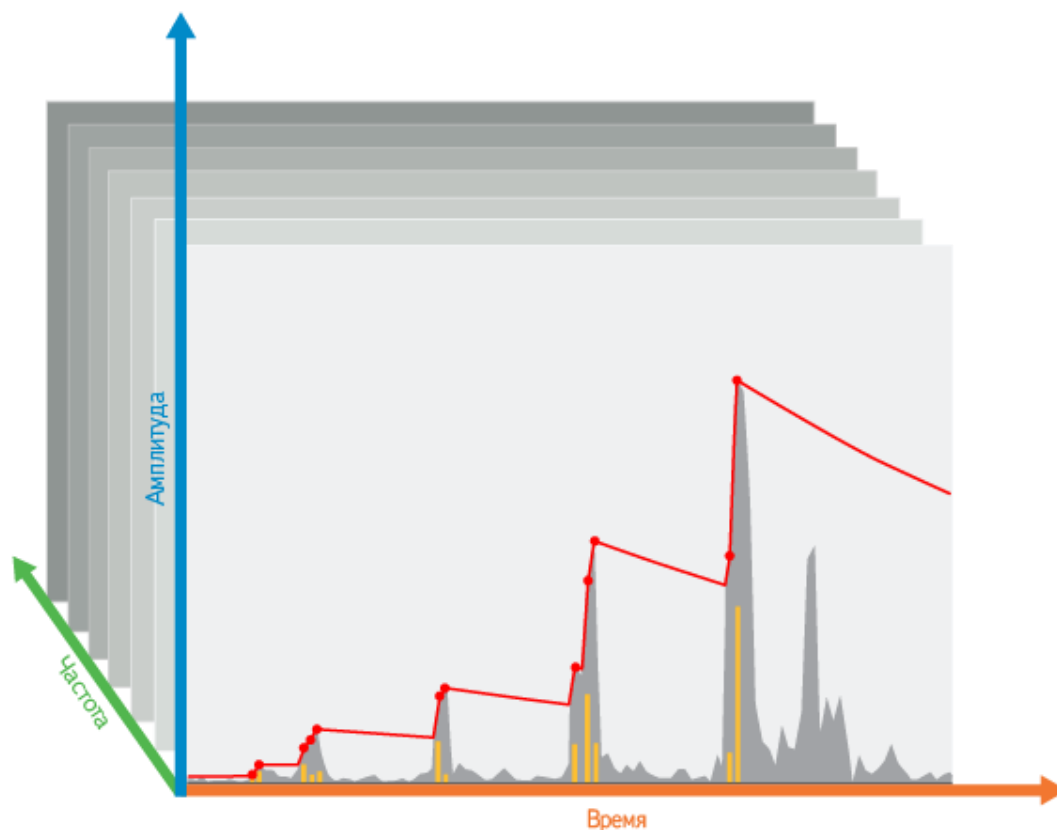


Рисунок 1.8 Пример работы порога

- среди всех частот, выбираем самые контрастные пики (рисунок 1.9), т.е. самые большие локальные максимумы среди срезанных верхушек.

При отборе пиков есть несколько параметров: скорость опускания порога, число выбираемых пиков в каждом временном интервале и окрестность влияния пиков на соседей. На основе этих критериев подбирается такая комбинация, при которой остается минимальное число пиков.

- предварительный отбор треков по очень разряженному набору контрастных пиков.
- выбирается гарантированно наилучший ответ.

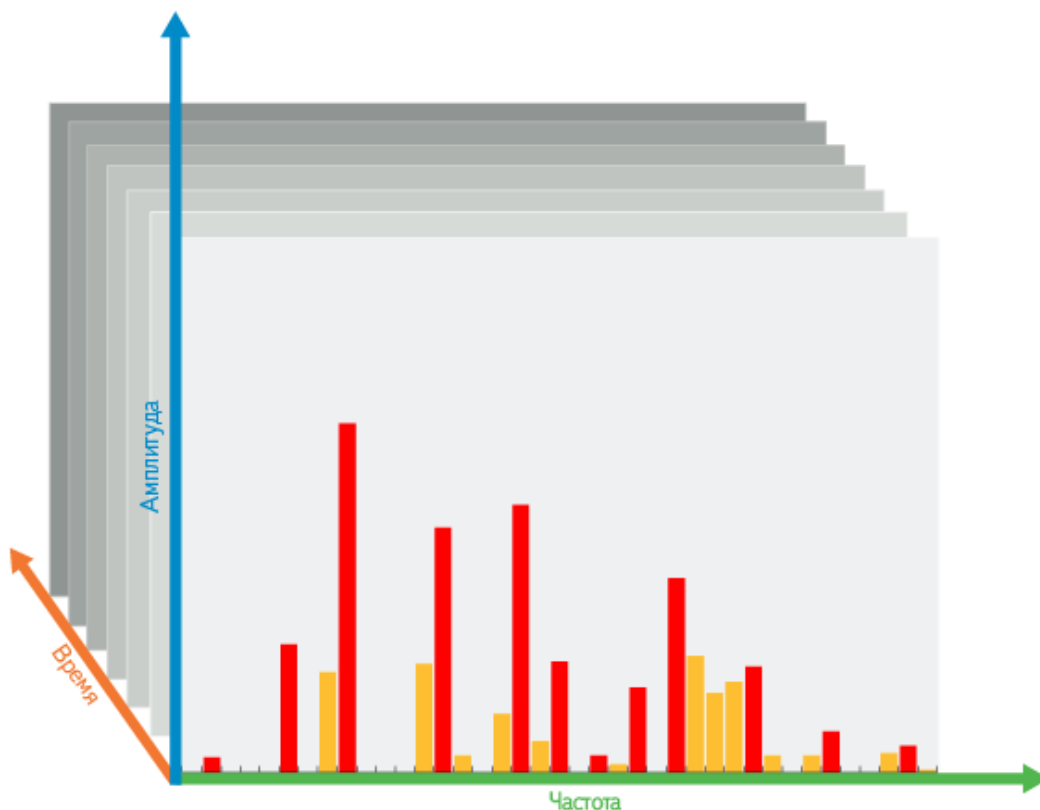


Рисунок 1.9 Контрастные пики

1.3.3 Echoprint – свободный сервис распознавания музыки

Алгоритм распознавания Echoprint основан на анализе векторов цветности нескольких последовательных сегментов, называемым “Fingerprint”.

Перв. примен.														
Справ. №														
<div>Подпись и дата</div> <div>Инв. № дубл.</div> <div>Взам. инв. №</div> <div>Подпись и дата</div> <div>Инв. № подл.</div>														
<p>Сформирована база данных известных треков, причем каждый трек Т состоит из идентификатора и метаданных (исполнителя, альбома, названия трека). Выполнение соответствия предполагает получение неизвестного аудио-запроса Q и поиск соответствующего трека в справочной базе данных.</p> <p>Для создания базы данных каждый трек разбивается на 60-секундные сегменты, причем смежные секции перекрываются на 30 секунд. Это помогает удалить предубеждение, когда более длинные песни предоставляют больше совпадений для набора хэшей запросов.</p> <p>Коды для 60-секундного сегмента представлены в качестве условий документа D в инвертированном индексе. В качестве идентификатора документа используется комбинация уникального идентификатора дорожки и номера сегмента. В нашем базовом хранилище данных используется Apache Solr с помощью специального обработчика запросов для быстрого поиска кода-запроса в список идентификаторов документов.</p> <p>30-секундный запрос содержит около 800 хеш-ключей. Сервер запросов возвращает документы с наибольшим количеством совпадений каждого кода в запросе. На практике обнаруживается, что редко один документ имеет значительно больше совпадений, чем все другие документы в индексе, однако верхние совпадения (используется 15) в этой метрике будут содержать фактическое совпадение, если оно существует. Вычисляется гистограмму всех разностей временного смещения t на соответствующий ключ в наборе результатов. Затем используется общая сумма двух верхних гистограмм, чтобы сообщить «истинный балл». Это позволяет нам гарантировать, что коды происходят в порядке, даже если Q находится из другого раздела песни и, следовательно, имеет другое абсолютное смещение по времени.</p>														
<table border="1"> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> </table>										Изм.	Лист	№ докум.	Подпись	Дата
Изм.	Лист	№ докум.	Подпись	Дата										
<div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</div>														
<div>Лист</div>														
<div>25</div>														

