

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ им.
Р. Е. АЛЕКСЕЕВА»

МОДЕЛИ И МЕТОДЫ РАСПОЗНАВАНИЯ ОДНОМЕРНЫХ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

*Рекомендовано Учёным советом Нижегородского государственного
технического университета им. Р. Е. Алексеева в качестве учебного пособия
для студентов, обучающихся по направлениям подготовки 09.03.01,
09.04.01 – «Информатика и вычислительная техника»*

Нижний Новгород 2019

УДК 528.854
М 744

Авторы: В. Е. Гай, И. В. Поляков, М. О. Дербасов, О. В. Андреева

Рецензент
доктор технических наук, профессор Ю. С. Федосенко

М 744 **Модели и методы распознавания одномерных сигналов и изображений: учебное пособие / В. Е. Гай, [и др.]; Нижегород. гос. техн. ун-т. им. Р. Е. Алексеева. – Н. Новгород, 2019. – 108 с.**

ISBN 978-5-502-01162-4

Рассматривается теоретический подход к решению задач распознавания изображений и одномерных сигналов. Приводятся примеры решения задач распознавания. Предназначено для студентов всех форм обучения по направлениям 09.03.01, 09.04.01 – «Информатика и вычислительная техника»

Рис. 53. Табл. 14. Библиогр.: 8 назв.

УДК 528.854

ISBN 978-5-502-01162-4

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
1. СТРУКТУРА СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ	6
2. ФОРМИРОВАНИЕ СПЕКТРАЛЬНОГО ПРЕДСТАВЛЕНИЯ ЗВУКОВОГО СИГНАЛА С ПОЗИЦИЙ ТАВ	8
3. АЛГЕБРА ГРУПП	15
4. ПОЛНЫЕ ГРУППЫ	20
5. ЗАМКНУТЫЕ ГРУППЫ	24
6. ВЫБОР РАЗМЕРНОСТИ ПРИЗНАКОВОГО ОПИСАНИЯ	31
7. ФОРМИРОВАНИЕ ПРИЗНАКОВОГО ОПИСАНИЯ ЗВУКОВОГО СИГНАЛА	33
7.1. Гистограммы операторов, полных, замкнутых групп	33
7.2. Матрицы вероятностей переходов	35
7.3. Выбор систем признаков для описания сигналов	39
8. РЕШЕНИЕ ЗАДАЧ ОБРАБОТКИ И РАСПОЗНАВАНИЯ ОДНОМЕРНЫХ СИГНАЛОВ	42
8.1. Выбор длины сегмента сигнала	42
8.2. Оценка частоты основного тона	44
8.3. Оценка отношения сигнал / шум	47
8.4. Выделение активности голоса	51
8.5. Оценка стабильности описания одномерного сигнала	54
8.6. Фильтрация одномерного сигнала	56
8.6.1. Алгоритм фильтрации 1	56
8.6.2. Алгоритм фильтрации 2	59
8.7. Сжатие звукового сигнала	59
8.7.1. Алгоритм сжатия 1	59
8.7.2. Алгоритм сжатия 2	60
8.8. Алгоритм сравнения сигналов	60
8.9. Поиск сигналов в звуковых базах данных	65
8.9.1. Предварительная обработка и признаковое описание сигнала	66
8.9.2. Информационная модель системы хранения и поиска	67
8.9.3. Размещение сегментов в двоичном дереве	67
8.9.4. Формирование ключа по сегменту сигнала	68
8.9.5. Анализ возможных коллизий ключей поиска	69
8.9.6. Поиск звукового сигнала в базе данных	70
8.10. Прогнозирование времени отказа подшипника качения	71

9. РЕШЕНИЕ ЗАДАЧ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ.....	75
9.1. Обнаружение объектов на изображении с использованием локального признакового описания	75
9.2. Поиск похожих людей по фотографии лица, идентификация человека по фотографии лица.....	78
9.3. Обнаружение вредоносного воздействия на систему видеоконтроля	81
9.4. Формирование карты глубины на основе стереопары	83
9.4.1. Общий принцип работы методов построения карт глубины по стереопаре	84
9.4.2. Метод построения карты глубины	85
9.4.3. Сегментация.....	85
9.4.4. Выделение горизонтальных сегментов	86
9.4.5. Выделение границ объектов	86
9.4.6. Сопоставление сегментов.....	87
9.4.7. Формирование карты несоответствий	87
9.4.8. Построение карты глубины.....	88
9.4.9. Вычислительный эксперимент	88
9.5. Моделирование ассоциативной памяти	89
9.6. Распознавание изображений текстур	93
9.7. Обнаружение объектов на изображении с использованием глобального признакового описания	95
10. ЗАДАНИЯ ДЛЯ ВЫПОЛНЕНИЯ	102
10.1. Методика проведения эксперимента	102
10.2. Варианты заданий.....	102
10.2.1. Решение практических задач	102
10.2.2. Разработка новых систем признаков.....	104
ЗАКЛЮЧЕНИЕ	106
СПИСОК ЛИТЕРАТУРЫ.....	107

ВВЕДЕНИЕ

С ростом автоматизации производства возникла необходимость обеспечить работу многих технологических процессов без участия человека. Главным образом это процессы, связанные с рутинной, однообразной работой или опасные для человека. В таких процессах человека целесообразно заменить на автоматы, реагирующие на различные отклонения параметров технологического процесса от номинальных. Создание таких автоматов являлось первым шагом на пути к построению распознающих систем. Со временем такие автоматы становились все сложнее, а с появлением компьютеров, открылись широчайшие возможности для применения распознающих автоматов, работа которых основывается на цифровой обработке данных.

Не только указанная замена и освобождение человека от выполнения рутинных операций является причиной поиска путей создания систем распознавания образов. В некоторых случаях человек вообще не в состоянии решать поставленную задачу со скоростью, задаваемой обстоятельствами, независимо от качеств и психологического состояния, принимающего решение.

В течение достаточно продолжительного времени проблема распознавания образов привлекает внимание специалистов в области прикладной математики, а затем и информатики. В частности, можно отметить работы Р. Фишера, выполненные в 20-х годах и приведшие к формированию дискриминантного анализа, как одного из разделов теории и практики распознавания образов. В 40-х годах А. Н. Колмогоровым и А. Я. Хинчиным поставлена задача о разделении смеси двух распределений. Наиболее плодотворными явились 50-60-е годы XX века. В это время на основе массы работ появилась теория статистических решений. В результате этого появления найдены алгоритмы, обеспечивающие отнесение нового объекта к одному из заданных классов, что явилось началом планомерного научного поиска и практических разработок. В рамках кибернетики начало формироваться новое научное направление, связанное с разработкой теоретических основ и практической реализации устройств, а затем и систем, предназначенных для распознавания объектов, явлений, процессов. Иными словами, появилась новая научная дисциплина, которая получила название «Распознавание образов». Несмотря на многолетние усилия исследователей задача распознавания изображений и одномерных сигналов остается не решенной до сих пор.

Настоящее учебное пособие содержит описание теоретических и практических аспектов решения разноплановых задач распознавания образов и будет полезно студентам, аспирантам и инженерам, интересующимся данными задачами.

1. СТРУКТУРА СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ

С позиций системного анализа структуру системы распознавания можно представить в виде трёх этапов (рис. 1).

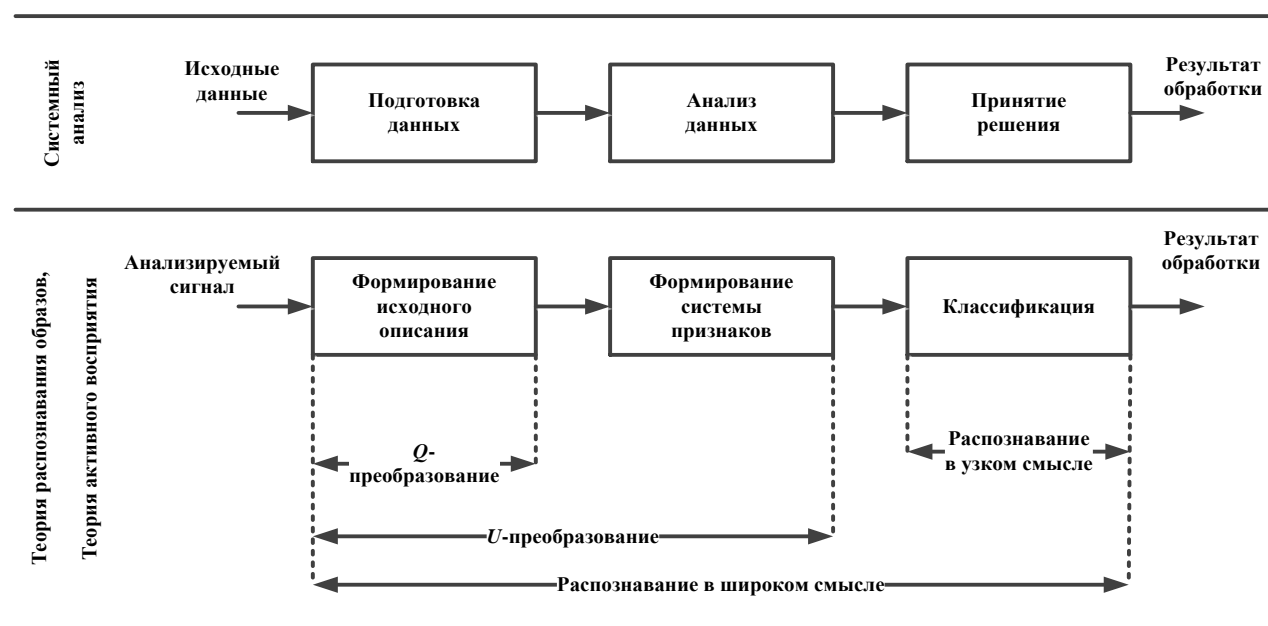


Рис. 1. Структура системы распознавания

Этап предварительной обработки обычно состоит в выполнении фильтрации сигнала, а также, в ряде случаев, в делении исходного сигнала на сегменты, по которым в дальнейшем вычисляются признаки. Для создания описания входного сигнала вычисляются следующие признаки: коэффициенты спектра Фурье, кепстральные коэффициенты, мел-частотные кепстральные коэффициенты, коэффициенты линейного предсказания, коэффициенты вейвлет-спектра и т. д. Этап принятия решения заключается в отнесении входного сигнала по имеющимся признакам к одному из известных классов. На этапе принятия решения используются нейронные сети, метод опорных векторов, решающие деревья, метод K -ближайших соседей.

При решении задачи распознавания имеющиеся данные обычно делятся на обучающую, тестовую и контрольную выборки. Обучающая выборка используется для оценки параметров классификатора, тестовая и контрольная – для оценки обобщающих возможностей классификатора.

Настоящее пособие содержит описание использования теории активного восприятия (ТАВ), первоначально разработанной для решения задач распознавания изображений, в обработке одномерных сигналов. ТАВ предназначена для разработки реализаций этапов предварительной обработки и формирования признакового описания сигнала.

В жизненном цикле системы распознавания можно выделить два этапа: создание и эксплуатация.

На первом этапе формулируются цели системы распознавания, осуществляется постановка задачи распознавания для некоторой предметной области, выбираются способы реализации элементов системы распознавания, выполняется сбор данных для исследований.

На втором этапе осуществляется функционирование распознающей системы, заключающееся в классификации неизвестных элементов среды, зафиксированных на её входе, и поддержании системы в работоспособном состоянии. Как правило, скорость выполнения классификации весьма высока и соизмерима со скоростью функционирования распознаваемых элементов, что дает право говорить о функционировании таких систем распознавания в реальном масштабе времени. Это обуславливается тем, что основной объем работ научно-производственного характера при решении проблемы распознавания выполняется на первом этапе.

Информацию (сведения, данные) об исследуемых элементах среды, получаемую на первом этапе жизненного цикла распознающих систем (на этапе разработки), принято называть априорной, а получаемую на втором этапе (в процессе распознавания поступающих на её вход неизвестных элементов) – апостериорной.

Контрольные вопросы

1. Основные элементы системы распознавания образов.
2. Этап предварительной обработки данных.
3. Этап формирования признакового описания.
4. Этап принятия решения.
5. Жизненный цикл системы распознавания образов.
6. Этап создания системы распознавания образов.
7. Этап использования системы распознавания образов.

2. ФОРМИРОВАНИЕ СПЕКТРАЛЬНОГО ПРЕДСТАВЛЕНИЯ ЗВУКОВОГО СИГНАЛА С ПОЗИЦИЙ ТАВ

Непрерывным звуковым сигналом называется непрерывная действительная функция $s_a(t)$ действительного аргумента t , определённая на конечном интервале времени \bar{T} ($t \in \bar{T} \subset R^1$).

Дискретным звуковым сигналом s_{dsc} называется числовая последовательность, каждый элемент которой n есть значение непрерывного звукового сигнала $s_a(t)$ в момент времени $n\Delta t$:

$$s_{dsc}[n] = \begin{cases} s_a(n\Delta t), & \text{если } t \in \bar{T} \subset R^1, \\ 0, & \text{если } t \notin \bar{T}, \end{cases}$$

где Δt – шаг дискретизации; $n = \overline{0, (N-1)}$; N – число отсчётов дискретного звукового сигнала.

Цифровой звуковой сигнал s_{dig} формируется в результате квантования по уровню отсчётов дискретного звукового сигнала s_{dsc} .

Учитывая, что неизвестны статистические характеристики звукового сигнала и помех, его искажающих, обработка звукового сигнала выполняется в условиях априорной неопределённости.

Цифровой звуковой сигнал s_{dig} можно рассматривать как системное образование, представляющее собой совокупность элементов (отсчётов), между которыми существуют связи. Следующие признаки позволяют считать цифровой звуковой сигнал системным образованием:

1) целостность и членимость звукового сигнала: с одной стороны, звуковой сигнал представляет собой целостное образование, с другой стороны, в цифровом звуковом сигнале можно выделить элементы (отсчёты);

2) наличие связей (отношений) между элементами звукового сигнала, которые определяют его интегративные свойства: без наличия связей между элементами сигнала, ограничений на возможные комбинации элементов невозможна передача информации с использованием звуковых сигналов. Звуковой сигнал с некоторой долей упрощения можно сравнить с набором детских кубиков, на боковые поверхности каждого нанесены части изображения; каждая сторона кубика несёт какую-то часть общего изображения; чтобы получить общее изображение, надо соединить эти части, составив кубики; произвольно складывать кубики нельзя; их можно складывать только по правилу, устанавливая между кубиками связи; эти правила представляют собой ограничения, накладываемые на комбинацию элементов;

3) наличие интегративных свойств, присущих звуковому сигналу как системе в целом, но не присущих его элементам (отсчётам) в отдельности;

4) наличие в звуковом сигнале как системе организации, выражающейся в том числе в устойчивости звукового сигнала к искажающим воздействиям.

Таким образом, цифровой звуковой сигнал s_{dig} представляет собой системное образование, а не является лишь суммой свойств составляющих его отсчётов.

Цифровой звуковой сигнал s_{dig} конечен, следовательно, он, а также любой его подынтервал допускает применение операции сложения:

$$m(\bar{T}) = \sum_{n=0}^{N-1} s_{dig}[n], \quad (1)$$

где N – число отсчётов сигнала s_{dig} ; $m(\bar{T})$ – сумма отсчётов сигнала.

Результат операции (1), с позиций теории активного восприятия (ТАВ), можно интерпретировать как структурный элемент сигнала, выявляемый на самом грубом уровне разрешения. В ТАВ, в приложении к изображению, данная операция названа Q -преобразованием. Q -преобразование соответствует этапу предварительной обработки сигнала. Данное преобразование имеет следующую информационную интерпретацию: отличие получаемого значения $m(\bar{T})$ от нуля указывает на присутствие сигнала на интервале наблюдения \bar{T} . Для корректной интерпретации результата операции (1) цифровой сигнал s_{dig} необходимо привести к положительной области значений (как один из возможных вариантов, к отрезку $[0;1]$). Например, результат применения операции (1) к сигналу $s_{dig} = \{-1, -1, 1, 1\}$ равен нулю. В этом случае результат операции (1) покажет отсутствие сигнала на интервале наблюдения, а не его наличие.

Звуковой сигнал s_{dig} в условиях априорной неопределённости, рассматриваемый как система, допускает декомпозицию с использованием операции дихотомии ($\bar{T} = \bar{T}_1 \cup \bar{T}_2$). Тогда по результатам дихотомии, используя (1), можно вычислить два структурных элемента: $m(\bar{T}_1)$ и $m(\bar{T}_2)$, а для выявления отношения между структурными элементами, с позиций ТАВ, использовать операцию вычитания:

$$\mu = m(\bar{T}_1) - m(\bar{T}_2). \quad (2)$$

Рассмотрим возможные значения μ :

1) $\mu = 0$: сигнал симметричен относительно центра, следовательно, на данном уровне разрешения локализовать положение экстремума нельзя и необходимо перейти на следующий (более точный) уровень;

2) $\mu < 0$: экстремум расположен слева от центра интервала \bar{T} ;

3) $\mu > 0$: экстремум расположен справа от центра интервала \bar{T} .

Таким образом, дихотомия интервала наблюдения сигнала позволяет проверить гипотезу об однородности (симметрии) сигнала.

Использование операции (1) на грубых уровнях анализа сигнала позволяет учитывать только его низкочастотные компоненты, в этом случае шумы и помехи как бы исчезают.

Деление отрезка на четыре равных интервала (уточнение положения экстремума, рис. 2, в) и использование операции (1) ставит в соответствие каждому из интервалов сумму отсчётов сигнала s_{dig} : $m_{2,1}$, $m_{2,2}$, $m_{2,3}$, $m_{2,4}$. Шарик на рис. 1 обозначает экстремум в сигнале. Отношение между данными величинами вычисляется следующим образом:

$$\begin{aligned}\mu_1 &= (m_{2,2} + m_{2,3}) - (m_{2,1} + m_{2,4}), \\ \mu_2 &= (m_{2,1} + m_{2,3}) - (m_{2,2} + m_{2,4}).\end{aligned}\quad (3)$$

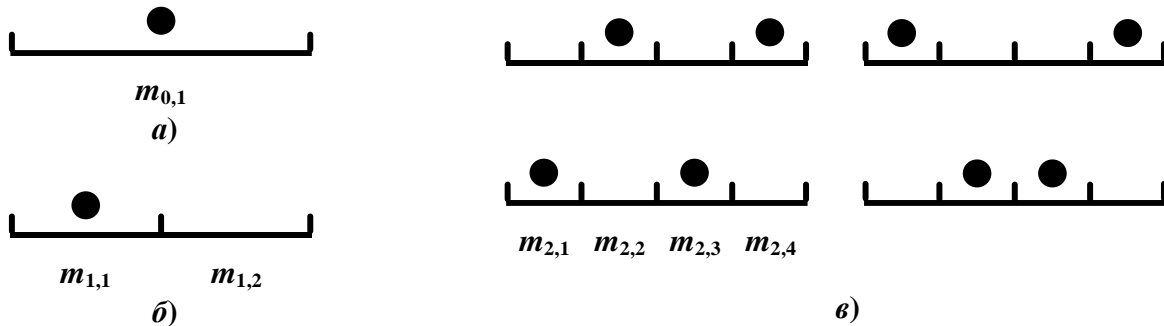


Рис. 2. Дихотомия отрезка: *a* – на ноль частей; *б* – пополам; *в* – на четыре части

Анализ значений μ_1 и μ_2 позволяет определить положение экстремума: в трёх случаях получаем, что в сигнале находится два экстремума (т. е. один экстремум на нулевом уровне, рис. 2, *a*, разделился на два экстремума на втором уровне, рис. 2, *в*), в одном случае (два шарика по центру) получаем, что для локализации экстремума требуется переместиться на уровень ниже (выполнить более тонкий анализ, рис. 3).

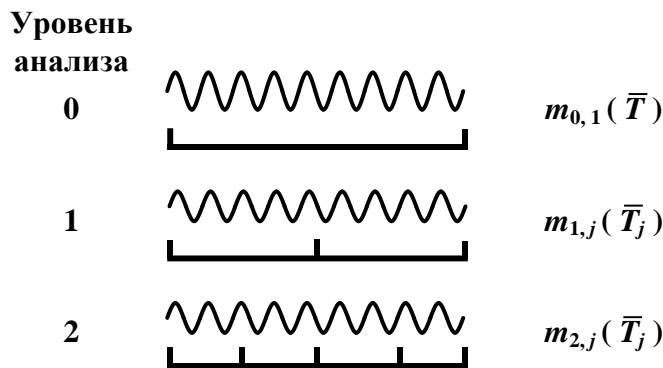


Рис. 3. Уровни анализа сигнала

Ранее были приведены операции анализа структуры сигнала на трёх уровнях разрешения. На первом уровне вычислялась и интерпретировалась сумма отсчётов сигнала в целом (один структурный элемент, см. формулу (1)), на втором – отношение между двумя структурными элементами (см. формулу (2)), на третьем – отношение между четырьмя элементами (см. формулу (3)).

Необходимо отметить, что на каждом следующем уровне анализа число вычисляемых отношений, по сравнению с суммарным количеством предыдущих найденных отношений, удваивается. Данные операции можно записать в следующем виде:

$$\mu_i = \sum_{j=0}^3 m_j F_i[j], i = \overline{0,3}, \quad (4)$$

где m_i – результат вычисления суммы отсчётов по i -му интервалу сигнала $i = \overline{1,4}$, $F_0 = (+1, +1, +1, +1)$; $F_1 = (+1, +1, -1, -1)$; $F_2 = (-1, +1, +1, -1)$; $F_3 = (+1, -1, +1, -1)$; $F_i[j]$ – j -й отсчёт i -го фильтра.

Формула (4) описывает U -преобразование – базовое преобразование ТАВ, выполняющее разложение сигнала по системе дискретных базисных функций (фильтров) F (рис. 4). В результате вычисления U -преобразования формируется спектральное разложение сигнала. U -преобразование соответствует этапу формирования признакового описания сигнала.

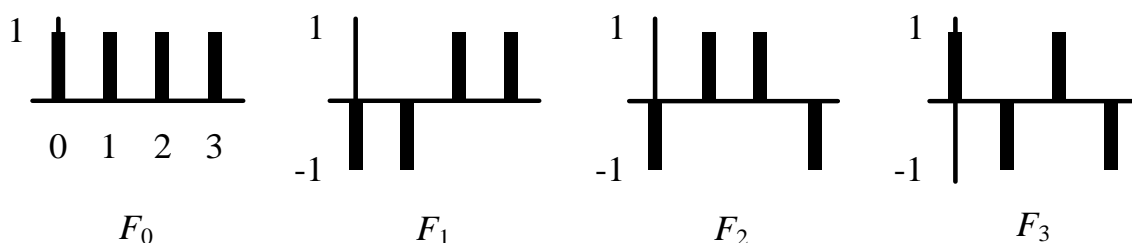


Рис. 4. Дискретные базисные функции

Результатом применения указанных фильтров к анализируемому сигналу является вектор μ спектральных коэффициентов. Приведённая система базисных функций обладает свойствами:

1) ортогональности – взаимная мощность любой пары функций равна нулю:

$$\sum_{i=0}^{N-1} F_k[i] \cdot F_l[i] = 0, k \neq l,$$

где N – число дискретных функций;

2) полноты – нельзя убрать и нельзя добавить новую функцию;

3) базисные функции являются двоичными – элементы базисных функций принимают значения «-1», «+1»;

4) модуль функции равен единице, а среднее – для всех, кроме F_0 , – равно нулю.

Приведённые фильтры внешне похожи на систему дискретных базисных функций Уолша системы Хармута. Специфика использования данных фильтров заключается в том, что они применяются после реализации Q -преобразования.

В общем виде U -преобразование записывается следующим образом:

$$\mu_i = \sum_{j=0}^{M-1} \left(\sum_{k \in T_i} s_{dig}[k] \right) F_i[j],$$

где T_i – i -й подынтервал сигнала s_{dig} ; M – число используемых фильтров.

Рассмотрим процесс вычисления двухуровневого U -представления тестового сигнала (сигнал включает 16 отсчётов).

Уровень 1. Q -преобразование. Сигнал разбивается на четыре части, амплитуды отсчётов, которые отнесены к каждой части, складываются (см. рис. 5, части отделены пунктиром). В результате генерируется вектор частичных сумм: $m = \{3; 3,25; 2; 2,5\}$.

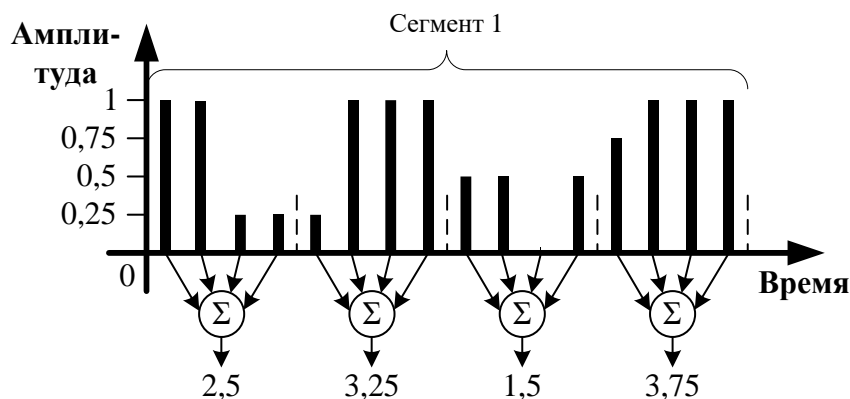


Рис. 5. Q -преобразование (первый уровень)

Уровень 1. Использование фильтров. Применяя к массиву сумм m фильтры, показанные на рис. 4, вычисляются значения спектральных коэффициентов. Для этого каждый фильтр поэлементно умножается на вектор частичных сумм. Например, для второго фильтра в результате умножения будет сгенерирован следующий массив: $\{-2,5; -3,25; 1,5; 3,75\}$. Далее выполняются сложение элементов полученного массива: $(-2,5) + (-3,25) + 1,5 + 3,75 = -0,5$. Результат сложения представляет собой спектральный коэффициент μ_i (табл. 1).

Таблица 1. Вычисление спектральных коэффициентов первого уровня

Результат умножения				Фильтр	Спектральные коэффициенты μ_i
2,5	3,25	1,5	3,75	F_0	11
-2,5	-3,25	1,5	3,75	F_1	-0,5
-2,5	3,25	1,5	-3,75	F_2	-1,5
2,5	-3,25	1,5	-3,75	F_3	-3

Уровень 2. Q -преобразование. Сигнал разбивается на четыре сегмента, каждый сегмент разбивается на четыре части, вычисляется сумма отсчётов, которые относятся к каждой части (рис. 6).

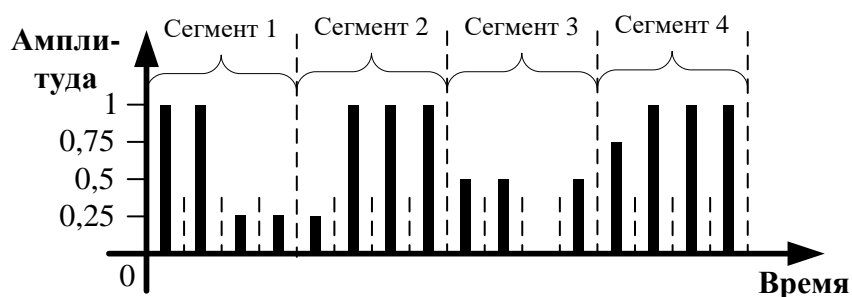


Рис. 6. Q -преобразование (второй уровень)

Уровень 2. Использование фильтров. С помощью фильтров, приведённых на рис. 4, и частичных сумм вычисляются значения спектральных коэффициентов (табл. 2).

Таблица 2. Вычисление спектральных коэффициентов второго уровня

Результат умножения								Фильтр	Спектральные коэффи- циенты μ_i		
Сегм. 1				Сегм. 2					Сегм. 1	Сегм. 2	
1	1	0,25	0,25	0,25	1	1	1	F_0	2,5	3,25	
-1	-1	0,25	0,25	-0,25	-1	1	1	F_1	-1,5	0,75	
-1	1	0,25	-0,25	-0,25	1	1	-1	F_2	0	0,75	
1	-1	0,25	-0,25	0,25	-1	1	-1	F_3	0	-0,75	
Сегм. 3				Сегм. 4				Фильтр	Сегм. 3	Сегм. 4	
0,5	0,5	0	0,5	0,75	1	1	1		F_0	1,5	3,75
-0,5	-0,5	0	0,5	-0,75	-1	1	1		F_1	0,5	0,25
-0,5	0,5	0	-0,5	-0,75	1	1	-1		F_2	-0,5	-0,5
0,5	-0,5	0	-0,5	0,75	-1	1	-1		F_3	-0,5	-0,25

Контрольные вопросы

1. Приведите определение непрерывного звукового сигнала.
2. Приведите определение дискретного звукового сигнала.
3. Приведите определение цифрового звукового сигнала.
4. Опишите свойства сигнала с позиций системного анализа.
5. Сформулируйте понятие U -преобразования сигнала.
6. Сформулируйте понятие Q -преобразования сигнала.
7. Опишите связь между U -преобразованием и фильтрами Уолша системы Хармута.
8. Опишите свойства фильтров, лежащих в основе U -преобразования.
9. Опишите понятие дихотомия относительно звукового сигнала.
10. Опишите свойства целостности и членимости звукового сигнала.
11. Опишите свойство интегративности звукового сигнала.
12. Опишите свойство наличия в звуковом сигнале организации.

3. АЛГЕБРА ГРУПП

Дальнейшие описания будут относиться к случаю, когда используется 16 фильтров. Приводимые далее алгоритмы и описания легко обобщаются на использование большего и меньшего числа фильтров.

В ТАВ каждому фильтру F_i ставится в соответствие бинарный оператор V_i . В этом случае компоненте $\mu_i \neq 0$ вектора μ соответствует оператор V_i или \bar{V}_i в зависимости от знака компоненты, а вектору μ соответствует подмножество операторов из V_i , имеющих аналогичную фильтрам конструкцию, но разное значение элементов матрицы: «+1» соответствует «1», «-1» соответствует «0».

Аналогичные процедуры можно реализовать для звукового сигнала. В этом случае фильтры переопределяются в одномерные булевы операторы по следующему алгоритму (рис. 7):

- 1) если $\mu_i > 0$, то имеем прямой оператор V_i ;
- 2) $\mu_i < 0$, то имеем инверсный оператор \bar{V}_i (рис. 8);
- 3) если $\mu_i = 0$, то оператор V_i отсутствует в описании сигнала, особенности сигнала по фильтру F_i не обнаружены (сигнал симметричен для данного фильтра).

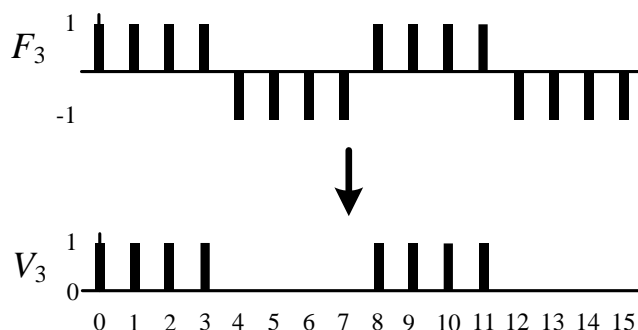


Рис. 7. Формирование образа оператора по фильтру

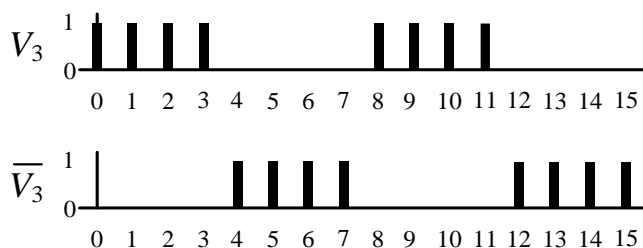


Рис. 8. Прямой и инверсный операторы

Задавая на множестве $\{V_i\}$ операции теоретико-множественного умножения и сложения, операцию, аналогичную отрицанию, два элемента:

$1 - V_0$ (e_0 , нижняя грань) и $0 - \overline{V_0}$ (e_1 , верхняя грань), получаем алгебру описания сигнала в одномерных булевых функциях: $A_V = \langle \{V_i\}: +, \times, \neg, e_1, e_0 \rangle$ (булеву алгебру или дистрибутивную решётку с дополнениями). Для любых V_i, V_j и $V_k \in V$ выполняются законы коммутативности, ассоциативности, идемпотентности и дистрибутивности.

На рис. 9 приведены 16 фильтров, а на рис. 10 – 16 операторов, соответствующих данным фильтрам.

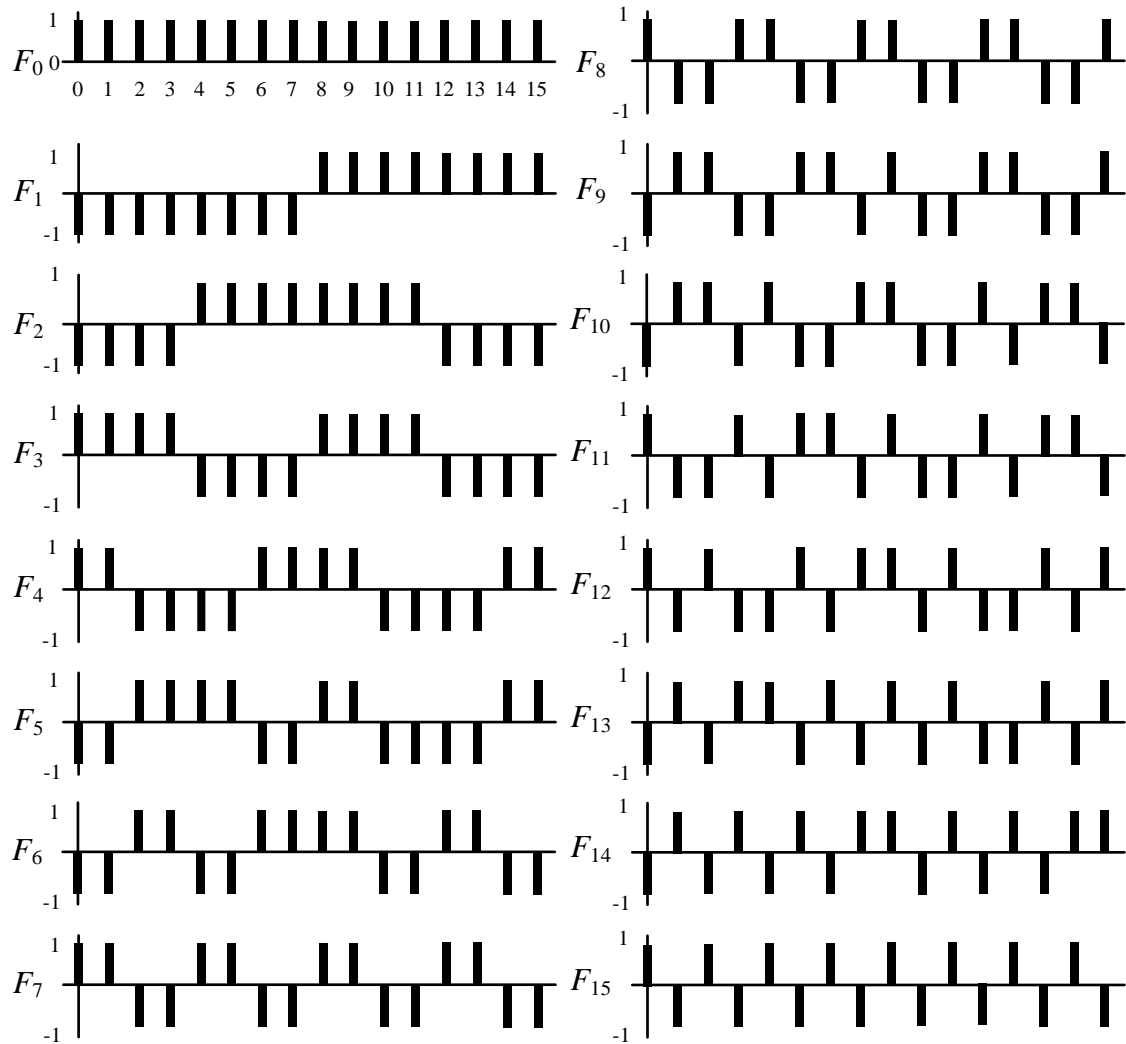


Рис. 9. Фильтры

На множестве одномерных операторов можно по аналогии с ТАВ сформировать алгебру групп, которая включает:

- 1) семейство структур P_n мощности 35 (для 16 операторов), названных полными группами;
- 2) семейство структур P_s мощности 105 (для 16 операторов), названных замкнутыми группами.

С помощью замкнутых и полных групп выполняется спектрально-корреляционный анализ. Полные группы позволяют выявить корреляционные связи между операторами. Замкнутые группы позволяют выявить корреляционные связи между полными группами.

Проводя аналогию между алгеброй групп и языком, можно ввести следующие соответствия: оператор – алфавит языка, полная группа – слово, замкнутая группа – словосочетание (рис. 11).

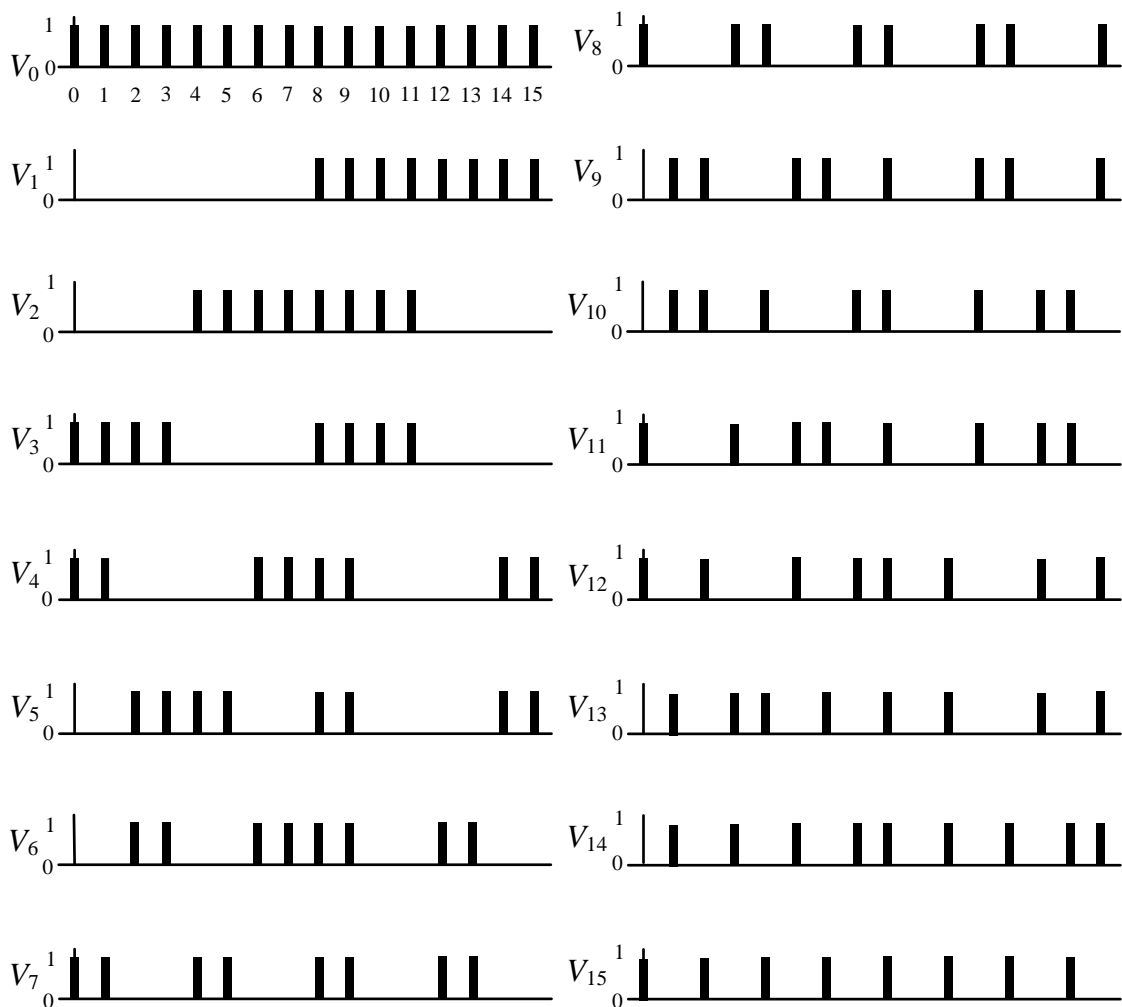


Рис. 10. Операторы

Каждой группе (полной или замкнутой), а также оператору можно поставить в соответствие характеристику – сумму элементов вектора частичных сумм, соответствующих единичным элементам образа группы (полной или замкнутой):

$$S = \sum_{j=0}^{M-1} I_j \cdot m_j, \quad (5)$$

где I – образ группы (оператора); m_j – j -й элемент вектора частичных сумм. Образ группы формируется в результате применения операций сложения и умножения к операторам, входящим в состав группы.

Рассмотрим пример вычисления суммы отсчётов, находящихся под образом оператора V_1 для сигнала, показанного на рис. 5, с помощью формулы (5). В этом случае m_j соответствует отсчёту сигнала. Образ оператора V_1 показан на рис. 10:

$$S = 0 \cdot 1 + 0 \cdot 1 + 0 \cdot 0,25 + 0 \cdot 0,25 + 0 \cdot 0,25 + 0 \cdot 1 + 0 \cdot 1 + 0 \cdot 1 + \\ + 1 \cdot 0,5 + 1 \cdot 0,5 + 1 \cdot 0 + 1 \cdot 0,5 + 1 \cdot 0,75 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 5,25.$$

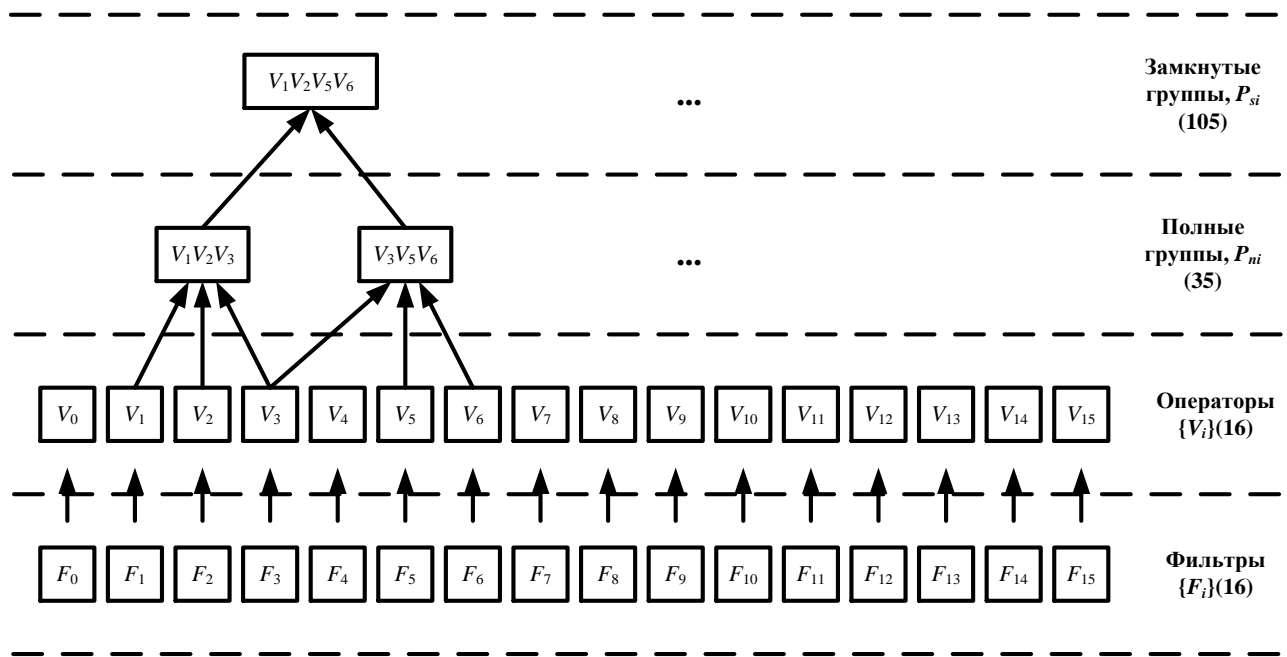


Рис. 11. Иерархия операторов и групп

В табл. 4. приведены значения спектральных коэффициентов и сумм отсчётов сигнала, находящихся под образом оператора, для сигнала, показанного на рис. 5.