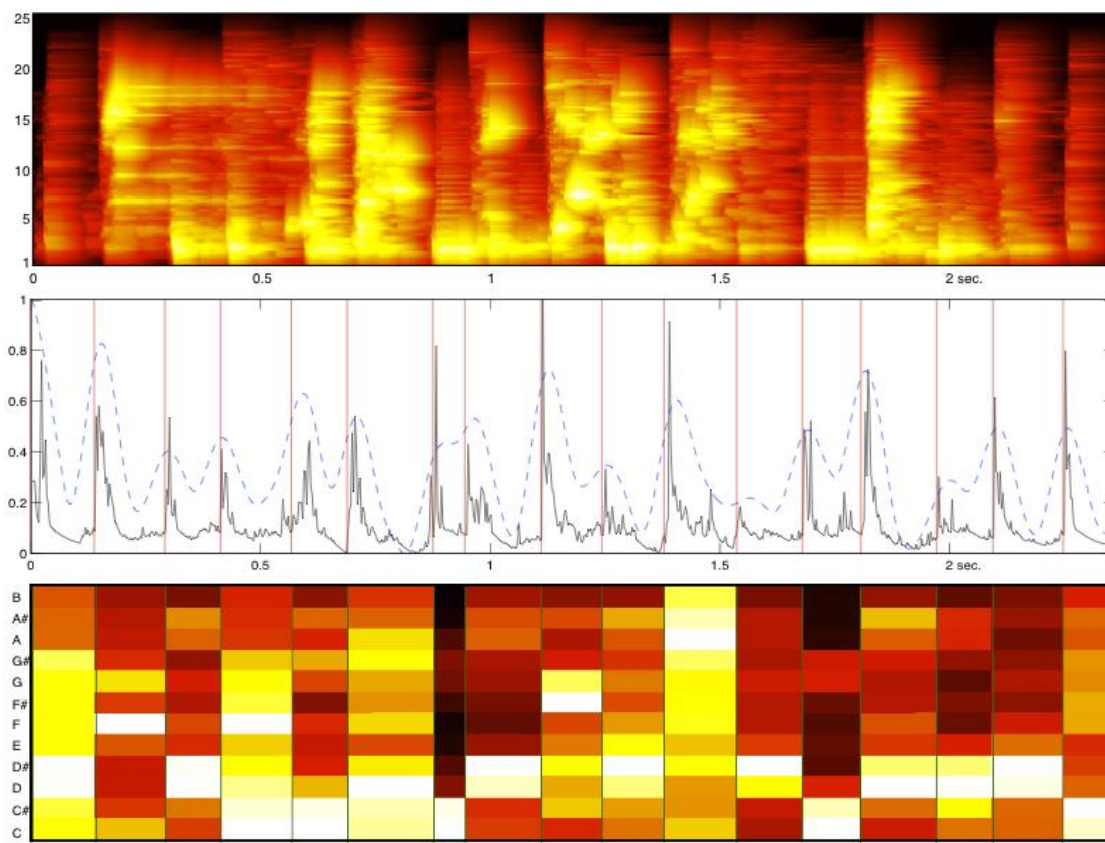


Рисунок 1.10 Пример векторов цветности для Fingerprint

Все возможные документы упорядочены по их истинному счету. Если в списке указано более одного документа с одного и того же трека, удаляются все, кроме документа, с наивысшим балл. Верхний документ в списке возвращается как положительный, если его истинный балл значительно выше, чем оценка всех других документов в списке результатов. Если промежуток между двумя верхними результатами незначителен, то возврат не возвращается.

Echoprint разработали слой базы данных для обработки сотен миллионов дорожек, разделив индекс (инвертированное хранилище документов Solr) и уровень хранения. Индекс занимает примерно 5 гигабайт дискового пространства на 100 000 треков, а для хранения требуется 15 гигабайт. Запросы могут быть выполнены на практике через 100 миллисекунд после вычисления кодовой строки для сигнала, который на текущем оборудовании занимает менее одной десятой секунды.

Перв. примен.	1.4 Выводы по главе 1				
	<p>В данной главе представлен краткий обзор существующих методов распознавания звукового сигнала. В данных обзорах приведены основные подходы к признаковому определению сигнала, к поиску схожих фрагментов в звуковом сигнале и базах данных. Большое количество работ на эту тему подтверждает актуальность решаемой задачи распознавания и всего исследования в целом.</p> <p>В результате обзора существующих методов распознавания сигналов, можно выделить ряд проблем:</p> <ul style="list-style-type: none"> - проблема формирования системы признаков, обеспечивающих однозначность классификации и последующего анализа на схожесть отдельных фрагментов сигнала. Этап выбора системы признаков необходим для сокращения размерности входного описания; - проблема величины входного сигнала. Многие алгоритмы, такие как Echoprint, анализируют музыкальный сегмент длительностью более 20 секунд для вычисления признакового описания; - проблема поиска музыкальной записи в звуковых базах данных. Она заключается в том, как быстро и удобно найти схожий с интересующим звуковой фрагмент для получения информации о нем. <p>В данном исследовании предлагается по-новому решить выделенные выше проблемы. Был выбран новый подход для признакового описания сигнала, отличающийся от описанных в обзорах. Построение системы признаков на основе Теории активного восприятия позволит сократить размерность входного описания и однозначно определить звуковой фрагмент сигнала. Также планируется разработать новый алгоритм хранения звуковых записей в базе данных. За основу взято бинарное дерево, которое упорядочит записи и позволит быстро выявлять схожие записи и предоставлять информацию и ним.</p> <p>Теоретические аспекты реализации данного алгоритма распознавания и поиска будут подробнее описаны в главе 2.</p>				
Справ. №					
Подпись и дата	Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.	<div> <div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</div> <div> <div>Изм.</div> <div>Лист</div> <div>№ докум.</div> <div>Подпись</div> <div>Дата</div> </div> </div>
					27

Глава 2. Информационная модель поиска сигнала в звуковых базах данных

2.1 Теория активного восприятия

Теория активного восприятия была предложена профессором В.А. Утробиним. Первоначально ТАВ применялась для анализа изображений, а позднее была адаптирована для применения к звуковым сигналам.

Данная теория позволяет реализовать два первых этапа распознавания - это формирование исходного описания сигнала и формирования системы признаков.

На этапе предварительного обработки сигнала происходит процесс интегрирования. Сигнал разбивается на части, затем амплитуды отсчетов для каждой части складываются (рисунок 2.1). Результатом интегрирования является массив S - множество сумм отсчетов сегмента звукового сигнала.