Таблица 4. Спектральное разложение сигнала на 16 фильтрах

μ_0	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7
11	-0,5	-1,5	-3,0	3,0	-1,5	-0,5	1,0
μ_8	μ_9	μ_{10}	μ_{11}	μ_{12}	μ_{13}	μ_{14}	μ_{15}
-0,5	1,0	0,0	1,5	0,5	-1,0	0,0	-1,5
V_0	\bar{V}_1	\bar{V}_2	\bar{V}_3	V_4	\bar{V}_5	\bar{V}_6	V_7
11	5,25	4,75	4	7	4,75	5,25	6
\bar{V}_8	V_9	V_{10}	V_{11}	V_{12}	\bar{V}_{13}	V_{14}	\bar{V}_{15}
5,25	6	—	6,25	5,75	5	—	4,75

Операторы V_{10} и V_{14} отсутствуют в описании сигнала, поскольку соответствующие им спектральные коэффициенты равны нулю.

Контрольные вопросы

1. Опишите алгоритм формирования операторов на основе фильтров.
2. Опишите понятие “полная группа”.
3. Опишите понятие “замкнутая группа”.
4. Опишите алгоритм вычисления массы оператора (полной, замкнутой группы).
5. Вычислите спектральное разложение случайного сигнала, а также массы операторов, полных и замкнутых групп.

4. ПОЛНЫЕ ГРУППЫ

Свойства полных групп:

1) полная группа образуется на тройке операторов (V_i, V_j, V_k);

2) известны полные группы на операции умножения (число инверсий операторов, входящих в описание полной группы, должно быть чётно) и полные группы на операции сложения (число инверсий операторов, входящих в описание полной группы, должно быть нечётно); операция «сложение» или «умножение» в данном случае обозначает способ, которым формируется образ полной группы на основе операторов: $V_i \cdot V_j \cdot V_k$ или $V_i + V_j + \overline{V_k}$;

3) число возможных образов полных групп, с учётом инверсий операторов, входящих в их состав, равно четырём; образ полной группы на операции умножения включает 4 единичных элемента и 12 нулевых, на операции сложения – 12 единичных элементов и 4 нулевых (при использовании для формирования групп 16 операторов);

4) каждая полная группа образована на тройке спектральных компонент, поэтому она позволяет выявлять корреляционные (энергетические) связи между этими компонентами: если полная группа присутствует в описании сигнала, тогда связь между этими компонентами есть;

5) полная группа на операции сложения выделяет сегмент сигнала, на котором сумма его отсчётов минимальна, на операции умножения – сегмент сигнала, на котором сумма его отсчётов максимальна; точность определения зависит от того, насколько хорошо образ полной группы соответствует форме сигнала.

Алгоритм генерации полных групп заключается в следующем: перебираются все возможные сочетания троек операторов V_i, V_j, V_k ; три оператора составляют полную группу, если $V_i + V_j + V_k \equiv e_1$. Оператор V_0 в переборе не участвует. После выполнения данного алгоритма на 16 операторах будут сгенерированы следующие полные группы: $V_1V_2V_3, V_1V_4V_5, V_1V_6V_7, V_1V_8V_9, V_1V_{10}V_{11}, V_1V_{12}V_{13}, V_1V_{14}V_{15}, V_2V_4V_6, V_2V_5V_7, V_2V_8V_{10}, V_2V_9V_{11}, V_2V_{12}V_{14}, V_2V_{13}V_{15}, V_3V_4V_7, V_3V_5V_6, V_3V_8V_{11}, V_3V_9V_{10}, V_3V_{12}V_{15}, V_3V_{13}V_{14}, V_4V_8V_{12}, V_4V_9V_{13}, V_4V_{10}V_{14}, V_4V_{11}V_{15}, V_5V_8V_{13}, V_5V_9V_{12}, V_5V_{10}V_{15}, V_5V_{11}V_{14}, V_6V_8V_{14}, V_6V_9V_{15}, V_6V_{10}V_{12}, V_6V_{11}V_{13}, V_7V_8V_{15}, V_7V_9V_{14}, V_7V_{10}V_{13}, V_7V_{11}V_{12}$.

Данный список можно рассматривать как список «моделей» полных групп, поскольку после определения наличия каждой группы в сигнале у операторов, входящих в состав полных групп, появятся инверсии, а некоторые полные группы будут отсутствовать в описании сигнала, поскольку в их составе будут операторы, принимающие значение 0.

На рис. 12 показаны варианты записи полных групп в зависимости от знаков, входящих в их состав операторов.

V_i	V_j	V_k	Знак		V_i	V_j	V_k
0	0	0	+		$\bar{V}_i + \bar{V}_j + \bar{V}_k$		
1	0	0	×		$V_i \times \bar{V}_j \times \bar{V}_k$		
0	1	0	×		$\bar{V}_i \times V_j \times \bar{V}_k$		
1	1	0	+	→	$V_i + V_j + \bar{V}_k$		
0	0	1	×		$\bar{V}_i \times \bar{V}_j \times V_k$		
1	0	1	+		$V_i + \bar{V}_j + V_k$		
0	1	1	+		$\bar{V}_i + V_j + V_k$		
1	1	1	×		$V_i \times V_j \times V_k$		

0 – инверсный
 1 – прямой
 + – нечётное кол-во инв.
 × – чётное кол-во инв.

Варианты образов полных групп

Рис. 12. Варианты записи полных групп

Рассмотрим алгоритм вычисления полных групп, описывающих сигнал, показанный на рис. 5:

- 1) вычисляется спектр сигнала s_{dig} ;
- 2) каждому элементу спектра ставится в соответствие оператор из множества $\{V_i\}$ (элементы спектра нумеруются с нуля, номера операторов и элементов спектра совпадают).
- 3) проверяется наличие i -й полной группы в сигнале s_{dig} :
 - а) из спектра выбираются три элемента, номера которых соответствуют номерам операторов, входящих в i -ю группу;
 - б) определяется количество инверсных операторов: если значение j -го элемента спектра меньше нуля, то считается, что соответствующий элементу оператор – инверсный; рассматривается две ситуации:
 - число инверсных операторов чётно: вычисляется образ полной группы на операции умножения. Пусть из спектра выбраны три элемента, соответствующие операторам V_1, V_4, V_5 . На рис. 13 показан процесс вычисления образа полной группы $\bar{V}_1 V_4 \bar{V}_5$; образ данной группы выделяет область сигнала с максимальной суммой отсчётов;
 - число инверсных операторов нечётно: вычисляется образ полной группы на операции сложения. На рис. 14 показан процесс вычисления

образа полной группы $\overline{V_1 V_2 V_3}$; образ данной полной группы выделяет области сигнала с минимальной суммой отсчётов;

4) вычисляется сумма отсчётов сигнала, находящихся под образом группы.

Шаги 3-4 выполняются для всех 35 полных групп. В результате, формируется два массива сумм отсчётов, находящихся под образами полных групп: для полных групп на операции сложения и операции умножения.

В состав полной группы входит три оператора. Для каждой «модели» полной группы можно, с учётом инверсий операторов (рис. 12), поставить в соответствие четыре полных группы на операции сложения и четыре полных группы на операции умножения.

Таким образом, всего, с учётом инверсий, существует $4 \cdot 35 = 140$ полных групп на операции сложения и столько же полных групп на операции умножения. В связи с этим каждой полной группе можно поставить в соответствие число от 1 до 140.

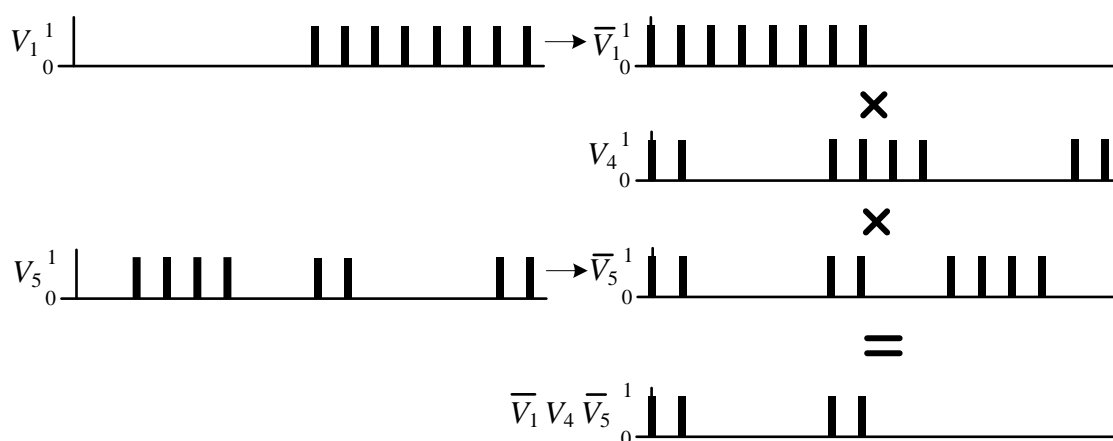


Рис. 13. Вычисление образа полной группы на операции умножения

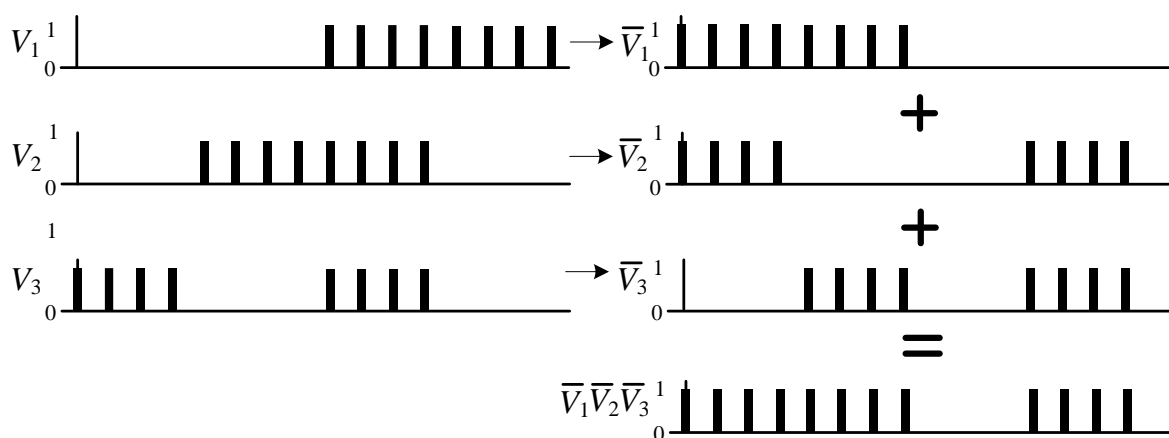


Рис. 14. Вычисление образа полной группы на операции сложения

Несмотря на то, что всего, с учётом инверсий, существует 140 полных групп на операции сложения и 140 на операции умножения, по сигналу можно вычислить не более 35 групп (на операции сложения и умножения). В табл. 5 показаны полные группы, для сигнала, показанного на рис. 5.

Таблица 5. Полные группы, входящие в описание сигнала

Группа	$\overline{V_1} \overline{V_2} \overline{V_3}$	$\overline{V_1} \overline{V_4} \overline{V_5}$	$\overline{V_1} \overline{V_6} \overline{V_7}$	$\overline{V_1} \overline{V_8} \overline{V_9}$	$\overline{V_1} \overline{V_{10}} \overline{V_{11}}$	$\overline{V_1} \overline{V_{12}} \overline{V_{13}}$	$\overline{V_1} \overline{V_{14}} \overline{V_{15}}$
Пояснение	+	×	×	×	0	×	0
Сумма отсчётов	9,5	4	3,25	3,25	-	3,25	-
Группа	$\overline{V_2} \overline{V_4} \overline{V_6}$	$\overline{V_2} \overline{V_5} \overline{V_7}$	$\overline{V_2} \overline{V_8} \overline{V_{10}}$	$\overline{V_2} \overline{V_9} \overline{V_{11}}$	$\overline{V_2} \overline{V_{12}} \overline{V_{14}}$	$\overline{V_2} \overline{V_{13}} \overline{V_{15}}$	$\overline{V_3} \overline{V_4} \overline{V_7}$
Пояснение	×	×	0	+	0	+	+
Сумма Отсчётов	4	3,75	-	9,25	-	9,25	10
Группа	$\overline{V_3} \overline{V_5} \overline{V_6}$	$\overline{V_3} \overline{V_8} \overline{V_{11}}$	$\overline{V_3} \overline{V_9} \overline{V_{10}}$	$\overline{V_3} \overline{V_{12}} \overline{V_{15}}$	$\overline{V_3} \overline{V_{13}} \overline{V_{14}}$	$V_4 \overline{V_8} \overline{V_{12}}$	$V_4 \overline{V_9} \overline{V_{13}}$
Пояснение	+	×	0	×	0	+	+
Сумма Отсчётов	9,5	4	-	4	-	9,25	9,5
Группа	$V_4 \overline{V_{10}} \overline{V_{14}}$	$V_4 \overline{V_{11}} \overline{V_{15}}$	$\overline{V_5} \overline{V_8} \overline{V_{13}}$	$\overline{V_5} \overline{V_9} \overline{V_{12}}$	$\overline{V_5} \overline{V_{10}} \overline{V_{15}}$	$\overline{V_5} \overline{V_{11}} \overline{V_{14}}$	$\overline{V_6} \overline{V_8} \overline{V_{14}}$
Пояснение	0	+	+	+	0	0	0
Сумма Отсчётов	-	9,75	9	9	-	-	-
Группа	$V_6 \overline{V_9} \overline{V_{15}}$	$\overline{V_6} \overline{V_{10}} \overline{V_{12}}$	$\overline{V_6} \overline{V_{11}} \overline{V_{13}}$	$V_7 \overline{V_8} \overline{V_{15}}$	$V_7 \overline{V_9} \overline{V_{14}}$	$V_7 \overline{V_{10}} \overline{V_{13}}$	$V_7 \overline{V_{11}} \overline{V_{12}}$
Пояснение	×	0	×	×	0	0	×
Сумма Отсчётов	3,5	-	3,5	3,5	-	-	3,5

В табл. 5 используются следующие обозначения: «0» – соответствующая группа не входит в описание сигнала, так как содержит нулевой оператор, «+» – полная группа на операции сложения, «×» – полная группа на операции умножения.

Контрольные вопросы

1. Опишите свойства полных групп.
2. Опишите алгоритм генерации полных групп.
3. Опишите алгоритм вычисления состава полных групп по сигналу.

5. ЗАМКНУТЫЕ ГРУППЫ

Свойства замкнутых групп:

1) замкнутые группы образованы на четвёрке операторов (V_i, V_j, V_p, V_m) с описанием $V_i V_j + V_p \overline{V_m}$ (число инверсий операторов нечетно);

2) замкнутые группы, образованные на четвёрке операторов с чётным числом инверсий, называются замкнутыми множествами; допускается два варианта записи замкнутого множества: на операции сложения ($V_i + V_j + V_p + V_m$, образ составляет 14 элементов) и операции умножения ($V_i V_j V_p V_m$, образ составляет 2 элемента);

3) число возможных образов замкнутых групп, с учётом инверсий операторов, равно восьми; образ замкнутой группы составляет 8 элементов;

4) каждая замкнутая группа образована на паре полных групп путем их склейки на общем операторе, что позволяет выявлять корреляционные (энергетические) связи между данными анализа этих полных групп. Результатом является иерархия анализа спектральных компонент (откликов фильтров);

5) для замкнутой группы $V_i V_j V_k V_l$ существует восемь вариантов записи с учётом инверсий: $\overline{V_i} V_j + V_k V_l$, $V_i \overline{V_j} + V_k V_l$, $V_i V_j + \overline{V_k} V_l$, $V_i V_j + V_k \overline{V_l}$, $V_i \overline{V_j} + \overline{V_k} \overline{V_l}$, $\overline{V_i} V_j + \overline{V_k} \overline{V_l}$, $\overline{V_i} \overline{V_j} + V_k \overline{V_l}$, $\overline{V_i} \overline{V_j} + \overline{V_k} V_l$;

6) если один из операторов замкнутой группы – инверсный, то сумма операторов представляет собой единичный вектор;

7) образы замкнутых групп, получаемые при разном порядке операторов, одинаковые;

8) замкнутая группа выделяет сегмент сигнала, на котором сумма отсчётов сигнала максимальна;

9) в ТАВ определены два варианта записи замкнутого множества: на операции сложения и операции умножения;

10) для замкнутого множества $V_i V_j V_k V_l$ существует восемь вариантов записи с учётом инверсий: $V_i V_j V_k V_l$, $\overline{V_i} \overline{V_j} V_k V_l$, $\overline{V_i} V_j \overline{V_k} V_l$, $V_i \overline{V_j} \overline{V_k} V_l$, $\overline{V_i} V_j V_k \overline{V_l}$, $V_i \overline{V_j} V_k \overline{V_l}$, $\overline{V_i} \overline{V_j} \overline{V_k} \overline{V_l}$ (знак операции: сложение или умножение, в этой записи знак не указан).

Алгоритм генерации замкнутых групп заключается в следующем: перебираются все возможные сочетания троек операторов V_i, V_j, V_k, V_l , четыре оператора составляют замкнутую группу, если:

1) любые три оператора из четырёх, составляющих замкнутую группу, не образуют полную группу (операторы рассматриваются без инверсий);

2) сумма четырёх прямых операторов не составляет единицу (e_1 , единичный вектор);

3) пусть N – число операторов, тогда образ замкнутой группы (число инверсий операторов нечётно, т.е. образ строится с нечётным числом операторов), включает $((1/2) \cdot N)$ нулей и $((1/2) \cdot N)$ единиц.

После выполнения данного алгоритма на 16 операторах будут сгенерированы следующие замкнутые группы: $V_8V_9V_{10}V_{11}$, $V_4V_5V_6V_7$, $V_{12}V_{13}V_{14}V_{15}$, $V_8V_9V_{14}V_{15}$, $V_2V_3V_4V_5$, $V_{10}V_{11}V_{12}V_{13}$, $V_8V_9V_{12}V_{13}$, $V_2V_3V_6V_7$, $V_{10}V_{11}V_{14}V_{15}$, $V_8V_{11}V_{12}V_{15}$, $V_1V_2V_5V_6$, $V_9V_{10}V_{13}V_{14}$, $V_8V_{11}V_{13}V_{14}$, $V_1V_2V_4V_7$, $V_9V_{10}V_{12}V_{15}$, $V_8V_{10}V_{13}V_{15}$, $V_1V_3V_4V_6$, $V_9V_{11}V_{12}V_{14}$, $V_8V_{10}V_{12}V_{14}$, $V_1V_3V_5V_7$, $V_9V_{11}V_{13}V_{15}$, $V_6V_7V_8V_9$, $V_2V_3V_{12}V_{13}$, $V_4V_5V_{10}V_{11}$, $V_4V_5V_8V_9$, $V_2V_3V_{14}V_{15}$, $V_6V_7V_{10}V_{11}$, $V_4V_7V_8V_{11}$, $V_1V_2V_{13}V_{14}$, $V_5V_6V_9V_{10}$, $V_5V_6V_8V_{11}$, $V_1V_2V_{12}V_{15}$, $V_4V_7V_9V_{10}$, $V_5V_7V_8V_{10}$, $V_1V_3V_{12}V_{14}$, $V_4V_6V_9V_{11}$, $V_4V_6V_8V_{10}$, $V_1V_3V_{13}V_{15}$, $V_5V_7V_9V_{11}$, $V_4V_5V_{12}V_{13}$, $V_6V_7V_{14}V_{15}$, $V_2V_3V_{10}V_{11}$, $V_2V_3V_8V_9$, $V_6V_7V_{12}V_{13}$, $V_4V_5V_{14}V_{15}$, $V_5V_6V_{13}V_{14}$, $V_4V_7V_{12}V_{15}$, $V_1V_2V_9V_{10}$, $V_1V_2V_8V_{11}$, $V_4V_7V_{13}V_{14}$, $V_5V_6V_{12}V_{15}$, $V_4V_6V_{12}V_{14}$, $V_5V_7V_{13}V_{15}$, $V_1V_3V_9V_{11}$, $V_1V_3V_8V_{10}$, $V_5V_7V_{12}V_{14}$, $V_4V_6V_{13}V_{15}$, $V_2V_5V_{10}V_{13}$, $V_3V_4V_{11}V_{12}$, $V_1V_6V_9V_{14}$, $V_1V_6V_8V_{15}$, $V_3V_4V_{10}V_{13}$, $V_2V_5V_{11}V_{12}$, $V_1V_7V_8V_{14}$, $V_3V_5V_{10}V_{12}$, $V_2V_4V_{11}V_{13}$, $V_2V_4V_{10}V_{12}$, $V_3V_5V_{11}V_{13}$, $V_1V_7V_9V_{15}$, $V_3V_4V_8V_{15}$, $V_1V_6V_{10}V_{13}$, $V_2V_5V_9V_{14}$, $V_2V_5V_8V_{15}$, $V_1V_6V_{11}V_{12}$, $V_3V_4V_9V_{14}$, $V_2V_4V_8V_{14}$, $V_1V_7V_{11}V_{13}$, $V_3V_5V_9V_{15}$, $V_3V_5V_8V_{14}$, $V_1V_7V_{10}V_{12}$, $V_2V_4V_9V_{15}$, $V_3V_7V_8V_{12}$, $V_1V_5V_{10}V_{14}$, $V_2V_6V_9V_{13}$, $V_2V_6V_8V_{12}$, $V_1V_5V_{11}V_{15}$, $V_3V_7V_9V_{13}$, $V_2V_7V_8V_{13}$, $V_1V_4V_{11}V_{14}$, $V_3V_6V_9V_{12}$, $V_3V_6V_8V_{13}$, $V_1V_4V_{10}V_{15}$, $V_2V_7V_9V_{12}$, $V_2V_6V_{10}V_{14}$, $V_3V_7V_{11}V_{15}$, $V_1V_5V_9V_{13}$, $V_1V_5V_8V_{12}$, $V_3V_7V_{10}V_{14}$, $V_2V_6V_{11}V_{15}$, $V_1V_4V_8V_{13}$, $V_3V_6V_{10}V_{15}$, $V_2V_7V_{11}V_{14}$, $V_2V_7V_{10}V_{15}$, $V_3V_6V_{11}V_{14}$, $V_1V_4V_9V_{12}$.

Оператор V_0 в переборе не участвует. На рис. 15 показаны варианты записи замкнутых групп («Г») или множеств («М») в зависимости от знаков, входящих в их состав операторов.

V_i	V_j	V_k	V_l	Γ/\mathbf{M}	V_i	V_j	V_k	V_l	V_i	V_j	V_k	V_l	Γ/\mathbf{M}	V_i	V_j	V_k	V_l
0	0	0	0	M	$\bar{V}_i \oplus \bar{V}_j \oplus \bar{V}_k \oplus \bar{V}_l$	0	0	0	1	Γ	$\bar{V}_i \times \bar{V}_j + \bar{V}_k \times \bar{V}_l$	0	Γ	$\bar{V}_i \times \bar{V}_j + \bar{V}_k \times \bar{V}_l$			
1	0	0	0	Γ	$V_i \times \bar{V}_j + \bar{V}_k \times \bar{V}_l$	1	0	0	1	M	$V_i \oplus \bar{V}_j \oplus \bar{V}_k \oplus \bar{V}_l$	1	M	$V_i \oplus \bar{V}_j \oplus \bar{V}_k \oplus \bar{V}_l$			
0	1	0	0	Γ	$\bar{V}_i \times V_j + \bar{V}_k \times \bar{V}_l$	0	1	0	1	M	$\bar{V}_i \oplus V_j \oplus \bar{V}_k \oplus \bar{V}_l$	0	M	$\bar{V}_i \oplus V_j \oplus \bar{V}_k \oplus \bar{V}_l$			
1	1	0	0	M	$V_i \oplus V_j \oplus \bar{V}_k \oplus \bar{V}_l$	1	1	0	1	Γ	$V_i \times V_j + \bar{V}_k \times \bar{V}_l$	1	Γ	$V_i \times V_j + \bar{V}_k \times \bar{V}_l$			
0	0	1	0	Γ	$\bar{V}_i \times \bar{V}_j + V_k \times \bar{V}_l$	0	0	1	1	M	$\bar{V}_i \oplus \bar{V}_j \oplus V_k \oplus \bar{V}_l$	0	M	$\bar{V}_i \oplus \bar{V}_j \oplus V_k \oplus \bar{V}_l$			
1	0	1	0	M	$V_i \oplus \bar{V}_j \oplus V_k \oplus \bar{V}_l$	1	0	1	1	Γ	$V_i \times \bar{V}_j + V_k \times \bar{V}_l$	1	Γ	$V_i \times \bar{V}_j + V_k \times \bar{V}_l$			
0	1	1	0	M	$\bar{V}_i \oplus V_j \oplus V_k \oplus \bar{V}_l$	0	1	1	1	Γ	$\bar{V}_i \times V_j + V_k \times \bar{V}_l$	0	Γ	$\bar{V}_i \times V_j + V_k \times \bar{V}_l$			
1	1	1	0	Γ	$V_i \times V_j + V_k \times \bar{V}_l$	1	1	1	1	M	$V_i \oplus V_j \oplus V_k \oplus \bar{V}_l$	1	M	$V_i \oplus V_j \oplus V_k \oplus \bar{V}_l$			

Рис. 15. Варианты записи замкнутых групп и множеств

Рассмотрим алгоритм вычисления замкнутых групп, описывающих сигнал:

1) вычислить для сигнала s_{dig} спектр на основе U -преобразования;
 2) каждому элементу спектра ставится в соответствие оператор из множества $\{V_i\}$ (элементы спектра нумеруются с нуля, номера операторов и элементов спектра совпадают);

3) выполняется проверка наличие i -й замкнутой группы в сигнале s_{dig} :

а) из спектра выбираются четыре элемента, номера которых соответствуют номерам операторов, входящих в i -ю группу;

б) определяется количество инверсных операторов: если значение j -го элемента спектра меньше нуля, то считается, что соответствующий элементу оператор – инверсный; рассматривается две ситуации:

- число инверсных операторов чётно: вычисляется образ замкнутого множества. Пусть из спектра выбраны четыре элемента, соответствующие операторам V_8, V_9, V_{10}, V_{11} . На рис. 16 показан пример вычисления образа замкнутого множества на операции сложения: $V_8 + V_9 + V_{10} + V_{11}$, и на операции умножения: $V_8 V_9 V_{10} V_{11}$;

- число инверсных операторов нечётно: вычисляется образ замкнутой группы. На рис. 16 показан пример вычисления образа замкнутой группы $\bar{V}_8 V_9 + V_{10} V_{11}$;

4) вычисляется сумма отсчётов сигнала, находящихся под образом группы.

Каждой замкнутой группе, присутствующей в сигнале, в зависимости от числа инверсий можно поставить в соответствие число от 1 до 840. Данная операция выполняется отдельно для замкнутых групп и множеств.

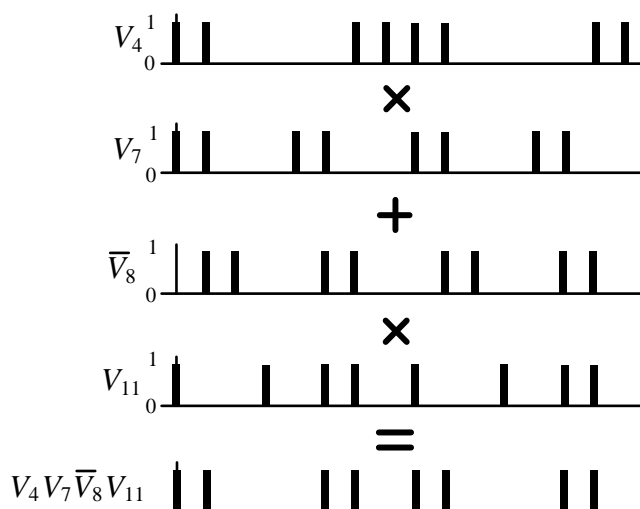


Рис. 16. Вычисление образа замкнутой группы

Несмотря на то, что всего, с учётом инверсий, существует 840 замкнутых групп и $840 + 840$ замкнутых множеств, по сигналу можно вычислить не более 140 замкнутых групп и 280 замкнутых множеств (140 – на операции сложения и 140 – на операции умножения). В табл. 6 показаны полные группы, описывающие сигнал, показанный на рис. 5.

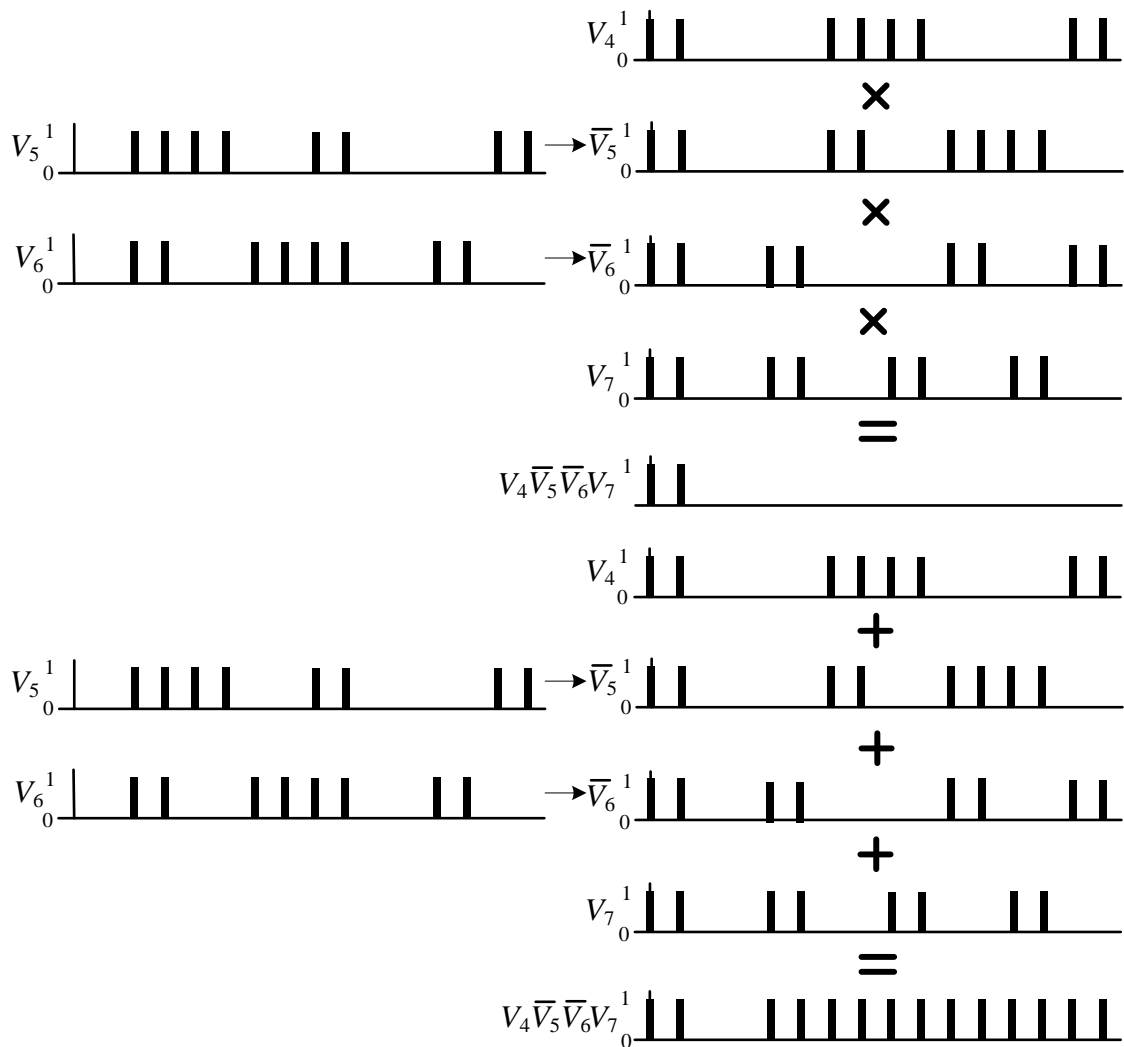


Рис. 17. Вычисление образа замкнутого множества

В табл. 6 используются следующие обозначения: «0» – соответствующая группа не входит в описание сигнала, так как содержит нулевой оператор, «Гр.» – замкнутая группа на операции сложения, «Мн.» – замкнутое множество. Для замкнутых множеств в поле «Сумма отсчётов» сначала указывается сумма при записи множества на операции умножения, а в скобках – на операции сложения.

Таблица 6. Замкнутые группы, входящие в описание сигнала

Группа	$\overline{V_8} \overline{V_9} V_{10} V_{11}$	$V_4 \overline{V_5} \overline{V_6} V_7$	$V_{12} \overline{V_{13}} V_{14} \overline{V_{15}}$	$\overline{V_8} \overline{V_9} V_{14} \overline{V_{15}}$	$\overline{V_2} \overline{V_3} V_4 \overline{V_5}$
Пояснение	0	Мн.	0	0	Гр.
Сумма отсчётов	-	2 (10,5)	-	-	7,75
Группа	$V_{10} V_{11} V_{12} \overline{V_{13}}$	$\overline{V_8} \overline{V_9} V_{12} \overline{V_{13}}$	$\overline{V_2} \overline{V_3} \overline{V_6} V_7$	$V_{10} V_{11} V_{14} \overline{V_{15}}$	$\overline{V_8} V_{11} V_{12} \overline{V_{15}}$
Пояснение	0	Мн.	Гр.	0	Мн.
Сумма Отсчётов	-	1,25 (10,5)	7	-	2 (10)
Группа	$\overline{V_1} \overline{V_2} \overline{V_5} \overline{V_6}$	$V_9 V_{10} \overline{V_{13}} V_{14}$	$\overline{V_8} V_{11} \overline{V_{13}} V_{14}$	$\overline{V_1} \overline{V_2} V_4 V_7$	$V_9 V_{10} V_{12} \overline{V_{15}}$
Пояснение	Мн.	0	0	Мн.	0
Сумма Отсчётов	2 (10)	-	-	2 (10,5)	-
Группа	$\overline{V_8} V_{10} \overline{V_{13}} \overline{V_{15}}$	$\overline{V_1} \overline{V_3} V_4 \overline{V_6}$	$V_9 V_{11} V_{12} V_{14}$	$\overline{V_8} V_{10} V_{12} V_{14}$	$\overline{V_1} \overline{V_3} \overline{V_5} V_7$
Пояснение	0	Гр.	0	0	Гр.
Сумма Отсчётов	-	7,25	-	-	7
Группа	$V_9 V_{11} \overline{V_{13}} \overline{V_{15}}$	$\overline{V_6} \overline{V_7} \overline{V_8} V_9$	$\overline{V_2} \overline{V_3} V_{12} \overline{V_{13}}$	$V_4 \overline{V_5} V_{10} V_{11}$	$V_4 \overline{V_5} \overline{V_8} V_9$
Пояснение	Мн.	Мн.	Гр.	0	Мн.
Сумма Отсчётов	1,5 (10,75)	2 (9,75)	7	-	2 (10,5)
Группа	$\overline{V_2} \overline{V_3} V_{14} \overline{V_{15}}$	$\overline{V_6} \overline{V_7} V_{10} V_{11}$	$V_4 V_7 \overline{V_8} V_{11}$	$\overline{V_1} \overline{V_2} \overline{V_{13}} V_{14}$	$\overline{V_5} \overline{V_6} V_9 V_{10}$
Пояснение	0	0	Гр.	0	0
Сумма Отсчётов	-	-	7	-	-
Группа	$\overline{V_5} \overline{V_6} \overline{V_8} V_{11}$	$\overline{V_1} \overline{V_2} V_{12} \overline{V_{15}}$	$V_4 V_7 V_9 V_{10}$	$\overline{V_5} \overline{V_7} \overline{V_8} V_{10}$	$\overline{V_1} \overline{V_3} V_{12} V_{14}$
Пояснение	Гр.	Гр.	0	0	0
Сумма Отсчётов	6,5	6,5	-	-	-
Группа	$V_4 \overline{V_6} \overline{V_9} V_{11}$	$V_4 \overline{V_6} \overline{V_8} V_{10}$	$\overline{V_1} \overline{V_3} \overline{V_{13}} \overline{V_{15}}$	$\overline{V_5} \overline{V_7} V_9 V_{11}$	$V_4 \overline{V_5} V_{12} \overline{V_{13}}$
Пояснение	Гр.	0	Мн.	Гр.	Мн.
Сумма Отсчётов	7	-	2 (10,5)	6,75	2 (10,5)
Группа	$\overline{V_6} \overline{V_7} V_{14} \overline{V_{15}}$	$\overline{V_2} \overline{V_3} V_{10} V_{11}$	$\overline{V_2} \overline{V_3} \overline{V_8} V_9$	$\overline{V_6} \overline{V_7} V_{12} \overline{V_{13}}$	$V_4 \overline{V_5} V_{14} \overline{V_{15}}$
Пояснение	0	0	Гр.	Мн.	0
Сумма Отсчётов	-	-	7	2 (9,75)	-
Группа	$\overline{V_5} \overline{V_6} \overline{V_{13}} V_{14}$	$V_4 V_7 V_{12} \overline{V_{15}}$	$\overline{V_1} \overline{V_2} V_9 V_{10}$	$\overline{V_1} \overline{V_2} \overline{V_8} V_{11}$	$V_4 V_7 \overline{V_{13}} V_{14}$
Пояснение	0	Гр.	0	Гр.	0
Сумма Отсчётов	-	7	-	6,5	-

Продолжение табл. 6

Группа	$\overline{V_5} \overline{V_6} V_{12} \overline{V_{15}}$	$V_4 \overline{V_6} V_{12} V_{14}$	$\overline{V_5} V_7 \overline{V_{13}} \overline{V_{15}}$	$\overline{V_1} \overline{V_3} V_9 V_{11}$	$\overline{V_1} \overline{V_3} \overline{V_8} V_{10}$
Пояснение	Гр.	0	Гр.	Мн.	0
Сумма Отсчётов	6,5	-	6,75	2 (10,5)	-
Группа	$\overline{V_5} V_7 V_{12} V_{14}$	$V_4 \overline{V_6} \overline{V_{13}} \overline{V_{15}}$	$\overline{V_2} \overline{V_5} V_{10} \overline{V_{13}}$	$\overline{V_3} V_4 V_{11} V_{12}$	$\overline{V_1} \overline{V_6} V_9 V_{14}$
Пояснение	Гр.	Гр.	0	Гр.	0
Сумма Отсчётов	0	7	-	7,5	-
Группа	$\overline{V_1} \overline{V_6} \overline{V_8} \overline{V_{15}}$	$\overline{V_3} \overline{V_4} V_{10} \overline{V_{13}}$	$\overline{V_2} \overline{V_5} V_{11} V_{12}$	$\overline{V_1} V_7 \overline{V_8} V_{14}$	$\overline{V_3} \overline{V_5} V_{10} V_{12}$
Пояснение	Мн.	0	Мн.	0	0
Сумма Отсчётов	2 (9,75)	-	2 (10,25)	-	-
Группа	$\overline{V_2} V_4 V_{11} \overline{V_{13}}$	$\overline{V_2} V_4 V_{10} V_{12}$	$\overline{V_3} \overline{V_5} V_{11} \overline{V_{13}}$	$\overline{V_1} V_7 V_9 \overline{V_{15}}$	$\overline{V_3} V_4 \overline{V_8} \overline{V_{15}}$
Пояснение	Мн.	0	Гр.	Мн.	Гр.
Сумма Отсчётов	2 (10,75)	-	7,25	2 (10)	7,5
Группа	$\overline{V_1} \overline{V_6} V_{10} \overline{V_{13}}$	$\overline{V_2} \overline{V_5} V_9 V_{14}$	$\overline{V_2} \overline{V_5} \overline{V_8} \overline{V_{15}}$	$\overline{V_1} \overline{V_6} V_{11} V_{12}$	$\overline{V_3} V_4 V_9 V_{14}$
Пояснение	0	0	Мн.	Мн.	0
Сумма Отсчётов	-	-	2 (10,25)	2 (9,75)	-
Группа	$\overline{V_2} V_4 \overline{V_8} V_{14}$	$\overline{V_1} V_7 V_{11} \overline{V_{13}}$	$\overline{V_3} \overline{V_5} V_9 \overline{V_{15}}$	$\overline{V_3} \overline{V_5} \overline{V_8} V_{14}$	$\overline{V_1} V_7 V_{10} V_{12}$
Пояснение	0	Мн.	Гр.	0	Гр.
Сумма Отсчётов	-	2 (10)	7,25	-	0
Группа	$\overline{V_2} V_4 V_9 \overline{V_{15}}$	$\overline{V_3} V_7 \overline{V_8} V_{12}$	$\overline{V_1} \overline{V_5} V_{10} V_{14}$	$\overline{V_2} \overline{V_6} V_9 \overline{V_{13}}$	$\overline{V_2} \overline{V_6} \overline{V_8} V_{12}$
Пояснение	Мн.	Мн.	0	Гр.	Гр.
Сумма Отсчётов	2 (10,75)	2 (10,25)	-	6,5	6,25
Группа	$\overline{V_1} \overline{V_5} V_{11} \overline{V_{15}}$	$\overline{V_3} V_7 V_9 \overline{V_{13}}$	$\overline{V_2} V_7 \overline{V_8} \overline{V_{13}}$	$\overline{V_1} V_4 V_{11} V_{14}$	$\overline{V_3} \overline{V_6} V_9 V_{12}$
Пояснение	Гр.	Мн.	Гр.	0	Мн.
Сумма Отсчётов	6,75	1,75 (10,75)	6,5	-	2 (10,25)
Группа	$\overline{V_3} \overline{V_6} \overline{V_8} \overline{V_{13}}$	$\overline{V_1} V_4 V_{10} \overline{V_{15}}$	$\overline{V_2} V_7 V_9 V_{12}$	$\overline{V_2} \overline{V_6} V_{10} V_{14}$	$\overline{V_3} V_7 V_{11} \overline{V_{15}}$
Пояснение	Мн.	0	Гр.	0	Мн.
Сумма Отсчётов	2 (10,25)	-	6,5	-	2 (10,75)
Группа	$\overline{V_1} \overline{V_5} V_9 V_{13}$	$\overline{V_1} \overline{V_5} \overline{V_8} V_{12}$	$\overline{V_3} V_7 V_{10} V_{14}$	$\overline{V_2} \overline{V_6} V_{11} \overline{V_{15}}$	$\overline{V_1} V_4 \overline{V_8} \overline{V_{13}}$
Пояснение	Гр.	Гр.	0	Гр.	Гр.
Сумма Отсчётов	6,5	6,25	-	6,75	6,75

Окончание табл. 6

Группа	$\overline{V_3} \overline{V_6} V_{10} \overline{V_{15}}$	$\overline{V_2} V_7 V_{11} V_{14}$	$\overline{V_2} V_7 V_{10} \overline{V_{15}}$	$\overline{V_3} \overline{V_6} V_{11} V_{14}$	$\overline{V_1} V_4 V_9 V_{12}$
Пояснение	0	0	0	0	Гр.
Сумма Отсчётов	-	-	-	-	6,75

Контрольные вопросы

1. Опишите свойства замкнутых групп.
2. Опишите алгоритм генерации замкнутых групп
3. Алгоритм вычисления состава замкнутых групп по сигналу.