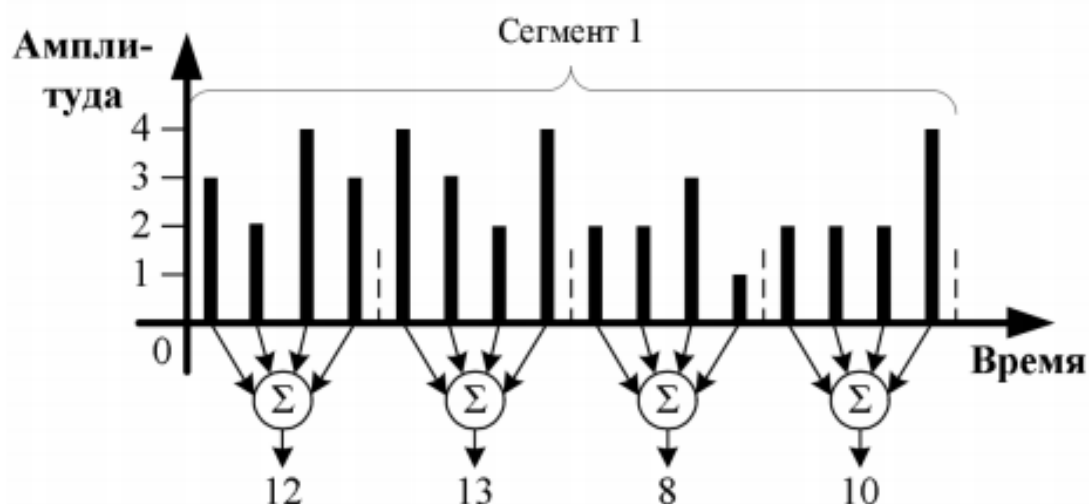
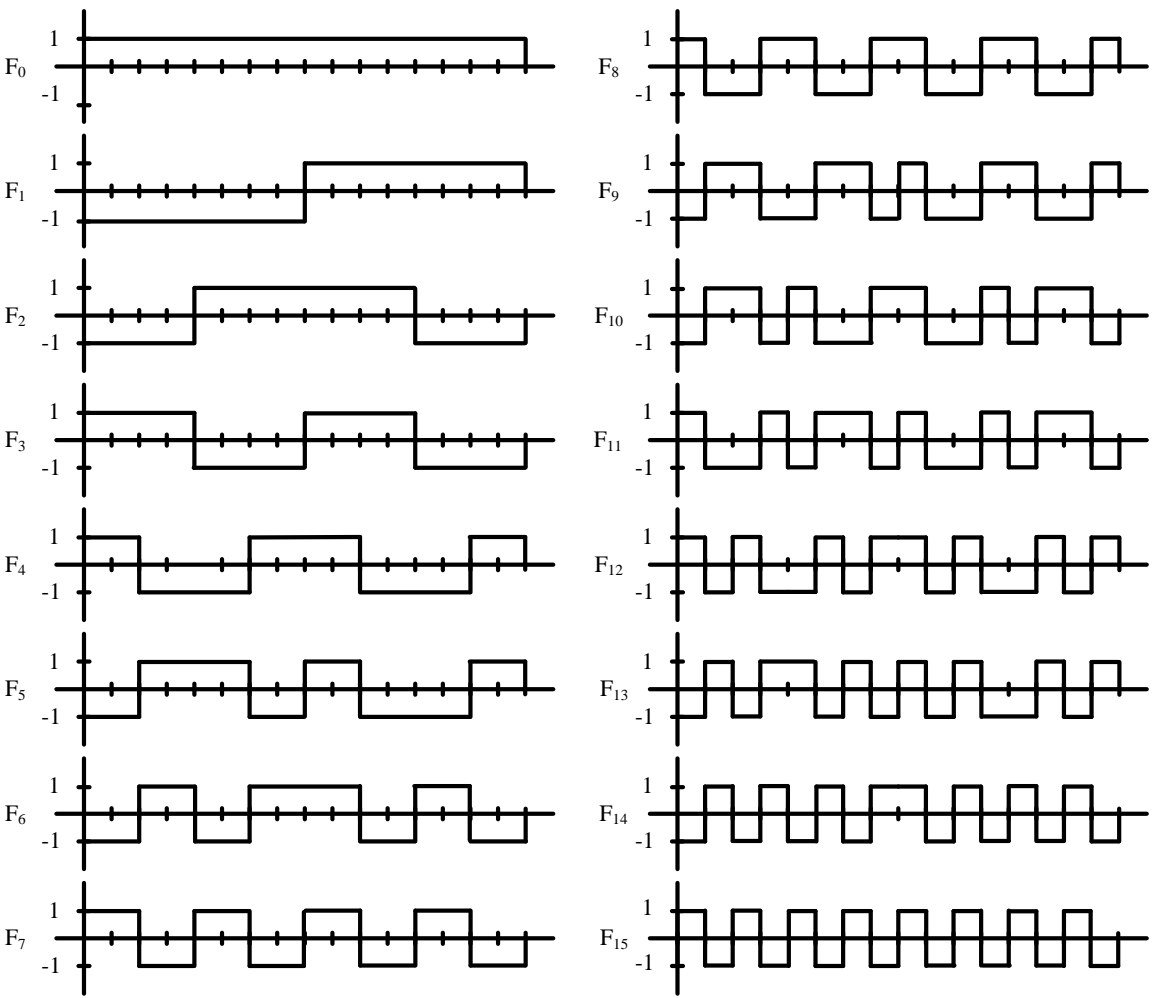


Рисунок 2.1 Интегрирование сигнала

На этапе формирования признакового описания, используется дифференцирование. К массиву сумм S применяются фильтры (рисунок 2.2), в результате чего вычисляются значения спектральных коэффициентов. Для этого каждый фильтр поэлементно умножается на массив сумм S .

Пара этих преобразований определяет U -преобразование, которое является главным в ТАВ.

Перв. примен.																
	Рисунок 2.2 Фильтры															
Справ. №	<p>Теория активного восприятия не ограничивается только формированием спектрального представления сигнала. В ее состав входит “Алгебра групп”. Алгебра групп используется для формирования описания речевого сигнала. В алгебре A_V существуют алгебраические группы (V_i - бинарные операторы, рисунок 2.3):</p> <p>1) P_{ni} (названы полными), образованы на тройках операторов (V_i, V_j, V_k), для которых справедливы соотношения: $V_i + V_j + V_k \equiv e_I$ - единица; $V_i V_j V_k$ - образ (на операции умножения; число инверсий должно быть четно) на планигоне и описание группы P_{ni}. Число возможных образов полных групп, с учётом инверсий операторов, равно четырём.</p> <p>2) P_{si} (названы замкнутыми), образованы на четверке операторов (V_i, V_j, V_p, V_m), где $(V_i, V_j, V_k) \in P_{ni}$, $(V_n, V_m, V_k) \in P_{nj}$, с описанием $V_i V_j + V_p V_m$ (где необходимое число инверсий</p>															
Подпись и дата																
Инв. № дубл.																
Взам. инв. №																
Подпись и дата																
Инв. № подл.																

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист 29
------	------	----------	---------	------	---	------------

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					


операторов нечетно) и единиц – $V_i + V_j + V_p + V_m \equiv e_1$. Число возможных образов замкнутых групп, с учётом инверсий операторов, равно восьми.

Рисунок 2.3 Бинарные операторы

С учётом инверсий всего существует 140 полных групп на операции умножения, 140 полных групп на операции сложения, 840 замкнутых групп, 840 замкнутых множеств на операции умножения и 840 замкнутых множеств на операции сложения.

В итоге, для объединения данных, полученных от разных сегментов анализируемого сигнала, вычисляется гистограмма элементов структур, использованных при создании описания сегмента.

						ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист
							30
Изм.	Лист	№ докум.	Подпись	Дата			

Перв. примен.	2.2 Система поиска сигнала в звуковой базе данных с использованием теории активного восприятия								
	<p>Данная система будет решать две функциональные задачи - это</p> <ul style="list-style-type: none">- формирование базы данных сигналов;- поиск музыкального фрагмента по существующей базе данных. <p>База данных будет представлена в виде бинарного дерева, в узлах которого будут храниться признаковые описания сегментов музыкальных композиций.</p> <p>Задача поиска объекта в базе данных будет решаться в три этапа:</p> <ul style="list-style-type: none">- предварительная обработка сигнала;- формирование системы признаков сигнала;- принятие решения. <p>К сформированной базе данных будет применен алгоритм поиска музыкальной композиции по звуковому фрагменту.</p> <p>Для входного запроса также формируется бинарное дерево путем деления пополам сегментов музыкальной композиции, применения к ним перечисленных выше этапов обработки и формирования признакового описания. Затем, на этапе принятия решения осуществляется оценка результатов поиска интересующей композиции по заданному музыкальному фрагменту и вывод результата.</p>								
Справ. №									
Подпись и дата	Рисунок 2.4 Информационная модель системы								
Инов. № дубл.	Ниже будут подробно описаны все перечисленные этапы.								
Взам. инв. №									
Подпись и дата									
Инов. № подл.									
<table><tr><td>Изм.</td><td>Лист</td><td>№ докум.</td><td>Подпись</td><td>Дата</td></tr></table>					Изм.	Лист	№ докум.	Подпись	Дата
Изм.	Лист	№ докум.	Подпись	Дата					
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)									
Лист 31									

2.2.1 Предварительная обработка сигнала

Исходный сигнал представлен в MP3 формате. Для этого формата характерным параметром является “битрейт”. Битрейт - это количество бит, используемых для хранения одной секунды мультимедийного контента, в данном случае музыкальной композиции. В форматах аудио сигнала, использующих сжатие с потерей качества, параметр “битрейт” выражает степень сжатия потока и, тем самым, определяет размер канала, для которого сжат поток данных. Чаще всего битрейт звука измеряется в килобитах в секунду [14].

На этапе предварительной обработки каждая музыкальная композиция приводится к единому “битрейту” для повышения точности распознавания системы.

Далее текущий трек делится на последовательность сегментов одинакового размера. Каждый из сегментов делится еще на 16 частей. Определяются амплитуды отсчетов данных частей и складываются в массив сумм, который в дальнейшем используется для формирования системы признаков.

$$S = \sum_{i=1}^N A_i \quad (2.1)$$

где A – амплитуда отсчета, N – количество отсчетов.

Сигнал g

Сигнал f

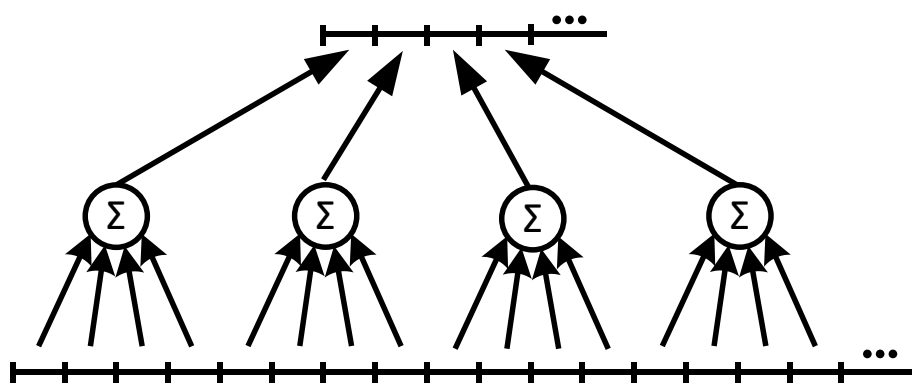


Рисунок 2.5 Схема предварительной обработки сигнала

Перв. примен.	<h3>2.2.2 Формирование системы признаков</h3> <p>После проведения этапа предварительной обработки формируется признаковое описание входного сигнала.</p> <p>Признаковое описание музыкальной композиции представляет собой набор признаковых описателей сегментов текущего сигнала, полученных на предыдущем этапе. В качестве описателя используется коэффициенты спектрального представления, полученные в результате применения U-преобразования.</p> <p>По каждому сегменту звукового сигнала вычисляется U-преобразование. Из этих признаковых описаний каждого сегмента формируется описание для полной музыкальной композиции. Коэффициенты спектрального представления, полученного в результате U-преобразования, являются описанием данного сегмента и выглядят следующим образом:</p> $Sig_i = \{S_n\}, n = \overline{1, F} \quad (2.2)$ <p>где S_n – коэффициент спектрального представления сегмента сигнала; F – количество фильтров, применяемых при вычислении U-преобразования.</p> <p>Следовательно, признаковое описание для целой музыкальной композиции будет иметь вид:</p> $Sig = \{Sig_n\} \quad (2.3)$ <p>Данное признаковое описание является достаточным для описания музыкального сигнала. Оно состоит из 16 цифр с точкой и поэтому не требует больших затрат памяти при хранении.</p>														
	Справ. №														
Подпись и дата	<h3>2.2.3 Принятие решения</h3> <p>После формирования признаковых описаний, исходный сегмент сигнала и описатель помещается в вершину бинарного дерева. Далее текущий сегмент сигнала разбивается на две равных части (рисунок 2.6).</p>														
	Инов. № дубл.														
Подпись и дата															
	Взам. инв. №														
Инов. № подл.															
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</div> <div>Лист 33</div>
Изм.	Лист	№ докум.	Подпись	Дата											

Перв. примен.	<pre> graph TD Sig1[Sig 1] --> Sig11[Sig 1.1] Sig1 --> Sig12[Sig 1.2] </pre>				
Справ. №	<p>Рисунок 2.6 Разбиение сегментов сигнала на две части</p> <p>Для каждой из получившихся частей снова запускается процесс вычисления признакового описания на основе U-преобразования. После формирования признакового описателя сегмент помещается в один из узлов-детей дерева.</p> <p>Для принятия решения, в какой узел относить текущий сегмент сигнала, используется расстояние Хэмминга.</p> <p>Расстояние Хэмминга позволяет определить число позиций, в которых символы двух одинаковых последовательностей различны [15].</p> <p>В данном случае расстояние Хэмминга заключается в посимвольном сравнении числовых массивов признакового описания с заданными шаблонами и определения количества расхождений.</p> <p>Таким образом значения признаковых описаний сегментов делятся на две группы путем сравнения с заданными шаблонами и выявления наибольшего соответствия.</p> <p>Заданные шаблоны имеют вид:</p> $T_1 = \{1,1,1,1,1,1,1,0,0,0,0,0,0\}$ $T_2 = \{0,0,0,0,0,0,0,1,1,1,1,1,1\}$ <p>Для сравнения признакового описания сегмента с шаблоном требуется провести предварительное преобразование к единому виду. Все положительные значения коэффициентов признакового описания соответствуют 1, а отрицательные - 0, как показано в примере в таблице 2.1</p>				
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист
Изм.	Лист	№ докум.	Подпись	Дата		34

Таблица 2.1 Пример приведения в соответствие с заданным шаблоном знаков коэффициентов U-преобразования

Коэффициенты	-3,54	2,31	-13,5	...	8,1
Результат	0	1	0	...	1

После распределения сегментов сигнала между двумя узлами дерева, процесс формирования бинарного дерева продолжается. К сегментам, содержащимся в текущем дочернем узле, применяются все вышеперечисленные действия, начиная с деления нового сегмента пополам и заканчивая классификацией его в последующий узел дерева. Эта процедура продолжается заданное число раз, пока бинарное дерево не станет заданного размера.

В результате получается бинарное дерево, в узлах которого хранится массив сегментов, содержащих признаковое описание, информацию о названии песни и исполнителе, идентификатор.