6. ВЫБОР РАЗМЕРНОСТИ ПРИЗНАКОВОГО ОПИСАНИЯ

Рассмотрим вопрос выбора количества фильтров, операторов и групп, используемых при построении описания звукового сигнала. В табл. 7 приведена зависимость числа полных и замкнутых групп от количества используемых операторов (фильтров): с увеличением числа используемых операторов (фильтров) увеличивается число полных и замкнутых групп. Поэтому при выборе количества операторов (фильтров) для формирования признакового описания сигнала необходимо учитывать, что с увеличением размерности признакового пространства размер выборки должен увеличиваться нелинейно («проклятие размерности»). При увеличении объёма обучающей выборки также возрастают требования к объёму используемой памяти и увеличивается время работы алгоритмов принятия решения.

В связи с этим для формирования описания сигнала допустимо использовать 8, 16 или 32 оператора (фильтра). Использование указанного количества признаков обеспечивает компромиссное решение между точностью описания сигнала, размером обучающей выборки и временем работы алгоритмов принятия решения. В колонке «Полные группы с инверсиями» показано число полных групп на операции сложения и умножения, в колонке «Замкнутые группы с инверсиями» – число замкнутых групп и замкнутых множеств (на операции сложения и умножения).

Таблица 7. Связь между количеством операторов и групп

№	Операторы	Полные группы	Замкнутые группы	Полные группы с инверсиями	Замкнутые группы с инверсиями
1	4	1	0	4 + 4	0
2	8	7	7	21 + 21	56 + 2 · 56
3	16	35	105	140 + 140	840 + 2 · 840
4	32	155	1085	620 + 620	8680 + 2 · 8680
5	64	651	9765	2604 + 2604	78120 + 2 · 78120
6	128	2667	82677	10668 + 10668	661416 + 2 · 661416
7	256	10795	680085	43180 + 43180	5440680 + 2 · 5440680

Количество замкнутых групп для заданного числа операторов связано с количеством полных групп следующим соотношением:

$$N_C = N_f (2^{\log_2 N_o - 2} - 1),$$

где N_c – количество замкнутых групп; N_f – количество полных групп; N_o – количество операторов. Минимальное количество операторов в табл. 7 выбрано равным четырём, так как полную группу можно построить только по трём операторам.

Контрольные вопросы

1. Приведите формулу для определения количества замкнутых групп по количеству операторов;
2. Опишите проблему проклятия размерности;
3. Опишите алгоритм выбора размерности признакового описания для решения задачи распознавания.

7. ФОРМИРОВАНИЕ ПРИЗНАКОВОГО ОПИСАНИЯ ЗВУКОВОГО СИГНАЛА

Полные и замкнутые группы удобно использовать для описания сегмента сигнала. Для получения описания сигнала в целом предлагается использовать «интегрирующие» описания, объединяющие описания отдельных сегментов сигнала.

Используя спектральное представление сигнала μ , можно сформировать описание сигнала на основе операторов, полных и замкнутых групп:

$$DV = GV[\mu], P_{na} = GP_{na}[\mu, V], P_{nm} = GP_{nm}[\mu, V], \\ P_s = GP_s[\mu, V, P_{na}, P_{nm}], P_c = GP_c[\mu, V, P_{na}, P_{nm}],$$

где GV – оператор вычисления по спектральному представлению сигнала признакового описания DV на основе операторов; GP_{na} (GP_{nm}) – на основе полных групп на операции сложения P_{na} (умножения, P_{nm}); GP_c (GP_s) – на основе замкнутых групп P_s (замкнутых множеств, P_c); P_{na} – множество полных групп на операции сложения; P_{nm} – множество полных групп на операции умножения; P_s – множество замкнутых групп; P_c – множество замкнутых множеств.

7.1. Гистограммы операторов, полных, замкнутых групп

Метод, предлагаемый для создания признакового описания в виде гистограммы:

1) отсчёты сигнала s_{dig} разбиваются на множество сегментов $s_{dig} = \{s_{dig, k}\}$, $k = \overline{1, N}$, длиной M отсчётов, со смещением в T отсчётов;

2) к каждому сегменту $s_{dig, k}$ применяется U -преобразование, в результате формируется спектральное представление каждого сегмента: $\mu_k = U[s_{dig, k}]$, $\mu = \{\mu_k\}$;

3) по вычисленному спектральному представлению μ_k сегмента $s_{dig, k}$ формируется описание с помощью одной или нескольких структур, входящих в алгебру групп. В алгебре групп существуют следующие структуры, которые могут использоваться для создания описания (допустимо использование сочетаний данных структур):

а) операторы (значимыми являются 15 операторов из 16, зная, что операторы могут принимать значения -1, 0 и 1, т.е. инверсный оператор, оператор отсутствует в описании сигнала и прямой оператор, пространство признаков на основе операторов включает 45 элементов; учитывая, что нулевые значения операторов встречаются не часто, можно использовать только значения -1 и 1, и, следовательно, сократить пространство признаков до 30 элементов);

б) полные группы (P_n), известны полные группы на операции умножения (P_{nm} , 140 элементов), полные группы на операции сложения (P_{na} , 140 элементов);

в) замкнутые группы (P_s , 840 элементов), замкнутые множества (P_{ca} – на операции сложения, 840 элементов, P_{cm} – на операции умножения, 840 элементов);

4) для объединения данных, полученных от разных сегментов анализируемого сигнала, вычисляется гистограмма элементов структур, использованных при создании описания сегмента сигнала (см. рис. 18, 19):

$$h_V = H [DV, \Gamma], h_{na} = H [P_{na}, \Gamma], h_{nm} = H [P_{nm}, \Gamma], h_s = H [P_s, \Gamma],$$

$$h_c = H [P_c, \Gamma], h_{nam} = H [P_{na}, P_{nm}, \Gamma], h_{sc} = H [P_s, P_c, \Gamma],$$

где h_V – гистограмма операторов; h_{na} – гистограмма полных групп на операции сложения; h_{nm} – гистограмма полных групп на операции умножения; h_s – гистограмма замкнутых групп; h_c – гистограмма замкнутых множеств; h_{nam} – гистограмма полных групп на операции сложения и умножения; H – оператор вычисления гистограммы заданной размерности; Γ – размерность гистограммы: $1d$ – одномерная гистограмма; $2d$ – двумерная гистограмма; $3d$ – трёхмерная гистограмма.

В двумерной гистограмме учитываются частота возможных появлений пар групп (операторов) в описании одного сегмента сигнала, в трёхмерной – троек.

Далее приведены алгоритмы формирования гистограмм различных размерностей. Признакомое описание, на основе которого формируется гистограмма h_D , обозначается как D . В качестве данного описания могут использоваться операторы, полные и замкнутые группы (рис. 18, 19).

Алгоритм формирования одномерной гистограммы признаков:

$$\forall i = \overline{1, M}$$

$$\forall j = \overline{1, |D_i|}$$

$$h_D(D_{i,1}[j]) = h_{D,1}(D_i[j]) + 1.$$

Алгоритм формирования двумерной гистограммы признаков:

$$\forall i = \overline{1, M}$$

$$\forall j = \overline{1, |D_i|}$$

$$\forall k = \overline{(j+1), |D_i|}$$

$$h_{D,2}(D_i[j], D_i[k]) = h_{D,2}(D_i[j], D_i[k]) + 1.$$

Алгоритм формирования трёхмерной гистограммы признаков:

$$\forall i = \overline{1, M}$$

$$\forall j = \overline{1, |D_i|}$$

$$\forall k = \overline{(j+1), |D_i|}$$

$$\forall l = \overline{(k+1), |D_i|}$$

$$h_{D,3}(D_i[j], D_i[k], D_i[l]) = h_{D,3}(D_i[j], D_i[k], D_i[l]) + 1.$$

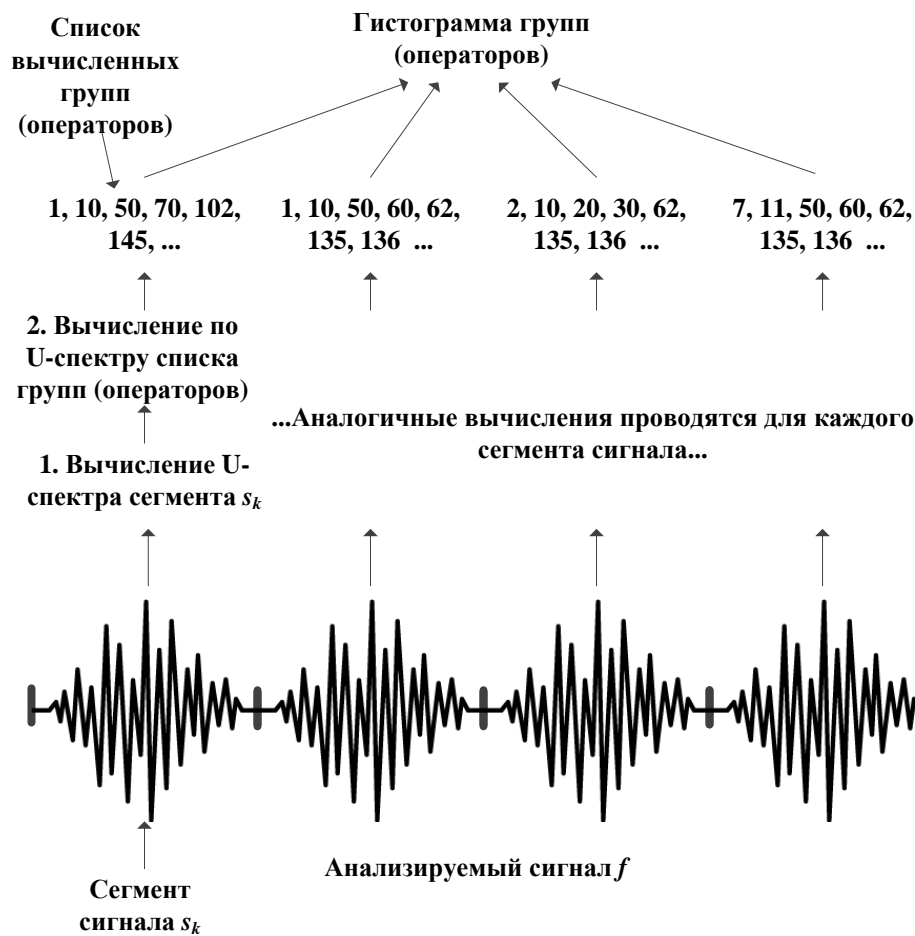


Рис. 18. Формирование одномерной гистограммы признаков

7.2. Матрицы вероятностей переходов

Предлагается модель признакового описания, в которой учитываются связи между соседними сегментами сигнала. Метод вычисления признакового описания в таком случае состоит в формировании матрицы вероятностей переходов между описаниями соседних сегментов.

Предлагаются следующие системы признаков, основанные на матрицах вероятностей переходов:

1) система признаков *PVI*, описывающая вероятности переходов между значениями операторов (оператор может принимать три возможных значения: прямое, инверсное и равное нулю), вычисленными по соседним сегментам сигнала, без учёта связей между различными операторами, размерность пространства признаков – $3 \times 3 \times 15$ (рис. 20);

2) система признаков *PVD*, описывающая вероятности переходов между значениями операторов, вычисленными по соседним сегментам сигнала, с учётом связи между операторами, размерность пространства признаков – 45×45 (рис. 21);

3) система признаков, описывающая вероятности переходов между описаниями сегментов, представленных в виде полных групп, размерность пространства признаков – 140×140 ; при использовании полных групп допустимо использовать только несколько максимальных по сумме отсчётов, находящихся под их образами, групп (граф переходов для полных групп подобен графу переходов для операторов, рис. 21);

4) система признаков, описывающая вероятности переходов между описаниями сегментов, представленных в виде замкнутых групп, размерность пространства признаков – 840×840 ; при использовании замкнутых групп допустимо использовать группы, сумма отсчётов под образом которых максимальна.

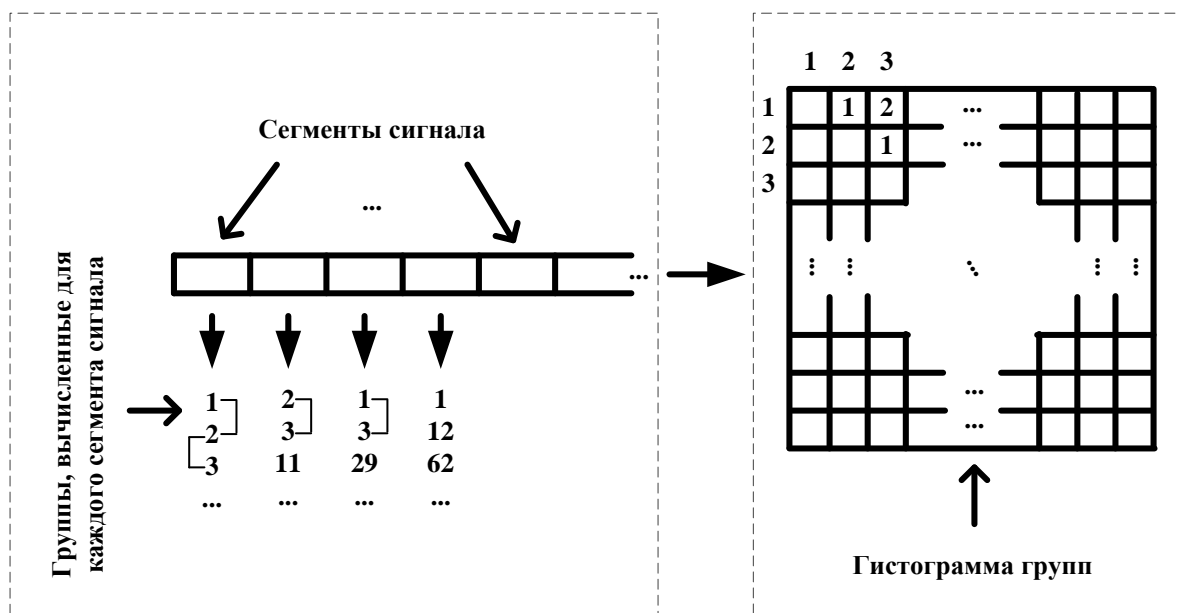


Рис. 19. Формирование двумерной гистограммы признаков

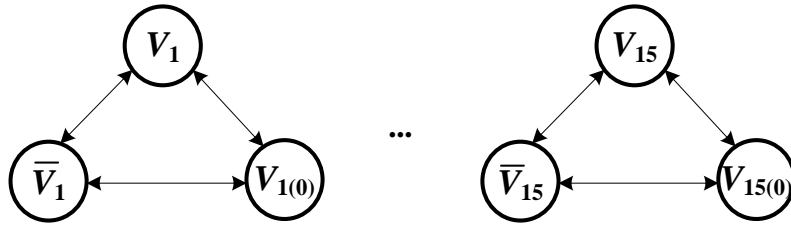


Рис 20. Графы переходов между операторами
(без учёта связей между операторами)

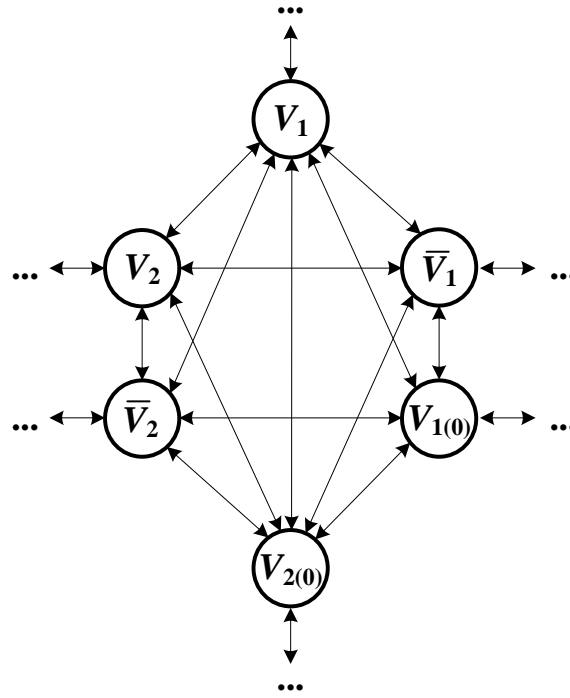


Рис 21. Граф переходов между операторами
(с учётом связей между операторами)

Алгоритм формирования матрицы вероятностей переходов на основе полных (замкнутых) групп, D_i – описание i -го сегмента сигнала в виде полных (замкнутых) групп:

$$\forall i = \overline{1, (N-1)}$$

$$\forall k = \overline{1, |D_i|}$$

$$\forall l = \overline{1, |D_{i+1}|}$$

$$MPG_{D_i}[D_i[k], D_{i+1}[l]] = MPG_{D_i}[D_i[k], D_{i+1}[l]] + 1.$$

После выполнения описанного алгоритма выполняется нормировка строк матрицы таким образом, чтобы сумма элементов в строке равнялась единице.

Далее будут использоваться следующие обозначения: PP_{na} (PP_{nm}) – матрица вероятностей переходов между полными группами на операции

сложения (умножения); PP_s – между замкнутыми группами; PP_{ca} – между замкнутыми множествами на операции сложения; PP_{cm} – между замкнутыми множествами на операции умножения.

При формировании описания сигнала на основе матрицы вероятностей переходов можно учитывать связи не только между i и $(i + 1)$ сегментом сигнала, но и между другими соседними сегментами.

Учитывать связь не только между i и $(i + 1)$ сегментом, но также между i -м сегментом и другими его соседями, например, между i -м и $(i + 1)$, $(i + 2)$, $(i - 1)$, $(i - 2)$ сегментами.

Пример формирования признакового описания сигнала в виде матрицы вероятностей переходов размером 140×140 элементов для полных групп показан на рис. 22. При вычислении значений матрицы рассматривались связи между описаниями только пары соседних сегментов.

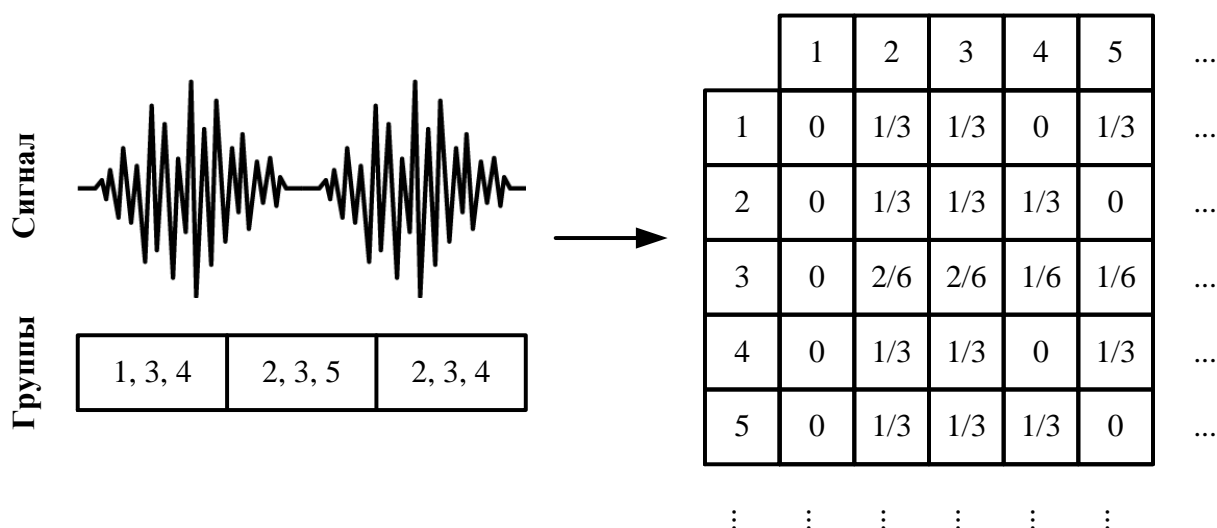


Рис. 22. Матрица вероятностей переходов между группами

Пример формирования признакового описания сигнала в виде 15 независимых матриц вероятностей переходов для операторов показан на рис. 23 (отображены только матрицы для двух операторов). При вычислении значений матрицы рассматривались связи между описаниями только пары соседних сегментов (i -го и $(i+1)$ -го) и учитываются вероятности переходов только для одного оператора.

Пример формирования признакового описания сигнала в виде матрицы вероятностей переходов для операторов размером 45×45 показан на рис. 24. При вычислении значений матрицы рассматривались связи между описаниями только пары соседних сегментов (i -го и $(i+1)$ -го). В описаниях сегментов приведены значения только операторов V_1 и V_2 .

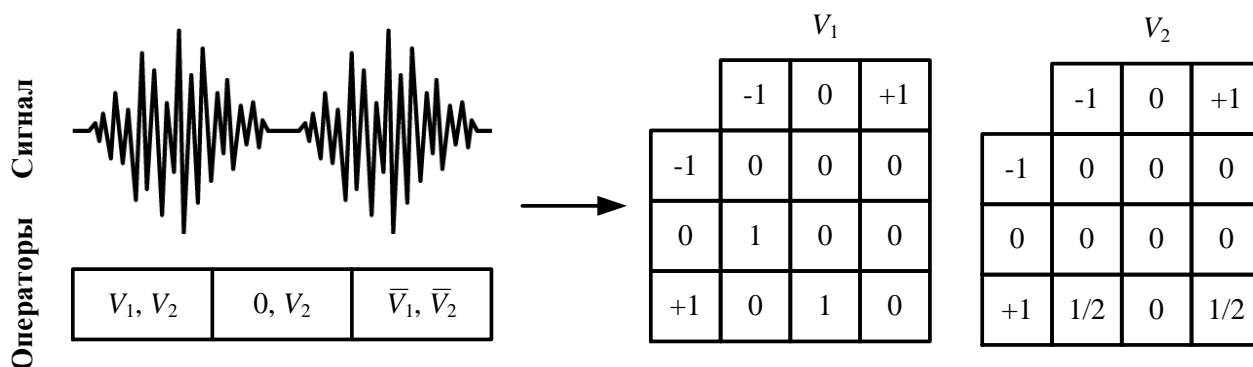


Рис. 23. Матрицы вероятностей переходов между операторами

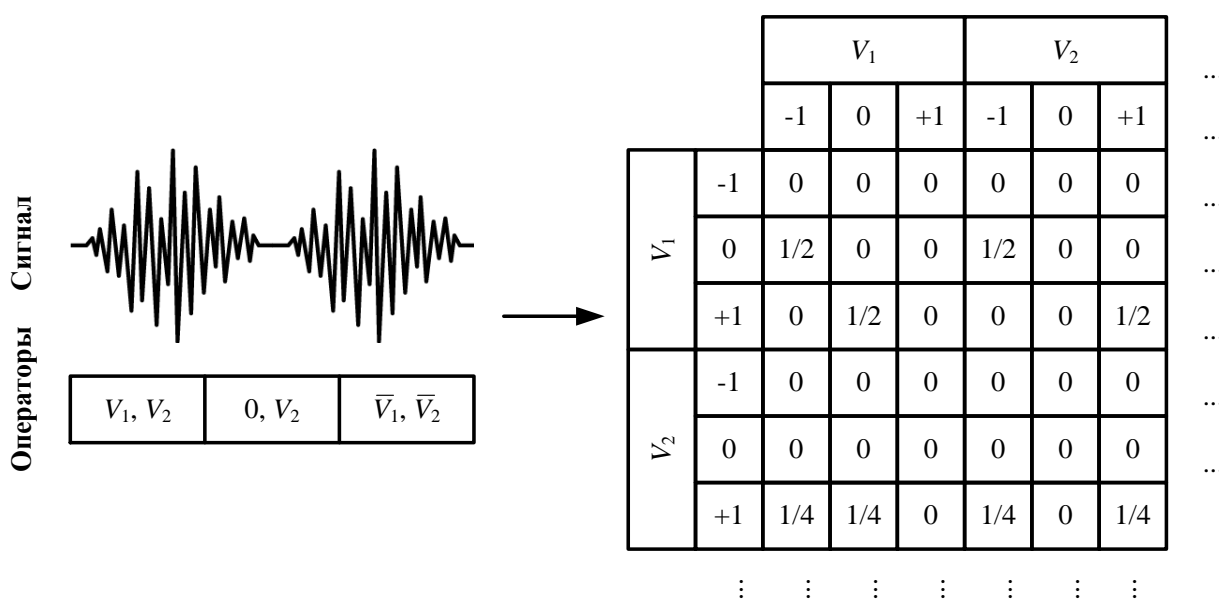


Рис. 24. Матрица вероятностей переходов между операторами

7.3. Выбор систем признаков для описания сигналов

В табл. 8 приведена размерность систем признаков, построенных на основе вероятностей переходов.

Таблица 8. Число элементов в системах признаков, построенных на основе вероятностей переходов

Количество операторов	PVD	PVI	$PP_{na} (PP_{nm})$	$PP_s (PP_{ca}, P_{cm})$
8	63	21^2	21^2	56^2
16	135	45^2	140^2	840^2
32	279	93^2	620^2	8680^2

В табл. 9 приведена размерность систем признаков, построенных на основе гистограмм.

Размерность приведённых в табл. 9 систем признаков ограничена тремя измерениями, так как с увеличением количества используемых измерений рост числа элементов в признаковом описании сигнала описывается показательной функцией.

Таблица 9. Число элементов в системах признаков, построенных на основе гистограмм

	Количество операторов	$1d$	$2d$	$3d$
h_V	8	14^1	14^2	14^3
	16	30^1	30^2	30^3
	32	62^1	62^2	62^3
h_{na}, h_{nm}	8	21^1	21^2	21^3
	16	140^1	140^2	140^3
	32	620^1	620^2	620^3
h_s, h_c	8	56^1	56^2	56^3
	16	840^1	840^2	840^3
	32	8680^1	8680^2	8680^3
h_{nam}	8	42^1	42^2	42^3
	16	280^1	280^2	280^3
	32	1240^1	1240^2	1240^3
h_{sc}	8	112^1	112^2	112^3
	16	1680^1	1680^2	1680^3
	32	17360^1	17360^2	17360^3

Серым цветом в табл. 8 и табл. 9 выделены системы признаков, использование которых накладывает высокие требования на производительность и объём памяти вычислительной машины.

Можно отметить следующие свойства систем признаков, построенных на основе гистограмм:

- 1) нулевая главная диагональ;
- 2) симметричность относительно главной диагонали.

Таким образом, без потери информации можно не учитывать элементы, находящиеся на главной диагонали и под главной диагональю.

Также при построении систем признаков, приведённых в табл. 8 и 9, необходимо учитывать:

- 1) несовместимость операторов: два оператора называются несовместимыми, если они имеют одинаковые номера, но разные знаки (например, операторы V_1 и \bar{V}_1);

2) несовместимость групп: две группы (полные или замкнутые) называются несовместными, если в их состав входят операторы с одинаковыми номерами, но с разными знаками.

Учитывая несовместимость операторов и групп, в двумерных и трёхмерных гистограммах, а также в матрицах вероятностей переходов нулевыми элементами будут те, которые стоят на пересечении строки и столбца, соответствующих несовместимым операторам или группам.

В табл. 10 приведены результаты сокращения размерности некоторых из описанных систем признаков.

Таблица 10. Результаты сокращения размерности систем признаков

Система признаков	Размер до сжатия	Размер после сжатия	Коэффициент сжатия
$PP_{na} (PP_{nm})$	19600	14140	1,38
$h_{na}, h_{nm} (2d)$	19600	7000	2,80
$h_V (2d)$	900	420	2,14
$h_V (3d)$	27000	3640	7,41

Контрольные вопросы

1. Опишите алгоритм вычисления признакового описания в виде гистограммы операторов.
2. Опишите алгоритм вычисления признакового описания в виде гистограммы полных групп.
3. Опишите алгоритм вычисления признакового описания в виде гистограммы замкнутых групп.
4. Опишите алгоритм вычисления признакового описания в виде матрицы вероятностей переходов операторов.
5. Опишите алгоритм вычисления признакового описания в виде матрицы вероятностей переходов полных групп.
6. Опишите алгоритм вычисления признакового описания в виде матрицы вероятностей переходов замкнутых групп.

8. РЕШЕНИЕ ЗАДАЧ ОБРАБОТКИ И РАСПОЗНАВАНИЯ ОДНОМЕРНЫХ СИГНАЛОВ

8.1. Выбор длины сегмента сигнала

Метод выбора длины сегмента сигнала основан на использовании понятия дихотомии и стратегии восприятия от целого к частному и заключается в следующем:

1. Сначала на основе операции дихотомии и U -преобразования формируется многоуровневое представление анализируемого сигнала:

$$\mathbf{D} = \{D_i\}, D_i = U[f, L_i], i = \overline{1, N},$$

где f – анализируемый сигнал; L_i – длина сегмента сигнала на i -м уровне представления; D_i – спектральное представление сигнала на i -м уровне представления; N – количество уровней представления анализируемого сигнала;

2. Затем вычисляется среднеквадратическое отклонение спектральных коэффициентов на каждом уровне разложения:

$$\sigma_i = STD[D_i], i = \overline{1, N},$$

где $STD[\bullet]$ – оператор оценки стандартного отклонения;

3. В результате, выбирается длина сегмента сигнала, соответствующая уровню, на котором значение СКО максимально:

$$\sigma_s = \arg \max_{i \in \overline{1, N}} STD[D_i],$$

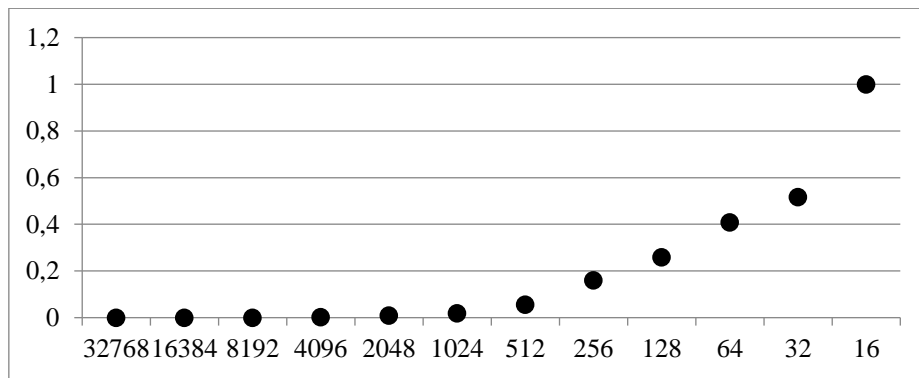
где σ_s – уровень представления с максимальным стандартным отклонением.

Поясним идею этого подхода. После операции суммирования получаем структурные элементы, после вычитания – отношения между структурными элементами, т.е. отклики фильтров. Если СКО откликов фильтра маленькое, это значит, что структурные элементы похожи друг на друга и не отличаются, следовательно, нужно опуститься на уровень вниз, чтобы уточнить сигнал.

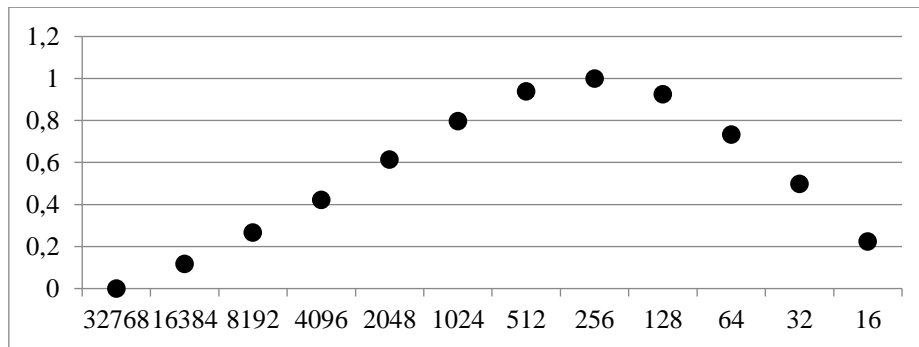
На рис. 25 показаны графики, построенные для сигналов различной физической природы, как видно, они отличаются.

На определённом уровне разложения получается пик СКО, который указывает на то, что отклик фильтров на данный сигнал максимален. Затем, как видно идёт спад. Это следствие того, что из-за уточнения, наряду с большими откликами, появляется очень маленьких откликов фильтра, т.е. шума. Следовательно, лучше прекратить анализ на том уровне, на котором достигнут пик СКО.

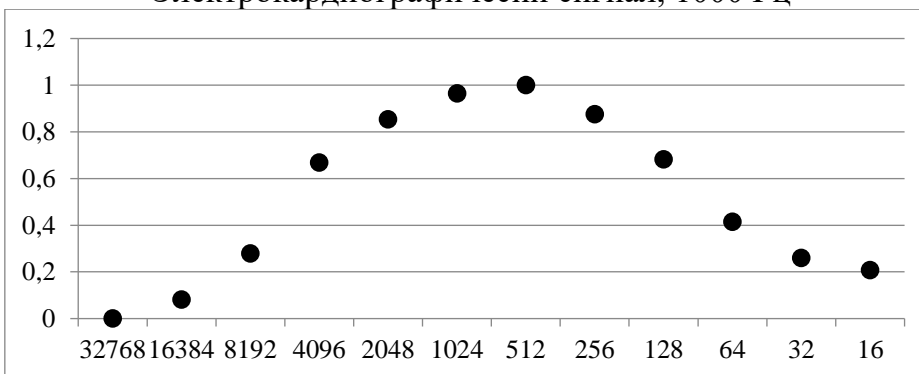
Рекомендованные значения длины сегмента сигнала сконцентрированы около максимума графика.



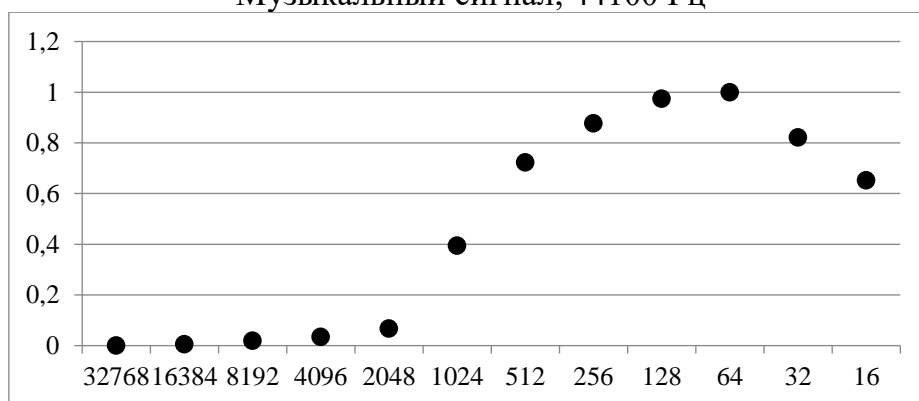
Виброакустический сигнал, 12000 Гц



Электрокардиографический сигнал, 1000 Гц



Музыкальный сигнал, 44100 Гц



Речевой сигнал, 16000 Гц

Рис. 25. График зависимости СКО откликов фильтров от длины сегмента сигнала

8.2. Оценка частоты основного тона

Задача оценки частоты голоса является одной из основных задач цифровой обработки речевых сигналов. Результаты её решения используются при распознавании и сжатии речи, идентификации / верификации диктора и т.д. Также алгоритмы вычисления частоты основного тона нашли своё применение при настройке музыкальных инструментов, создании нотной записи и классификации музыкальных произведений. Существующие алгоритмы определения основного тона выполняют обработку звукового сигнала во временной и частотной областях.

Рассмотрим алгоритм оценки частоты основного тона:

1. Построение на основе U -преобразования спектрального представления D сигнала f .

2. Формирование огрубленного представления $C = \{c_{ij}\}$ спектрального представления D (степень огрубления – параметр k); огрубление спектрального представления используется для создания единообразного описания каждого спектра;

3. Выбор ключевых сегментов сигнала f : анализируется каждый спектр c_{ij} огрублённого разложения: если $\forall t \in [1; L]$ верно, что $|c_{ij}(t)| \leq k$ или $|c_{ij}(t)| \geq l$ ($k < l$), то сегмент сигнала f_{ij} , для вычисления спектра d_{ij} , считается ключевым и будет использоваться в дальнейших вычислениях:

$$K_f = K_f \cup f_{ij},$$

где K_f – набор ключевых сегментов сигнала f .

Ключевой сегмент соответствует участку сигнала, на котором максимален отклик небольшого количества операторов (одного - двух). Это означает, что сегмент сигнала подходит под образ операторов, давших максимальные отклики. На рис. 26 показан пример ключевого сегмента и вычисленного по нему спектра (нулевой коэффициент спектра на рисунке не отображается, при построении спектра использовалось 16 фильтров).

4. Вычисление по каждому ключевому сегменту первых N по массе замкнутых групп:

$$M_f = \bigcup_{i \in \overline{1, N}} P_{cs, i},$$

где $P_{cs, i}$ – набор замкнутых групп, вычисленных по сегменту i .

5. Сравнение групп, сформированных по i -му и j -му сегментам сигнала: если состав групп совпадает не менее чем на M %, тогда считается, что сегменты сигнала подобны (с некоторой точностью) и расстояние $dist$ между сегментами сохраняется в массив H :

$$H(dist) = H(dist) + 1.$$

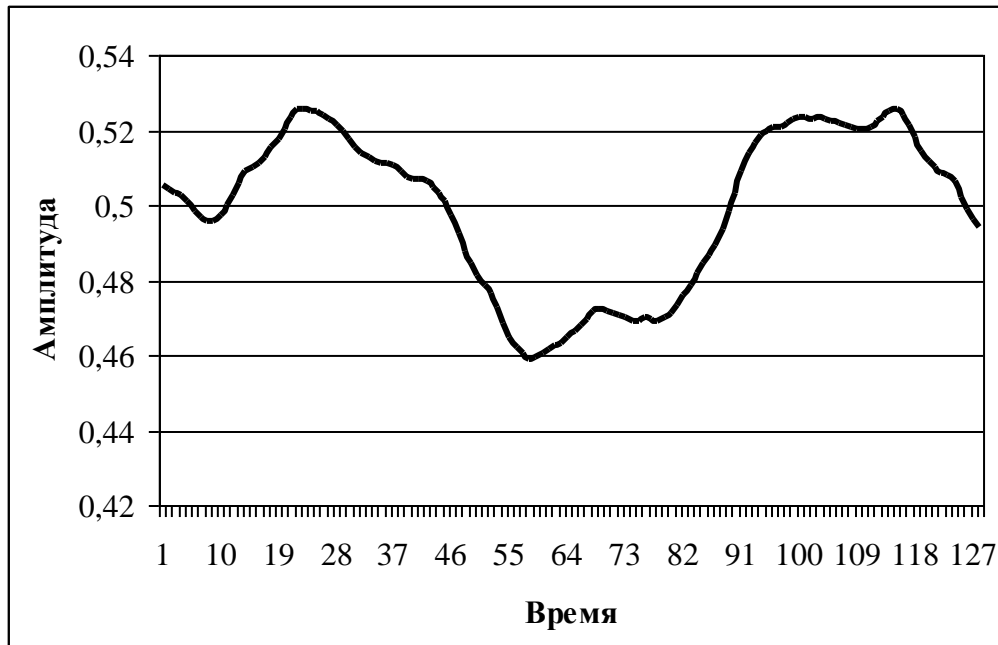
6. Определение номера максимального элемента X_{\max} гистограммы H , который соответствует наиболее часто встречающемуся расстоянию между похожими спектрами (расстояние измеряется в отсчётах):

$$X_{\max} = \max(H).$$

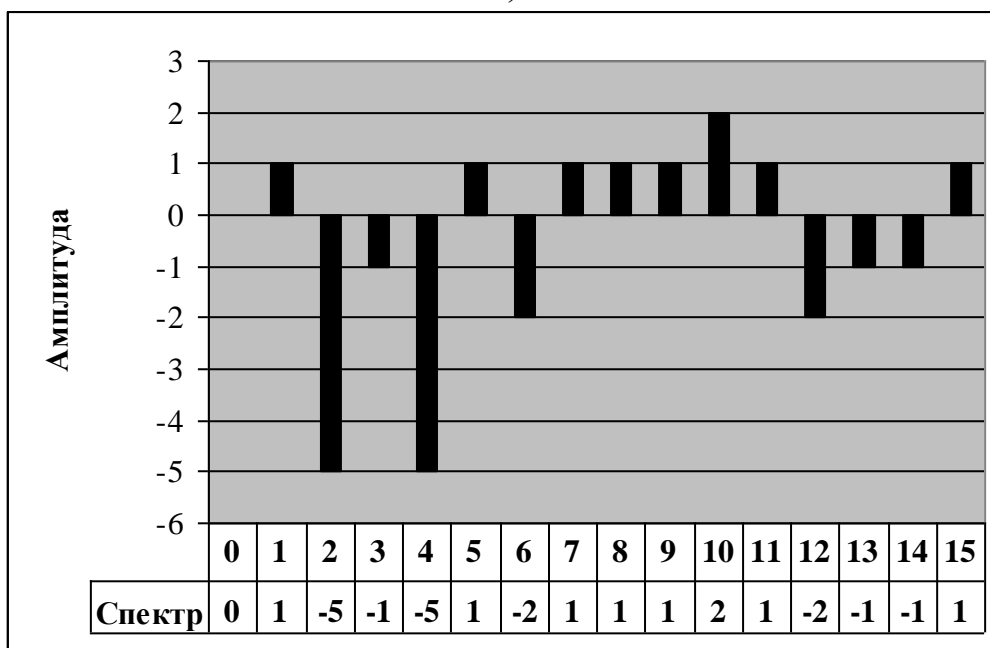
7. Вычисление частоты основного тона:

$$P = \frac{F_{\text{д}}}{X_{\max}},$$

где P – частота основного тона; $F_{\text{д}}$ – частота дискретизации сигнала.