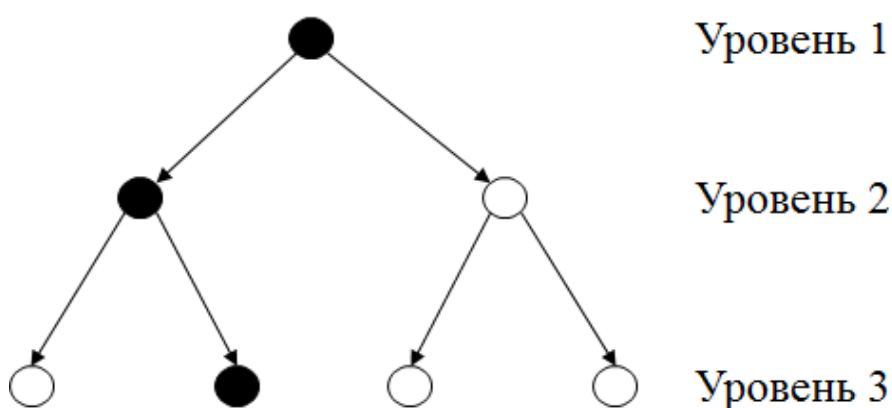
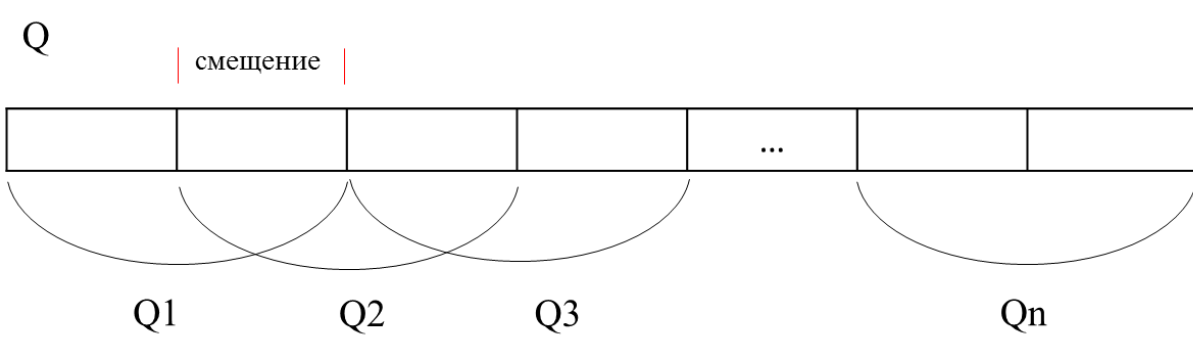


Рисунок 2.7 Размещение сегментов сигнала в узлах M-уровневого бинарного дерева

Алгоритм поиска названия композиции по заданному участку аудиозаписи заключается в переборе всех значений сегментов, находящихся на нижнем уровне бинарного дерева, и сравнения их с признаковым описанием входного фрагмента песни. Результатом этих действий будет набор последовательно идущих сегментов относящихся к одной музыкальной композиции и как следствие вывод о том, что входной запрос является фрагментом из песни, найденной в базе данных.

Данный процесс можно разделить на два этапа:

Перв. примен.	<div style="margin-bottom: 10px;"> 1. формирование признакового описания сигнала-запроса; 2. сравнение сигнала-запроса с сигналами из базы данных. </div> <p>На этапе формирования признакового описания к запрашиваемому сигналу Q применяются все преобразования описанные выше, результатом которых будет бинарное дерево запроса.</p> <p>Сигнал Q разбивается на несколько сегментов с перекрытием в несколько отсчетов (рисунок 2.8), по каждому из которых считается признаковое описание с помощью U-преобразования.</p> <div style="text-align: center; margin-top: 20px;">  </div>				
	Справ. №				
Подпись и дата		<p>Важно отметить, что величина сегмента запрашиваемого сигнала должна совпадать с величиной сегмента сигнала из готовой базы данных. В результате получается:</p> $Q = \{Q_i\} \quad (2.4)$ <p>где Q_i - признаковое описание одного сегмента сигнала-запроса.</p> <p>Далее на основе полученных признаковых описаний формируется бинарное дерево сигнала запроса. Важно отметить, что количество уровней бинарного дерева запрашиваемого сигнала должно совпадать с количеством уровней бинарного дерева базы данных.</p> <p>После окончания обработки запроса идет этап сравнения сигнала-запроса с музыкальными композициями из дерева базы данных. Для этого берется самый нижний уровень у бинарного дерева запроса и бинарного дерева базы данных.</p> <p>В базе данных каждой музыкальной композиции присваивается свой номер. После разбиения, каждому сегменту музыкальной композиции присваивается свой номер.</p>			
	Инв. № дубл.				
Взам. инв. №					
	Подпись и дата				
Инв. № подл.					

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист
						36
Изм.	Лист	№ докум.	Подпись	Дата		

Перв. примен.	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Дерево запроса</p> </div> <div style="text-align: center;"> <p>Дерево базы данных</p> </div> </div>													
	<div style="display: flex; justify-content: space-around; align-items: center;"> <table border="1" style="border-collapse: collapse;"> <tr> <td style="padding: 5px;">Q 1</td> <td style="padding: 5px;">Q 2</td> <td style="padding: 5px;">Q 3</td> <td style="padding: 5px;">...</td> <td style="padding: 5px;">Q n</td> </tr> </table> <div style="margin-left: 10px;">Запрос</div> </div> <div style="display: flex; justify-content: space-around; align-items: center; margin-top: 10px;"> <div style="text-align: center;"> <p>= = =</p> </div> <div style="text-align: center;"> <table border="1" style="border-collapse: collapse;"> <tr> <td style="padding: 5px;">Seg 1</td> <td style="padding: 5px;">Seg 2</td> <td style="padding: 5px;">Seg 3</td> <td style="padding: 5px;">...</td> <td style="padding: 5px;">Seg n</td> </tr> </table> </div> <div style="margin-left: 10px;">Экземпляр базы</div> </div>					Q 1	Q 2	Q 3	...	Q n	Seg 1	Seg 2	Seg 3	...
Q 1	Q 2	Q 3	...	Q n										
Seg 1	Seg 2	Seg 3	...	Seg n										

Рисунок 2.9 Сравнение сегментов деревьев запроса и базы данных

Далее идет сравнение признаков описаний каждого сегмента из дерева запроса с каждым сегментом из дерева базы. Важно учесть, что сравниваются только сегменты, которые находятся в соответствующих узлах в деревьях.

При обнаружении похожих признаков описаний сегментов во входящем сигнале и имеющимся в базе, проводится сравнение признаков описаний соседних сегментов сигналов.

Таким образом, находятся все последовательности сегментов, найденные в результате сравнения. Результирующий массив сегментов сортируется по номерам песен, а затем сортируются найденные сегменты, в рамках одной музыкальной композиции. После всех сортировок определяется максимальная последовательность сегментов, которая и является результатом поиска.

Инв. № подл.	Взам. инв. №	Инв. № дубл.	Подпись и дата						
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)				Лист 37

Перв. примен.															
Справ. №															
<div>2.3 Выводы по главе 2</div> <p>В данной главе были представлено теоретическое описание системы поиска сигнала в звуковой базе данных. Были предложены описания реализации каждого из этапов распознавания звукового сигнала. Ниже кратко описаны эти варианты:</p> <ul style="list-style-type: none"> - для формирования признакового описания сегментов сигнала была применена теория активного восприятия; - разработана структура <i>M</i>-уровневого бинарного дерева для хранения музыкальных композиций; - сформулирован и описан алгоритм поиска похожих сигналов в звуковой базе данных (бинарном дереве). <p>Описанные в этой главе подходы к решению каждой из задач распознавания звукового сигнала свидетельствуют о научной новизне данного исследования. Далее в главе 3 будут приведены результаты экспериментов, доказывающих работоспособность выбранного метода.</p>															
Подпись и дата															
Инв. № дубл.															
Взам. инв. №															
Подпись и дата															
Инв. № подл.															
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</div> <div>Лист 38</div>
Изм.	Лист	№ докум.	Подпись	Дата											