а)



б)

Рис. 26. Анализ сигнала: а – ключевой сегмент; б – спектр

Рассмотрим алгоритм огрубления спектрального представления D сигнала f . Входные данные алгоритма – представление D и степень огрубления k :

1) формирование множества чисел T : $\{0, 1/k, 2/k, \dots, (k-1)/k, 1\}$ (например, для $k = 3$ получим следующее множество: $T = \{0, 0.33, 0.66, 1\}$) и множества промежутков $W = \{(0; 1/k), [1/k; 2/k), \dots, [(k-1)/k, 1]\}$, каждому из которых назначается номер $N \in \overline{1, k}$;

2) обработка спектрального представления $D = \{d_{ij}\}$:

а) выбор максимального по абсолютной величине элемента спектра:

$$P = \max(d_{ij}), i = \overline{1, K}, j \in \overline{1, M_i};$$

б) умножение каждого элемента множества T на P (отклик нулевого фильтра не учитывается):

$$T_i = T_i \cdot P, i = \overline{1, |T|};$$

Допустим, $P = 23$, тогда $T = \{0, 7.66, 15.33, 23\}$;

в) огрубление спектрального коэффициента $d_{ij}(k)$:

$$c_{ij}(k) = \begin{cases} \text{sign}(d_{ij}(k)) \cdot N, & \text{если } |d_{ij}(k)| \in (a_N, b_N); \\ 0, & \text{если } d_{ij}(k) = 0. \end{cases}$$

В приведённом правиле определяется промежуток $((0; 7.66), [7.66, 15.33), [15.33, 23])$, к которому относится k -й спектральный коэффициент. Далее, в зависимости от полученного ответа, и с учётом знака коэффициента $d_{ij}(k)$ рассматриваемому элементу $c_{ij}(k)$ огрублённого спектра присваивается одно из значений: -3, -2, -1, 0, 1, 2, 3.

На рис. 27 показан пример огрубления спектрального разложения синуса на 4-м уровне (частота – 33 Гц, длина сигнала – 4 периода, $P = 488$).

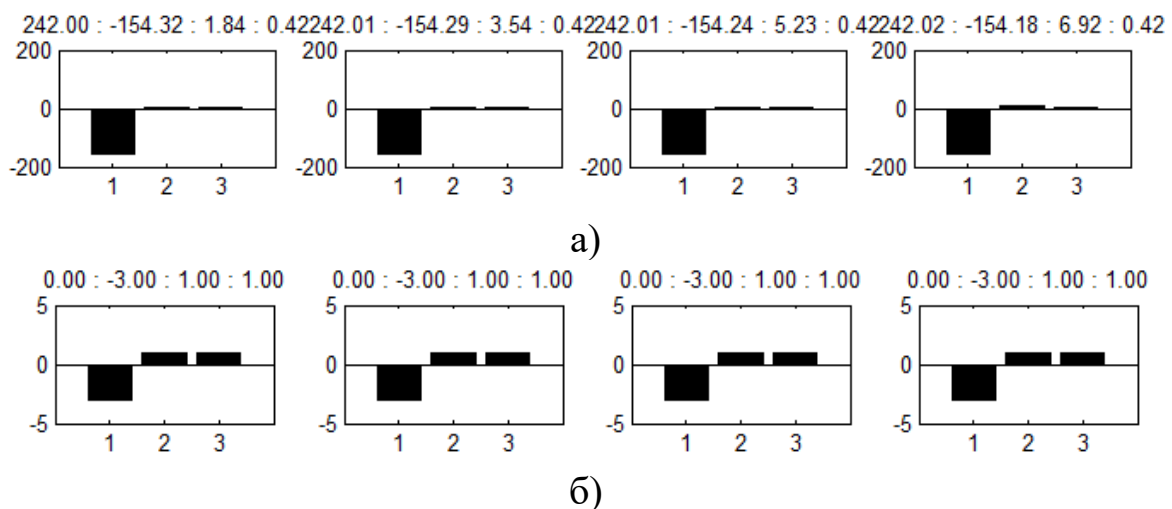


Рис. 27. Спектральное представление сигнала: а - исходное; б - огрублённое

Оценку качества вычисления частоты основного тона можно выполнить с помощью следующих показателей:

1) количество значительных ошибок оценки основного тона (gross pitch error):

$$G = \frac{1}{N_V} \sum_{i \in V} \delta_G(F_m(i), F_a(i)),$$

$$\delta_G(F_m(i), F_a(i)) = \begin{cases} 1, & |(F_m(i) - F_a(i)) / F_m(i)| > T, \\ 0, & \text{иначе;} \end{cases}$$

2) среднее значение незначительной ошибки оценки основного тона (fine pitch error): $S = \frac{1}{N_V} \sum_{i \in V} \delta_S(F_m(i), F_a(i))$,

$$\delta_S(F_m(i), F_a(i)) = \begin{cases} |F_m(i) - F_a(i)|, & |(F_m(i) - F_a(i)) / F_m(i)| \leq T, \\ 0, & \text{иначе;} \end{cases}$$

3) среднеквадратическое отклонение незначительной ошибки оценки основного тона: $S_\sigma = \sqrt{\frac{1}{N_V} \sum_{i \in V} \delta_\sigma(F_m(i), F_a(i), S)}$,

$$\delta_\sigma(F_m(i), F_a(i), S) = \begin{cases} (F_m(i) - F_a(i))^2 - S^2, & |(F_m(i) - F_a(i)) / F_m(i)| \leq T, \\ 0, & \text{иначе.} \end{cases}$$

Используемые обозначения: $F_m(i)$ – ручная оценка частоты основного тона i -го отсчёта сигнала; $F_a(i)$ – автоматическая оценка частоты основного тона i -го отсчёта сигнала; T – порог, разделяющий грубые и незначительные ошибки оценки частоты основного тона; V – множество отсчётов исходного сигнала, относящихся к вокализованным сегментам; N_V – мощность множества V .

8.3. Оценка отношения сигнал / шум

Системы обработки речевых сигналов (идентификации диктора, распознавания речи), работающие в условиях помех, должны обладать устойчивостью к возможным искажениям входного сигнала. Поэтому для настройки алгоритмов обработки сигналов на уровень шума такие системы

должны обладать возможностью оценки уровня искажения сигнала (отношения сигнал / шум, ОСШ). Причём такая оценка должна выполняться только по искажённому сигналу.

Рассмотрим характерные черты существующих методов оценки отношения сигнал / шум:

1) при оценке ОСШ, анализируемый сигнал разбивается на сегменты длительностью 20-30 мс., перекрытие между соседними сегментами составляет 50 %, затем вычисляется спектр каждого сегмента и по спектрам выполняется оценка уровня шума;

2) методы оценки ОСШ разрабатываются с учётом того, что анализируемый сигнал содержит и речь и паузы.

Анализ работ позволяет выделить следующие классы методов оценки ОСШ:

1) методы на основе определения активности голоса заключаются в сегментации сигнала на активную речь и паузы, оценке мощности сигнала и шума, и вычислении ОСШ;

2) методы, использующие управляемое рекурсивное усреднение, в таких методах оценка шума выполняется с помощью усреднения предыдущих значений спектральной мощности с помощью сглаживающего параметра, который зависит от вероятности присутствия сигнала в различных полосах частот: показано, что присутствие речи в некотором сегменте, в определённой полосе частот, может быть определено по отношению локальной энергии зашумлённой речи к её минимуму в данном сегменте. В случае если значение этого отношения меньше порога, можно сделать вывод об отсутствии речевого сигнала в сегменте;

3) методы оценки шума на основе минимальной статистики, методы, относящиеся к данному классу, основаны на двух предположениях. Первое заключается в независимости шума и речи, второе – в том, что спектр мощности зашумлённого речевого сигнала подобен спектру мощности шума. Поэтому оценка дисперсии шума состоит в вычислении минимума спектральной плотности зашумлённого речевого сигнала в сегменте фиксированной длины.

К недостаткам метода можно отнести необходимость выбора длины сегмента. При этом неверный выбор может значительно повлиять на результат оценки. Ещё одним из недостатков метода является внесение задержки в оценку параметров шума, так как длина сегмента (1.6 – 2.8 с) выбирается таким образом, чтобы гарантированно включать и участки речи и паузы;

4) методы оценки параметров шума на основе гистограммы используют наблюдение о том, что наиболее часто встречающееся значение энергии в некоторой полосе частот соответствует уровню шума в данной полосе

частот, т. е. уровень шума соответствует максимальному значению гистограммы энергии.

Проведённые исследования позволили установить свойства спектров (в рамках U -преобразования), относящихся к полезному сигналу и паузе. Полезный сигнал можно представить в виде совокупности вокализованных и невокализованных сегментов (рис. 28):

1) спектр сегмента сигнала, относящегося к вокализованному звуку, содержит элементы, значительно отличающиеся друг от друга по величине (участок 3);

2) спектр сегмента сигнала, относящегося к вокализованному звуку, содержит элементы, отличающиеся друг от друга по величине, но отличающиеся в меньшей степени, чем для вокализованного сегмента (участок 2);

3) элементы спектра сигнала, относящегося к паузе, близки к нулевым значениям (участок 1).

Рассмотрим таблицу (табл. 11), в которой приводятся значения спектральных коэффициентов для различных участков сигнала и их среднеквадратическое отклонение (СКО). Можно отметить, что в результате воздействия шума на сигнал:

1) на множестве спектральных коэффициентов (при определённом уровне шума) становится невозможно отличить невокализованные участки речевого сигнала от участков пауз (это также соответствует субъективному восприятию человеком зашумлённого речевого сигнала);

2) значения спектральных коэффициентов, относящихся к вокализованным сегментам, уменьшаются при увеличении уровня шума.

Известно, что при увеличении уровня шума ритмический рисунок слова или фразы является параметром, который разрушается в последнюю очередь. Следовательно, при оценке ОСШ можно использовать сегменты сигнала, для которых среднеквадратическое отклонение спектральных коэффициентов максимально (полезный сигнал) и минимально (помеха).

Пусть на конечном временном интервале $[0; T]$ принимается звуковой сигнал $f(t)$, являющийся функцией от речевого сигнала $s(t, \lambda)$ и некоррелированного шума $n(t)$:

$$f(t) = F(s(t, \lambda), n(t)), 0 \leq t \leq T,$$

где $\lambda = \{\lambda_1, \dots, \lambda_m\}$ – вектор параметров речевого сигнала. Предполагается, что непосредственному наблюдению доступен только принимаемый сигнал $f(t)$. Пусть $s_a(t, \lambda)$ – активная речь без пауз, тогда оценку отношения сигнал / шум (ξ) можно записать следующим образом:

$$\xi = \frac{E\{s_a^2(t, \lambda)\}}{E\{n^2(t)\}},$$

где E – оператор вычисления математического ожидания.

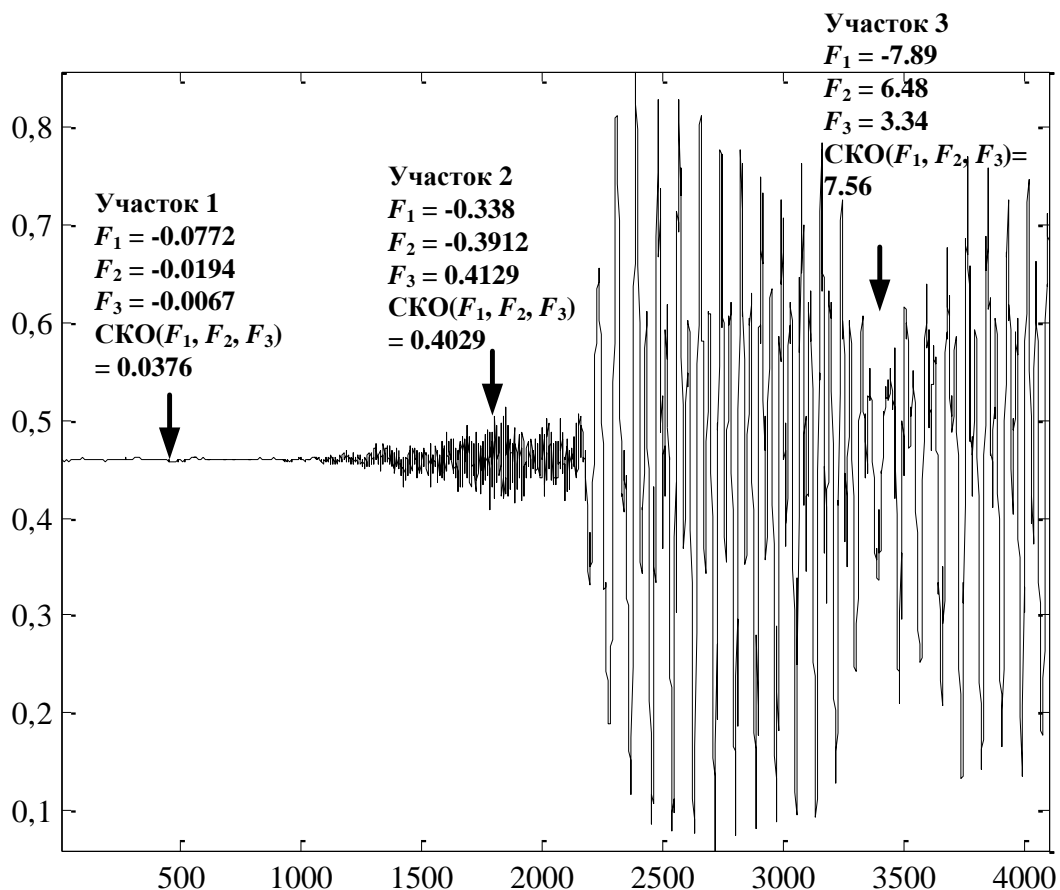


Рис. 28. Свойства полезного сигнала и паузы

Таблица 11. Влияние шума на речевой сигнал

Участок сигнала	Равномерный шум (ОСШ = 8 дБ)			Нормальный шум (ОСШ = 4 дБ)		
	1	2	3	1	2	3
Эксперимент № 1 (длина сегмента = 8 мс.)						
$CKO(F_1, F_2, F_3)$	0.4747	1.3124	13.0893	1.0637	1.2250	7.0924
F_1	0.6516	-1.1670	-3.3792	0.8600	0.5248	1.5579
F_2	0.8198	-0.0529	17.7059	1.6997	-1.9050	10.2402
F_3	-0.0735	1.4483	-6.2735	-0.4129	-0.4175	-3.8153
Эксперимент № 2 (длина сегмента = 8 мс.)						
$CKO(F_1, F_2, F_3)$	0.7981	0.5843	12.7010	1.3569	1.3808	5.7403
F_1	-0.7303	1.0742	-4.3398	0.7722	3.3808	0.2616
F_2	0.1213	0.9121	15.6942	-1.9404	1.2583	7.9577
F_3	-1.4737	1.9954	-7.8480	-0.5106	0.7893	-3.2680

Предлагаемый метод оценки ОСШ (на i -ом уровне разложения) включает в себя следующие операции:

1) определение среднеквадратического отклонения (СКО) спектральных коэффициентов, относящихся к полезному сигналу:

$$\hat{\sigma}_s = \max \left[\left\{ \sigma \left(\{d_{ij}(t)\} \right) \right\} \right], \quad j = \overline{1, 2^{i-1}}, \quad t = \overline{1, N};$$

2) определение СКО спектральных коэффициентов, относящихся к помехе:

$$\hat{\sigma}_n = \min \left[\left\{ \sigma \left(\{d_{ij}(t)\} \right) \right\} \right], \quad j = \overline{1, 2^{i-1}}, \quad t = \overline{1, N}.$$

3) вычисление оценки ОСШ $\hat{\xi}$:

$$\hat{\xi} = 10 \log_{10} \frac{(\hat{\sigma}_s - \hat{\sigma}_n)^2}{\hat{\sigma}_n^2}.$$

Предлагаемый метод имеет два параметра:

- 1) длина сегмента сигнала (в миллисекундах) по которому вычисляется спектр;
- 2) количество фильтров, используемых при вычислении спектра.

8.4. Выделение активности голоса

Звуковой сигнал, с точки зрения наблюдателя, содержит полезный сигнал (информационное сообщение) и помеху. Полезный сигнал – составляющая звукового сигнала, которая несёт для наблюдателя нужную информацию, т.е. информацию, которая требуется наблюдателю для принятия решения в рамках поставленной задачи, а помеха – вся остальная информация. Полезный сигнал в рамках данной работы рассматривается как системное образование. В этом случае он должен содержать структурные элементы и связи. Следовательно, задачу обнаружения полезного сигнала можно свести к задаче выделения структурных элементов сигнала и связей между ними.

Известно, что без наличия предварительных сведений о сигнале невозможно отличить полезный сигнал от помехи.

Задача выделения информационного сообщения из принятого сигнала в статистической радиотехнике соответствует задаче обнаружения сигнала.

Пусть на конечном временном интервале $[0; T]$ принимается звуковой сигнал $\xi(t)$, являющийся функцией от полезного сигнала $s(t, \lambda)$ и помехи $n(t)$:

$$f(t) = F(s(t, \lambda), n(t)), \quad 0 \leq t \leq T,$$

где $\lambda = \{\lambda_1, \dots, \lambda_m\}$ – вектор параметров сигнала. Предполагается, что непосредственному наблюдению доступен только принимаемый сигнал $f(t)$.

Пусть неизвестен сам факт наличия или отсутствия сообщения $s(t, \lambda)$ в принятом сигнале $f(t)$. Запишем принятый сигнал $f(t)$ (1) в следующем виде:

$$f(t) = \theta \cdot s(t, \lambda) + n(t), 0 \leq t \leq T.$$

Здесь θ – случайная величина, которая может принимать два значения: $\theta = 0$ (сообщение отсутствует) и $\theta = 1$ (сообщение присутствует). Требуется по принятой реализации $f(t)$ на интервале T решить: присутствует или отсутствует полезный сигнал $s(t, \lambda)$.

Описанная задача представляет собой типичную формулировку задачи обнаружения сигнала на фоне помехи.

Принятие решения о наличии или отсутствии полезного сигнала всегда сопровождается ошибками двух видов:

- 1) принимается неправильное решение о наличии сигнала (ошибка первого рода);
- 2) принимается ошибочное решение об отсутствии сигнала (ошибка второго рода).

Таким образом, целью работы является – выявление системных элементов сигнала и связей между ними.

Классический подход к обнаружению полезного сигнала с позиций статистической радиотехники, построенный на основе корреляционного приёмника (рис. 29), предполагает решение уравнения Фредгольма первого рода:

$$g(\bar{T}) = \int_{x \in \bar{T}} s(t, x) f(x) dx,$$

где $g(t)$ – наблюдаемый сигнал; $f(x)$ – априори неизвестное изменение во времени амплитуды регистрируемого сигнала на входе устройства; $s(t, x)$ – полезный сигнал.

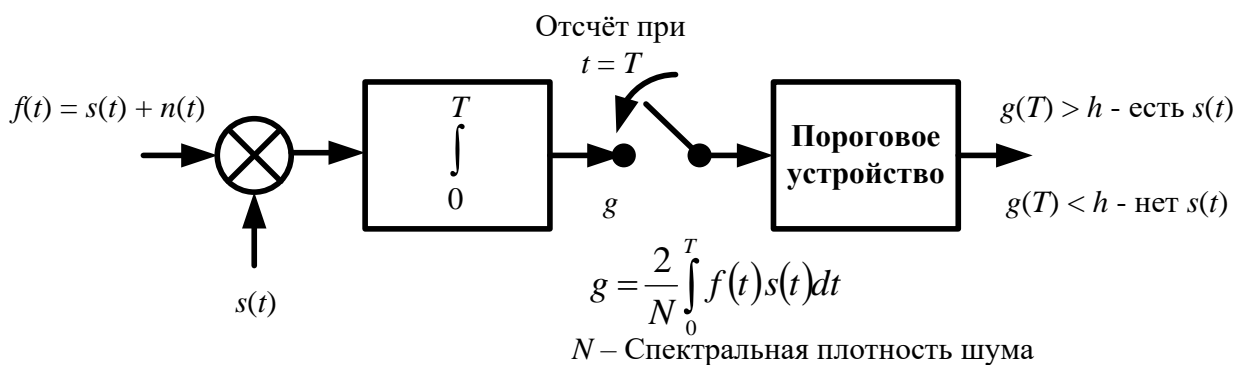


Рис. 29. Корреляционный приёмник

Незначительные изменения в регистрируемом сигнале $f(t)$ могут привести к недопустимо большим изменениям в решении. Поэтому результат, получаемый после выполнения фильтрации, в рамках использования уравнения Фредгольма, является некорректным. Под некорректно поставленной понимается такая задача, для которой не выполняется хотя бы одно из

условий, характеризующих корректно поставленные задачи. Понятие «корректная задача» было введено Ж. Адамаром в 1923 г.

С другой стороны, недостаток метода обнаружения полезного сигнала на основе корреляционного приёмника заключается в необходимости знания сигнала $s(t, x)$. Учитывая, что задача обнаружения решается в условиях априорной неопределённости, это решение может не работать.

Большинство алгоритмов, выполняющих сегментацию звукового сигнала, работают следующим образом:

1) выполняется оценка параметров фонового шума; при этом используется предположение о том, что характеристики фонового шума остаются практически постоянными в течение длительного промежутка времени;

2) для определения наличия или отсутствия сигнала в i -м сегменте, его характеристики в данном сегменте сравниваются с вычисленными характеристиками шума (используется некоторое решающее правило).

Рассмотрим предлагаемую схему информационных преобразований (рис. 30), основанную на ТАВ, лишённую недостатков корреляционного приёмника и адаптированную к специфике одномерных сигналов.