Перв. примен.	<h2 style="text-align: center;">Глава 3. Вычислительный эксперимент</h2> <p>Важной частью данного научного исследования является вычислительный эксперимент. На основании результатов, полученных в ходе проведения испытаний, можно сделать вывод о том, насколько хорошо данный метод справляется с поставленными задачами по сравнению с известными аналогами. В данной главе будет представлена информация о проводимом эксперименте и проведен анализ его результатов.</p>				
	Справ. №	<h3 style="text-align: center;">3.1 Описание эксперимента</h3> <p>В ходе проведения вычислительного эксперимента будет тестироваться эффективность распознавания музыкальных композиций и устойчивость алгоритма к внешним искажениям предложенного алгоритма.</p> <p>Звуковая база данных содержит 1000 различных композиций формата MP3, записанных в студийных условиях. Длительность музыкальных композиций составляет от 2,5 до 3 и более минут.</p> <p>В качестве входного параметра алгоритма будут использованы 1000 различных звуковых запроса, так же записанных в формате MP3. Испытания будут проводиться на запросах, полученных из студийных записей композиций, и с наложением искажений.</p> <p>Анализируется работа алгоритма в нормальных условиях. Нормальные условия выражены в отсутствии шума, а все запросы записаны в студийных условиях. В качестве изменяемых критериев будут взяты:</p> <ol style="list-style-type: none"> Число уровней разложения звуковой базы данных. По данному критерию взято 3 значения величины дерева базы данных - 3, 5, 7 уровней разложения. Величина сегмента сигнала, на который будут дробиться песни, содержащиеся в звуковой базе данных. Величины звуковых сегментов - 1 сек., 2 сек., 4 сек. Величина входного сигнала запроса. Длительность входных запросов совпадает с длительностью сегментов дробления базы - 1 сек., 2 сек., 4 сек. Величина смещения сегмента при разложении входного сигнала. Смещение в отсчетах взято - 10, 100 и 1000 отсчетов. 			
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.
	Изм.				

Перв. примен.															
Справ. №															
Подпись и дата															
Инв. № дубл.															
Взам. инв. №															
Подпись и дата															
Инв. № подл.															
<p>В ходе эксперимента выявлены лучшие конфигурации при которых система выдает максимальную эффективность. Для этих конфигураций проводится анализ устойчивости системы к шумам.</p> <p>Путем наложения искажения на входные запросы анализируется устойчивость алгоритма распознавания к звуковым искажениям во внешней среде. В качестве искажения будет выступать нормальный шум. Величины шума - 0 дБ, 10 дБ, 20 дБ.</p> <p>На выходе алгоритма могут быть следующие варианты результатов поиска:</p> <ol style="list-style-type: none"> 1. Найдено верное значение входной звуковой композиции, присутствующей в базе. 2. Найдено неверное значение входной звуковой композиции, присутствующей в базе. 3. Значение композиции, присутствующей в базе, не найдено. <p>Вариант 1 соответствует правильной работе исследуемой системы. Вариант 2 относится к ошибке 1-ого рода (ложноположительное срабатывание), вариант 3 - ошибка 2-ого рода (ложноотрицательное срабатывание).</p> <p>В результате проведения вычислительного эксперимента будут анализироваться:</p> <ol style="list-style-type: none"> 1. Точность распознавания звукового запроса; 2. Время формирования звуковой базы данных для различных уровней разложения; 3. Время потраченное обработку одного звукового запроса; 					Лист										
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<p>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</p> <p>40</p>
Изм.	Лист	№ докум.	Подпись	Дата											

Перв. примен.	<h3>3.2 Описание программного продукта</h3>																																															
	<p>Для проведения вычислительного эксперимента по данному исследованию был разработан программный продукт на языке R. Данный язык был выбран по причине того, что он обладает широким набором инструментов для сбора статистических данных и расширяемостью за счет установки дополнительных программных пакетов.</p> <p>Далее будут представлены программные модули, которые реализуют весь функционал системы.</p> <p>1. <i>createTree</i> - функция, реализующая создание дерева базы данных</p>																																															
Справ. №	<p>Таблица 3.1 Описание функции <i>createTree</i></p>																																															
	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 20%;">Тип параметра</th> <th style="width: 20%;">Название</th> <th style="width: 80%;">Значение</th> </tr> <tr> <td rowspan="3" style="text-align: center;">Входной</td> <td><i>segmentSize</i></td> <td>Величина сегмента разбиения</td> </tr> <tr> <td><i>offset</i></td> <td>Смещение сегмента в сигнале</td> </tr> <tr> <td><i>levelsCount</i></td> <td>Количество уровней разложения дерева</td> </tr> <tr> <td style="text-align: center;">Выходной</td> <td><i>tree</i></td> <td>Дерево сегментов</td> </tr> </table>					Тип параметра	Название	Значение	Входной	<i>segmentSize</i>	Величина сегмента разбиения	<i>offset</i>	Смещение сегмента в сигнале	<i>levelsCount</i>	Количество уровней разложения дерева	Выходной	<i>tree</i>	Дерево сегментов																														
Тип параметра	Название	Значение																																														
Входной	<i>segmentSize</i>	Величина сегмента разбиения																																														
	<i>offset</i>	Смещение сегмента в сигнале																																														
	<i>levelsCount</i>	Количество уровней разложения дерева																																														
Выходной	<i>tree</i>	Дерево сегментов																																														
Подпись и дата	<p>2. <i>createQueryTree</i> - функция, отвечающая за разложение входного сигнала-запроса в дерево</p>																																															
	<p>Таблица 3.2 Описание функции <i>createQueryTree</i></p>																																															
Инв. № дубл.	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 20%;">Тип параметра</th> <th style="width: 20%;">Название</th> <th style="width: 80%;">Значение</th> </tr> <tr> <td rowspan="4" style="text-align: center;">Входной</td> <td><i>signal</i></td> <td>Входной запрос</td> </tr> <tr> <td><i>segmentSize</i></td> <td>Величина сегмента разбиения</td> </tr> <tr> <td><i>offset</i></td> <td>Смещение сегмента в сигнале</td> </tr> <tr> <td><i>levelsCount</i></td> <td>Количество уровней разложения дерева</td> </tr> <tr> <td style="text-align: center;">Выходной</td> <td><i>tree</i></td> <td>Дерево сегментов</td> </tr> </table>					Тип параметра	Название	Значение	Входной	<i>signal</i>	Входной запрос	<i>segmentSize</i>	Величина сегмента разбиения	<i>offset</i>	Смещение сегмента в сигнале	<i>levelsCount</i>	Количество уровней разложения дерева	Выходной	<i>tree</i>	Дерево сегментов																												
	Тип параметра	Название	Значение																																													
Входной	<i>signal</i>	Входной запрос																																														
	<i>segmentSize</i>	Величина сегмента разбиения																																														
	<i>offset</i>	Смещение сегмента в сигнале																																														
	<i>levelsCount</i>	Количество уровней разложения дерева																																														
Выходной	<i>tree</i>	Дерево сегментов																																														
Взам. инв. №	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td style="text-align: center;">Изм.</td> <td style="text-align: center;">Лист</td> <td style="text-align: center;">№ докум.</td> <td style="text-align: center;">Подпись</td> <td style="text-align: center;">Дата</td> <td colspan="7" style="text-align: center; padding: 5px;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </td> <td style="width: 50px; text-align: center; vertical-align: middle;">Лист</td> </tr> <tr> <td colspan="11"></td> <td style="text-align: center;">41</td> </tr> </table>																	Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист												41						
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист																																				
											41																																					
Подпись и дата	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td style="text-align: center;">Изм.</td> <td style="text-align: center;">Лист</td> <td style="text-align: center;">№ докум.</td> <td style="text-align: center;">Подпись</td> <td style="text-align: center;">Дата</td> <td colspan="7" style="text-align: center; padding: 5px;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </td> <td style="width: 50px; text-align: center; vertical-align: middle;">Лист</td> </tr> <tr> <td colspan="11"></td> <td style="text-align: center;">41</td> </tr> </table>																	Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист												41						
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист																																				
											41																																					
Инв. № подл.	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td style="text-align: center;">Изм.</td> <td style="text-align: center;">Лист</td> <td style="text-align: center;">№ докум.</td> <td style="text-align: center;">Подпись</td> <td style="text-align: center;">Дата</td> <td colspan="7" style="text-align: center; padding: 5px;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </td> <td style="width: 50px; text-align: center; vertical-align: middle;">Лист</td> </tr> <tr> <td colspan="11"></td> <td style="text-align: center;">41</td> </tr> </table>																	Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист												41						
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист																																				
											41																																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td style="text-align: center;">Изм.</td> <td style="text-align: center;">Лист</td> <td style="text-align: center;">№ докум.</td> <td style="text-align: center;">Подпись</td> <td style="text-align: center;">Дата</td> <td colspan="7" style="text-align: center; padding: 5px;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </td> <td style="width: 50px; text-align: center; vertical-align: middle;">Лист</td> </tr> <tr> <td colspan="11"></td> <td style="text-align: center;">41</td> </tr> </table>																								Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист												41
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)							Лист																																				
											41																																					

3. *addChildren* - функция добавления узла в деревоТаблица 3.3 Описание функции *addChildren*

Тип параметра	Название	Значение
Входной	<i>node</i>	Узел дерева
	<i>levels</i>	Количество уровней разложения
	<i>indexation</i>	Индексация сегмента сигнала
Выходной	<i>indexation</i>	Индексация сегмента сигнала

4. *getHammingDistance* - функция вычисления расстояния Хэмминга для коэффициентов U-преобразования и шаблонаТаблица 3.4 Описание функции *getHammingDistance*

Тип параметра	Название	Значение
Входной	<i>signalCoefs</i>	Значения коэффициентов признакового описания сегмента
Выходной	<i>distance</i>	Число различных символов при сравнении коэффициентов признакового описания и шаблона

5. *findSignal* - функция поиска похожих сегментов в дереве запроса и дереве базы данных

Таблица 3.5 Описание функции *findSignal*

Тип параметра	Название	Значение
Входной	<i>node</i>	Значение узла дерева базы данных
	<i>targetNode</i>	Значения узла дерева запроса
	<i>results</i>	Результат поиска схожих сегментов из предыдущего набора узлов
Выходной	<i>results</i>	Результат поиска схожих сегментов из текущего набора узлов

6. *isEqual* - функция, сравнивающая признаковые описания сегментов

Таблица 3.6 Описание функции *isEqual*

Тип параметра	Название	Значение
Входной	<i>coef1</i>	Значение коэффициента сигнала в дереве запроса
	<i>coef2</i>	Значение коэффициента сигнала в дереве базы данных
Выходной	<i>result</i>	Результат сравнения коэффициентов

В результате работы всех описанных выше модулей на выходе системы получается название искомой композиции композиции.

Изм.	Лист	№ докум.	Подпись	Дата

3.3 Анализ результатов экспериментов

Сначала тестирование проводится для работы алгоритма в нормальных условиях. Результаты работы системы в нормальных условиях приведены в таблице 3.1.

Таблица 3.1 Результаты работы системы поиска сигнала в звуковых базах в нормальных условиях

Кол-во уровней дерева	Длительность запроса, с	Величина сегмента, с	Величина смещения в запросе, %	Точность поиска, %	Время поиска запроса, с	Время формирования базы данных, с
3	1	1	1000	12,5	0,21	2,4
			100	25	0,32	
			10	37,5	0,39	
	2	1	1000	8,3	0,45	3,4
			100	12,5	0,51	
			10	16,7	0,56	
		2	1000	4,2	0,22	2,5
			100	29,2	0,26	
			10	45,8	0,31	
	4	1	1000	4,2	0,67	3,7
			100	8,3	0,7	
			10	12,5	0,78	
		2	1000	8,3	0,44	2,6
			100	16,7	0,49	
			10	20,8	0,54	
		4	1000	8,3	0,21	1,8
			100	20,8	0,3	
			10	41,7	0,34	

Кол-во уровней дерева	Длитель- ность запроса, с	Величина сегмента, с	Величи- на смещения в запросе, %	Точность поиска, %	Время поиска запроса, с	Время формирова- ния базы данных, с	
5	1	1	1000	29,2	0,32	5,7	
			100	50	0,39		
			10	75	0,45		
	2	1	1000	12,5	0,55	5,4	
			100	25	0,58		
			10	33,3	0,49		
		2	1000	12,5	0,33	6,6	
			100	33,3	0,37		
			10	62,5	0,43		
		4	1	1000	16,7	0,71	5,1
				100	25	0,74	
				10	33,3	0,79	
	2		1000	29,2	0,59	6,6	
			100	37,5	0,53		
			10	50	0,48		
4	4		1000	33,3	0,34	5,9	
			100	58,3	0,39		
			10	79,2	0,43		

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

Таблица 3.1 Результаты работы системы поиска сигнала в звуковых базах в нормальных условиях

Кол-во уровней дерева	Длительность запроса, с	Величина сегмента, с	Величина смещения в запросе, %	Точность поиска, %	Время поиска запроса, с	Время формирования базы данных, с
7	1	1	1000	50	1,3	7,6
			100	66,6	1,38	
			10	83,3	1,46	
	2	1	1000	50	1,29	8,6
			100	62,5	1,34	
			10	79,2	1,39	
		2	1000	62,5	0,91	7,4
			100	79,2	0,95	
			10	91,6	1,01	
	4	1	1000	33,3	1,38	8,8
			100	41,7	1,41	
			10	62,5	1,47	
2		1000	50	1,19	7,9	
		100	62,5	1,22		
		10	79,2	1,25		
4		1000	75	0,86	6,7	
		100	83,3	0,9		
		10	95,8	0,94		

На основании данных из таблицы 3.1 можно сделать вывод, что наилучшая конфигурация для системы поиска сигнала в звуковых базах данных достигается при параметрах:

- Число уровней разложения дерева – 7;
- Длительность входного запроса – 4 с;
- Величина сегмента разбиения в базе данных – 4 с;
- Величина смещения в запросе – 10 отсчетов.

Изм.	Лист	№ докум.	Подпись	Дата

ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)

Лист
46

Перв. примен.	использует программное обеспечение <i>Shazam</i> , при аналогичных уровнях шума представлены в таблице 3.3.																
	Таблица 3.3 Сравнение точностей распознавания в условиях шума для разработанного алгоритма и алгоритма распознавания <i>Shazam</i>																
Справ. №	<table border="1"> <tr> <th>Уровень шума, дБ</th> <th>Разработанный метод, %</th> <th><i>Shazam</i>, %</th> </tr> <tr> <td>0</td> <td>87,5</td> <td>80</td> </tr> <tr> <td>10</td> <td>91,6</td> <td>95</td> </tr> <tr> <td>20</td> <td>95,8</td> <td>99</td> </tr> </table>					Уровень шума, дБ	Разработанный метод, %	<i>Shazam</i> , %	0	87,5	80	10	91,6	95	20	95,8	99
	Уровень шума, дБ	Разработанный метод, %	<i>Shazam</i> , %														
	0	87,5	80														
	10	91,6	95														
20	95,8	99															
По результатам, собранным в таблицах 3.1 и 3.2, можно сказать о том, что разработанный алгоритм поиска сигналов в звуковых базах данных полностью справляется с поставленной задачей.																	
На основании таблиц 3.3 можно сделать вывод, что разработанный метод является конкурентоспособным с существующими методами.																	
Подпись и дата	Инов. № дубл.	Взам. инв. №	Подпись и дата	Инов. № подл.													
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата								
Изм.	Лист	№ докум.	Подпись	Дата													
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)					Лист												
					48												

Перв. примен.	<h3 style="text-align: center;">3.4 Выводы по главе 3</h3> <p>В данной главе описан вычислительный эксперимент, проведенный по исследованию системы поиска сигнала в звуковых базах данных, и анализу полученных результатов. Помимо этого, в этой главе дано краткое описание программного продукта, написанного для проведения тестирования на языке R.</p> <p>В результате проведения эксперимента в нормальных условиях были выявлены наилучшие конфигурации системы, при которых алгоритм выдает максимальную точность распознавания. Для этих конфигураций были проведены испытания с наложением шума на музыкальную композицию. Были получены следующие значения точности распознавания сигнала в звуковой базе данных:</p> <p>Данные результаты говорят об устойчивости данного метода к внешним воздействиям окружающих шумов.</p> <p>Полученные результаты были сравнены с результатами работы известных алгоритмов. Было установлено, что разработанный метод имеет сопоставимую с результатами распознавания существующих методов точность распознавания. На основании этого можно сделать вывод о конкурентоспособности разработанного метода.</p>					
						Справ. №
<div>Подпись и дата</div> <div>Инв. № дубл.</div> <div>Взам. инв. №</div> <div>Подпись и дата</div> <div>Инв. № подл.</div>						<div>Лист</div> <div>49</div>
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	

Перв. примен.	<h2 style="text-align: center;">Заключение</h2> <p>В данной работе рассмотрен подход к построению модели и алгоритма поиска сигнала в звуковых базах данных. На практике данная разработка может применяться для поиска интересующих звуковых фрагментов. Например, в криминалистике, для установления сходства двух голосов. Либо в музыкальной индустрии для выявления плагиата среди исполнителей и композиторов.</p> <p>В отличие от известных методов поиска, которые в основном используют грубый перебор и сравнение признаков сигнала, в данном алгоритме идет последовательное, сначала грубое потом с каждым шагом более точное, нахождение звукового фрагмента в базе данных.</p> <p>Данная модель предполагает идентификатор сигнала, в который входит не только его признаковое описание, сформированное с помощью U-преобразования теории активного восприятия, но и адрес узла в структуре бинарного дерева, так называемый хэш. Бинарное дерево используется для формирования базы данных звуковых сигналов и помогает при поиске схожих звуковых фрагментов.</p> <p>Для исследования работоспособности и эффективности разработанного метода была написана программа на языке R и проведен ряд испытаний.</p> <p>По результатам эксперимента были установлены наилучшие конфигурации работы системы для нормальных условий и в условиях зашумленного сигнала. Полученная точность распознавания сигнала соответствует корректной работе разработанного метода при решении задач поиска сигнала в звуковых базах данных. Также было проведено сравнение полученных результатов распознавания с результатами известных алгоритмов. Из полученного сравнения можно сделать вывод, что данный алгоритм является конкурентноспособным.</p> <p>Исходя из сказанного выше можно сделать вывод, что задачи, поставленные в начале научного исследования, были в полной мере выполнены, а цель работы - достигнута.</p>					
						Справ. №
Подпись и дата	Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.		
Изм.	Лист	№ докум.	Подпись	Дата	<div style="text-align: center;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </div>	<div style="text-align: center;"> Лист 50 </div>

Перв. примен.	Список литературы																		
	<p>1. Jansen A., Church K., Hermansky H. Towards Spoken Term Discovery At Scale With Zero Resources // Human Language Technology Center of Excellence, Department of Electrical and Computer Engineering, Department of Computer Science Johns Hopkins University, Baltimore, Maryland - INTERSPEECH 2010, 1676-1679 с.</p> <p>2. Hazen T.J., Shen W., White C. Query-By-Example Spoken Term Detection Using Phonetic Posteriorgram Templates // MIT Lincoln Laboratory Lexington, Massachusetts, USA Johns Hopkins University Baltimore, Maryland, USA</p> <p>3. Преобразование Хафа – Википедия – URL: https://ru.wikipedia.org/wiki/Преобразование_Хафа</p> <p>4. Shazam: алгоритмы распознавания музыки, сигнатуры, обработка данных – Habrahub – URL: https://habrahabr.ru/company/wunderfund/blog/275043/</p> <p>5. Как Яндекс распознаёт музыку с микрофона – Habrahub – URL: https://habrahabr.ru/company/yandex/blog/181219/</p> <p>6. Wang A. An Industrial-Strength Audio Search Algorithm // Proc. 2003 ISMIR International Symposium on Music Information Retrieval, Baltimore, MD, Oct. 2003.</p> <p>7. Ellis D.P.W, Whitman B., Porter A. Echoprint – an open music identification service // LabROSA Columbia University, The Echo Nest, CIRMMT / Schulich School of Music McGill University</p> <p>8. Ellis D.P.W., Whitman B., Jehan T., Lamere P. The Echo Nest musical fingerprint. In Proceedings of the 2010 International Symposium on Music Information Retrieval, 2010.</p> <p>9. Гай В.Е. Метод обработки биометрической информации в системах контроля и управления доступом // Труды Нижегородского государственного технического университет им. Р.Е. Алексеева: НГТУ им. Р.Е. Алексеева. – Нижний Новгород, 2013, № 8 (94). – С. 65-32.</p> <p>10. Гай В.Е. Метод оценки частоты основного тона в условиях помех // Цифровая обработка сигналов, № 4, 2013, С. 65-71</p> <p>11. Гай В.Е. Идентификация диктора по спектральным характеристикам речевого сигнала (статья) // Датчики и системы, № 7, 2013, С. 40-45</p> <p>12. Утробин В.А. Информационные модели системы зрительного восприятия для задач компьютерной обработки изображений: учеб. пособие / В.А. Утробин. –</p>																		
Справ. №																			
Подпись и дата																			
Инв. № дубл.																			
Взам. инв. №																			
Подпись и дата																			
Инв. № подл.																			
<table border="1" style="width: 100%;"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<table border="1" style="width: 100%;"> <tr> <td colspan="2" style="text-align: center;">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)</td> </tr> <tr> <td colspan="2" style="text-align: center;">Лист 51</td> </tr> </table>	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)		Лист 51	
Изм.	Лист	№ докум.	Подпись	Дата															
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)																			
Лист 51																			

Перв. примен.	<p>Нижний Новгород: НГТУ им. Р. Е. Алексеева, 2001. – 234 с.</p> <p>13. Утробин В.А. Элементы теории активного восприятия изображений / В.А. Утробин // Труды Нижегородского государственного технического университета им. Р.Е. Алексеева. – 2010. – Т. 81. – № 2. – С. 61-69.</p> <p>14. Утробин В.А. Физические интерпретации элементов алгебры изображения // Успехи физических наук, Т. 174, № 10, 2004, С. 1089–1104.</p> <p>15. Загоруйко Н.Г. Методы распознавания и их применение / Н.Г. Загоруйко. – М.: Советское радио, 1972. – 208 с.</p> <p>16. Битрейт – Википедия – URL: https://ru.wikipedia.org/wiki/Битрейт</p> <p>17. Расстояние Хэмминга – Википедия – URL: https://ru.wikipedia.org/wiki/Расстояние_Хэмминга</p> <p>18. Faster Gibbs sampling MCMC from within R – Darren Wilkinson's research blog – URL: https://darrenjw.wordpress.com/2011/07/31/faster-gibbs-sampling-mcmc-from-within-r/</p> <p>19. Gai V., Polyakov I. Sound Identification Model in Terms of the Active Perception Theory // Institute of Radioelectronics and Information Technologies Nizhny Novgorod State Technical Universtiy n.a. R.E. Alekseev</p>																							
	Справ. №																							
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.																			
<table border="1"> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td rowspan="3"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ) </td> <td>Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td>52</td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td></td> </tr> </table>											ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист						52	Изм.	Лист	№ докум.	Подпись	Дата	
					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-013-2017 (ПЗ)	Лист																		
						52																		
Изм.	Лист	№ докум.	Подпись	Дата																				