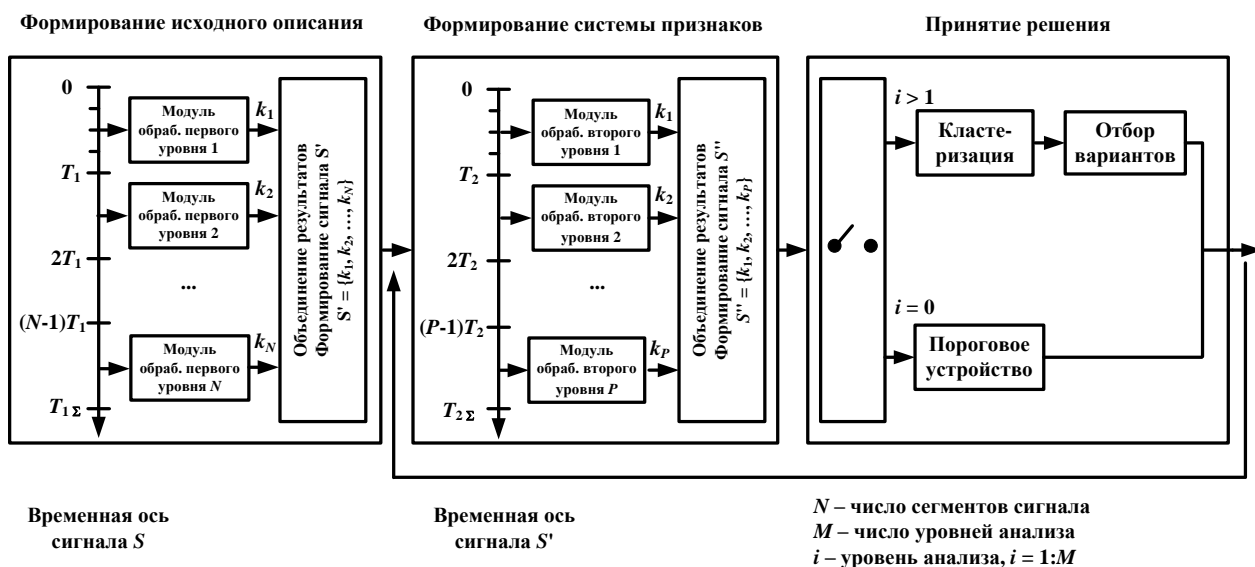


Рис. 30. Предлагаемая схема

Модули обработки, используемые на этапах формирования исходного описания и системы признаков, показаны на рис. 31 (модули выполняют обработку частей сигнала параллельно).

Этапы преобразования исходного сигнала включают в себя:

1) формирование исходного описания: вычисление U -преобразования сигнала S и построение его огибающей (сигнала S') по среднеквадратическому отклонению первого, второго и третьего коэффициентов;

2) формирование системы признаков: вычисление U -преобразования огибающей (формирование сигнала S'') – используется только нулевой коэффициент формируемого спектрального представления;

3) принятие решения: кластеризация сигнала S'' и отбор лучшего варианта сегментации.

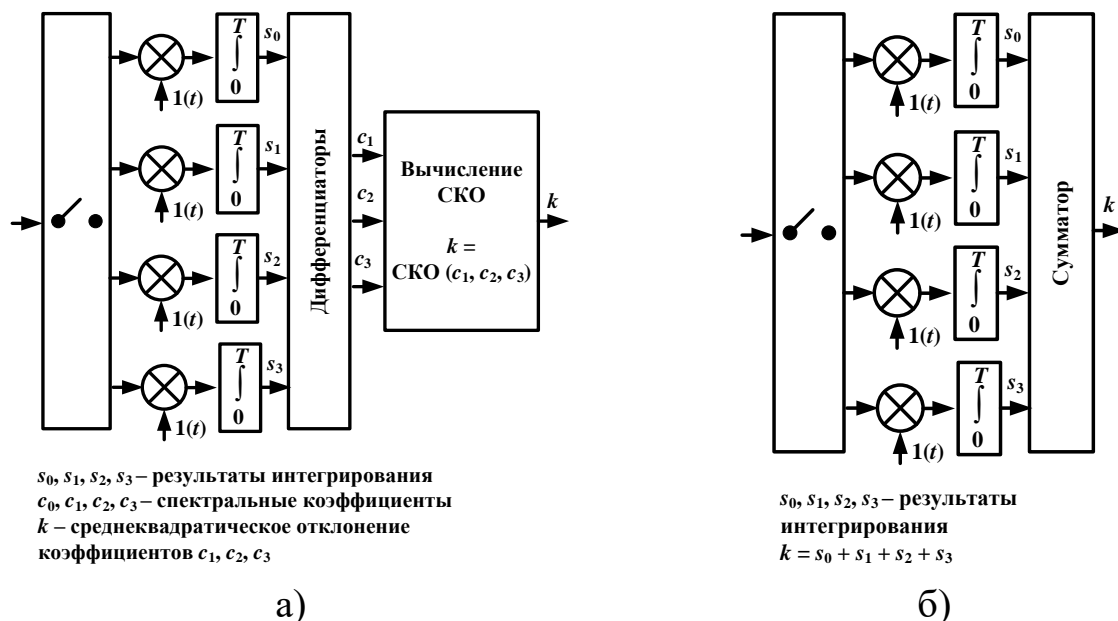


Рис. 31. Модули обработки: а – формирование исходного описания (модуль первого уровня); б – формирование системы признаков (модуль второго уровня)

Таким образом, предлагаемая схема состоит из нескольких элементов, подобных корреляционному приёмнику, однако, существенным образом от него отличающихся.

8.5. Оценка стабильности описания одномерного сигнала

Решение практически любой задачи цифровой обработки звукового сигнала начинается с создания описания сигнала. Существуют различные методы создания описания звукового сигнала. К ним относятся Фурье-преобразование, вейвлет-анализ, U -преобразование. Вопрос устойчивости описания звукового сигнала возникает при решении задачи обработки сигналов в условиях зашумлённости. Одной из таких задач является задача анализа медиаданных, которая заключается в поиске заданного звукового фрагмента в звуковом потоке или файле.

Определим показатели, которые можно использовать для выделения устойчивых сегментов сигнала A . Для этого выполним анализ сигнала A :

1) вычислим одноуровневое U -преобразование сигнала A (сигнал разбивается на M сегментов). По каждому сегменту сигнала вычислим набор полных (P_n) и замкнутых (P_s) групп. В результате, получим описание

сигнала, которое представляет собой совокупность значений откликов фильтров, а также полных и замкнутых групп, вычисленных по каждому сегменту сигнала: $D_A = \langle \bigcup_{i=1, M} F_i, P_{n,i}, P_{si}, \rangle$;

2) сформируем по сигналу A с помощью оператора искажения сигналы B и C (уровень искажений сигналов B и C должен быть одинаков, однако, учитывая, что на сигнал A накладывается случайный шум, сигналы B и C не будут эквивалентными): $B = \text{NOISE}[A]$, $C = \text{NOISE}[A]$, где $\text{NOISE}[\bullet]$ – оператор наложения шума на сигнал;

3) вычислим описания сигналов B и C : D_B, D_C . Структура описаний D_B и D_C эквивалентна структуре D_A ;

4) выполним сравнение пар описаний: D_A-D_B и D_A-D_C ; в результате, формируются массивы $C_{B, v}, C_{B, ni}, C_{B, pi}, C_{C, v}, C_{C, ni}, C_{C, pi}$ (i -ый элемент массива $C_{B, v}[i]$ – количество операторов, которые совпали по знакам для i -го сегмента описаний D_A-D_B ; $C_{B, ni}[i]$ – количество полных групп, совпавших для i -го сегмента описаний D_A-D_B ; $C_{B, pi}[i]$ – количество замкнутых групп, совпавших для i -го сегмента описаний D_A-D_B).

По результатам формирования указанных массивов можно сделать вывод об устойчивости сегментов по следующему правилу: если $C_B[i] > T$ и $C_C[i] > T$, то i -й сегмент сигнала устойчив к искажениям (T – пороговое значение).

После ряда экспериментов было установлено, что для выделения устойчивых сегментов сигнала могут использоваться только замкнутые группы. Анализируя описания искажённых и неискажённых сигналов, генерируемые с помощью замкнутых групп, удалось выделить показатели, позволяющие разделить сегменты сигнала на устойчивые и неустойчивые:

1) показатель «П1»: чем больше разница между максимальной и минимальной массами групп в наборе, описывающим сегмент сигнала S , тем более вероятно, что описание сегмента сигнала не будет меняться при его искажении:

$$D_1 = \text{MAX}(\text{MASS}(P, A)) - \text{MIN}(\text{MASS}(P, A)),$$

где P – набор групп, вычисленных по сигналу A , $\text{MAX}[\bullet]$ – оператор, вычисляющий максимальное значение среди чисел в массиве, $\text{MIN}[\bullet]$ – оператор, вычисляющий минимальное значение среди чисел в массиве, $\text{MASS}[\bullet]$ – оператор, вычисляющий массы групп в массиве P ;

2) показатель «П2»: чем больше сумма масс групп, входящих в набор групп P , описывающих сегмент сигнала A , тем больше вероятность того, что описание сегмента сигнала не будет меняться при его искажении:

$$D_2 = \text{SUM}(\text{MASS}(P, A)),$$

где $\text{SUM}[\bullet]$ – оператор вычисления суммы элементов массива.

8.6. Фильтрация одномерного сигнала

Задача очистки сигнала от шума с целью восстановления смысла сообщения, улучшения качества сигнала является одной из актуальных задач обработки сигналов.

Часто предполагается, что на исходный сигнал действует аддитивная или мультипликативная помеха.

8.6.1. Алгоритм фильтрации 1

Алгоритм фильтрации 1 основан на обнулении коэффициентов разложения и состоит из следующих шагов:

1) построение U -разложения D сигнала S (формируется один уровень разложения, длина сегмента – L , $L = W$, W – количество используемых фильтров);

2) модификация разложения D : обнуление в каждом спектре разложения указанных коэффициентов;

3) применение к модифицированному разложению D обратного преобразования: генерация сигнала S' .

Предположим, что разложение сигнала формируется с использованием четырёх фильтров, тогда существует 15 ($2^4 - 1$) комбинаций вариантов обнуления коэффициентов фильтров (табл. 12).

Таблица 12. Варианты обнуления коэффициентов

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1
2	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1
3	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0

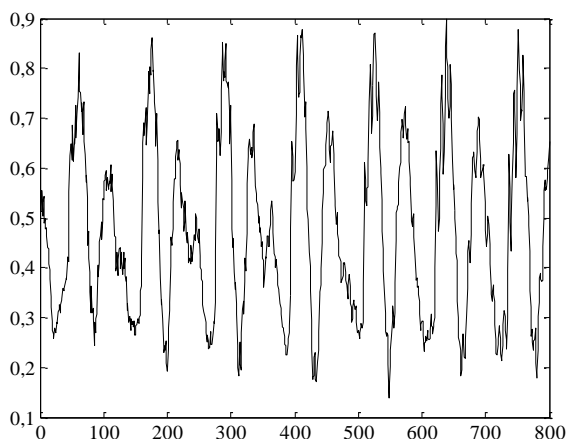


Рис. 32. Исследуемый сигнал

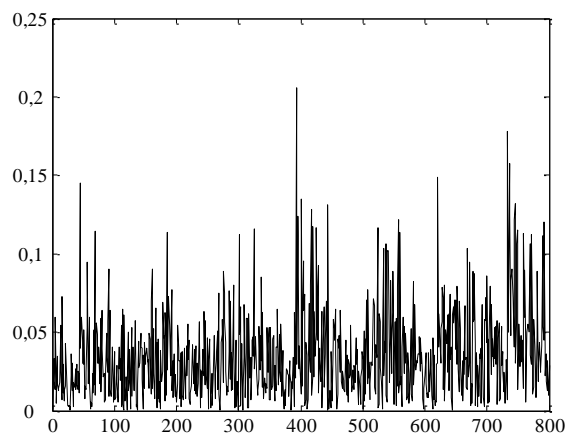


Рис. 33. Результат обнуления нулевого коэффициента

На рис. 33 показан сигнал S' (исходный сигнал показан на рис. 32), в каждом спектре которого, обнулён нулевой коэффициент. Выполнение данной операции привело к потере информации о форме сигнала (после обнуления сохраняется только высокочастотная составляющая сигнала). Таким образом, при решении задачи фильтрации обнулением коэффициентов, нельзя пользоваться комбинации с первой по восьмую (табл. 12).

При использовании 10, 11 и 12 комбинаций в получаемые сигналы вносятся артефакты (рис. 34, рис. 35, на рисунках сплошной линией показан восстановленный сигнал, пунктирной линией – исходный сигнал, отображается только часть сигнала). В данных комбинациях обнуляется первый спектральный коэффициент.

Можно отметить, что чем больше число знакоперемен (переходов через ноль) у фильтра, тем более высокочастотную составляющую он позволяет выделить из сигнала. Количество информации о форме сигнала, находящейся в коэффициентах, уменьшается с увеличением числа знакоперемен в фильтре, генерирующем коэффициент.

В связи с этим, не рекомендуется обнулять спектральные коэффициенты, полученные с применением нулевого и первого фильтров, в противном случае это приведёт не только к огрублению формы сигнала и сглаживанию шумовой составляющей, но и к внесению в сигнал различных артефактов.

Обнуление коэффициентов, полученных с применением второго и третьего фильтров, приводит только к «огрублению» сигнала, после выполнения обратного преобразования.

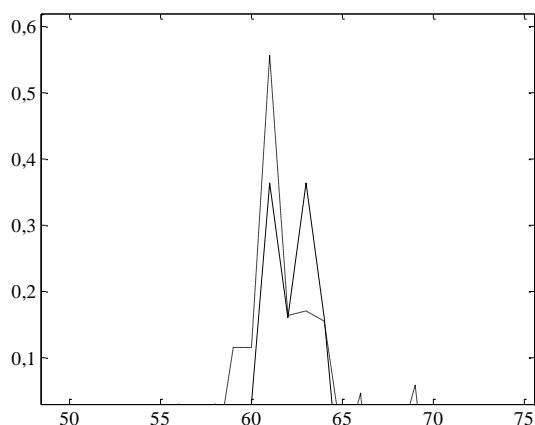


Рис. 34. Результат восстановления сигнала с использованием комбинации 1001 (10)

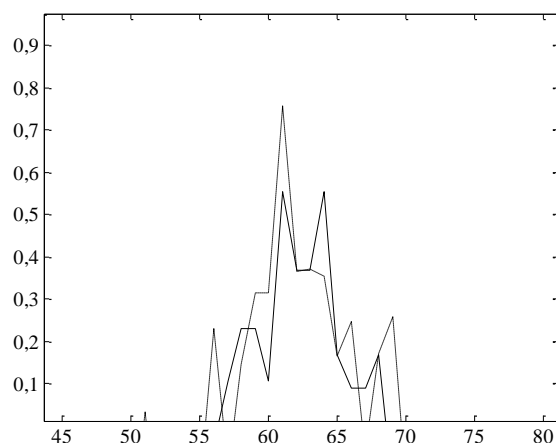


Рис. 35. Результат восстановления сигнала с использованием комбинации 1010 (11)

Сравнивая сигналы, полученные после применения 9 и 13, 14 и 15 комбинаций, можно отметить, что для 9-й комбинации генерируется более грубый сигнал, чем для других комбинаций.

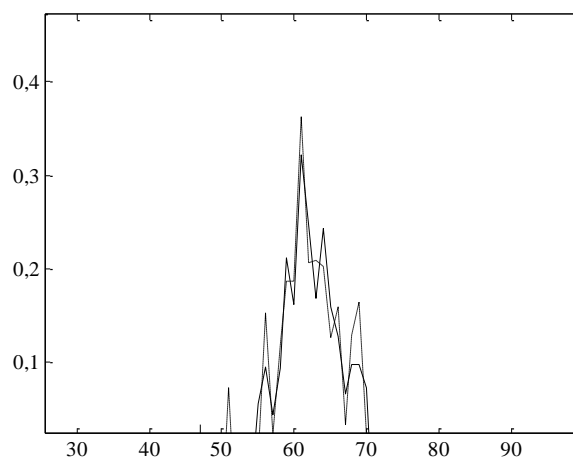


Рис. 36. Результат восстановления сигнала с использованием комбинации 1110 (15)

Таким образом, для фильтрации сигнала можно использовать следующие комбинации спектральных коэффициентов: 1000 (9), 1100 (13), 1101 (14), 1110 (15) (рис. 36).

Ранее проведён анализ результатов обнуления коэффициентов для четырёх фильтров. При использовании большего числа фильтров будет доступно больше вариантов комбинаций обнуляемых коэффициентов.

В результате выполнения алгоритма, уровень шума в сигнале S' должен быть меньше уровня шума в исходном сигнале S . Указанный результат достигается за счёт обнуления коэффициентов разложения, несущих информацию о высокочастотной составляющей сигнала, а, следовательно, и о шуме.

Примечания:

1) использование комбинаций, в которых во второй половине находятся единицы не рекомендуется, это приводит к добавлению различных артефактов к «огрублённому» сигналу, указанный эффект особенно заметен при использовании 8 и более фильтров;

2) сравнивая речевые сигналы до и после огрубления, можно отметить разницу между указанными сигналами, заключающуюся в повышении высоты голоса, полученную после огрубления сигнала;

3) результат применения комбинаций 1100 и 11110000 (8 фильтров) эквивалентен, поэтому для повышения производительности рекомендуется использовать меньшее количество фильтров.

8.6.2. Алгоритм фильтрации 2

Алгоритм фильтрации 2 основан на восстановлении сигнала с заданной степенью точности:

1) построение U -представления D сигнала S (формируется один уровень разложения, длина сегмента – L , $L = P \cdot W$, W – количество используемых фильтров);

2) выполнение обратного преобразования и генерация сигнала S' : меняя длину сегмента L , можно генерировать сигнал S' с различной степенью «огрубления». При $L = W$ сигнал восстанавливается без ошибок, при $L > W$ восстановленный сигнал будет более «грубым» (содержать меньше деталей), чем исходный;

3) интерполяция восстановленного сигнала до размера исходного.

Длина сегмента L должна быть кратна длине сигнала, в противном случае, восстановленный сигнал будет обрезан по длине. Это является следствием того, что в сигнале будут присутствовать отсчёты, не уложившиеся в целое число сегментов.

В качестве третьего алгоритма фильтрации можно предложить алгоритм, основанный на комбинации первого и второго алгоритмов. В таком алгоритме одновременно будет выполняться обнуление коэффициентов и огрубление сигнала.

Необходимо отметить, что после восстановления сигнала максимальное значение амплитуды восстановленного сигнала будет меньше амплитуды исходного сигнала. Данное примечание касается и первого алгоритма фильтрации.

8.7. Сжатие звукового сигнала

8.7.1. Алгоритм сжатия 1

Алгоритм сжатия 1 основан на обнулении коэффициентов разложения U сигнала и состоит из следующих шагов:

1) построение U -разложения D сигнала S (формируется один уровень разложения, длина сегмента – L , $L = W$, W – количество фильтров);

2) модификация разложения D : в каждом спектре разложения D обнуляются указанные коэффициенты;

3) выполнение обратного преобразования: генерация сигнала S' .

Предположим, что разложение сигнала формируется с использованием четырёх фильтров, тогда существует 15 вариантов обнуления коэффициентов фильтров (от 0000 до 1110, нумерация коэффициентов фильтров следующая: 0123).

Исследования показали, что обнуление нулевого коэффициента разложения сигнала приведёт к потере информации о форме сигнала (после обнуления сохраняется только высокочастотная составляющая сигнала).

Таким образом, для решения задачи сжатия с помощью обнуления коэффициентов, нельзя использовать комбинации, в которых обнулён нулевой коэффициент.

При использовании 10, 11 и 12 комбинаций в получаемые сигналы вносятся артефакты. Можно отметить, что чем больше число знакоперемен (переходов через ноль) у фильтра, тем более высокочастотную составляющую он позволяет выделить из сигнала. Таким образом, не рекомендуется обнулять спектральные коэффициенты, полученные с применением нулевого и первого фильтров, в противном случае это приведёт не только к огрублению формы сигнала и сглаживанию высокочастотной составляющей, но и к внесению в сигнал различных артефактов. Таким образом, при сжатии сигнала будут использованы следующие комбинации спектральных коэффициентов: 1000 (9), 1100 (13), 1101 (14), 1110 (15). Сжатие в предлагаемом алгоритме осуществляется за счёт того, что выполняется хранение только некоторых элементов спектра.

8.7.2. Алгоритм сжатия 2

Алгоритм сжатия 2 основан на огрублении сигнала:

- 1) построение U -представления D сигнала S (формируется один уровень разложения, длина сегмента – L , $L = P \cdot W$, P – целое число);
- 2) выполнение обратного преобразования и генерация сигнала S' : меняя длину сегмента L , можно генерировать сигнал S' с различной степенью «огрубления».
- 3) интерполяция сжатого сигнала до размера исходного сигнала.

При использовании данного алгоритма сжатие выполняется за счёт того, что по сегменту длиной L генерируется W коэффициентов ($W < L$).

В качестве третьего алгоритма сжатия можно предложить алгоритм, основанный на комбинации первого и второго алгоритмов. В таком алгоритме одновременно будет выполняться обнуление коэффициентов и огрубление сигнала.

8.8. Алгоритм сравнения сигналов

Одна из задач цифровой обработки сигналов – сравнение форм сигналов. Данная задача решается при оценке качества сжатия, восстановления сигнала, автоматического определения степени похожести музыкальных композиций, при разработке аудио кодека для проверки качества

кодирования звука, сравнения полученных результатов кодирования с другими кодеками и т.д.

Рассмотрим метрики, применяемые при сравнении сигналов:

1) максимальное отклонение амплитуд сигналов:

$$D = \max_{i \in [1, n]} |x(i) - y(i)|,$$

где x, y – сравниваемые сигналы; n – длина сигнала x (y). Данная метрика чувствительна к единичным отличиям в амплитудах сигналов;

2) среднеквадратичное отклонение амплитуд сигналов:

$$RMS = \sqrt{\frac{\sum_{i=1}^n (x(i) - y(i))^2}{n}};$$

3) отношение пикового уровня сигнала к шуму (peak signal-to-noise ratio, PSNR), модификация метрики RMS:

$$PSNR = 20 \cdot \log_{10} \left(\frac{A}{RMS} \right),$$

где A – максимальное возможное значение амплитуды сигнала. Один из недостатков данной метрики – высокая чувствительность к среднему отличию сигналов по амплитуде. В связи с этим PSNR используется совместно с расчетом средне интегрального значения для сравниваемых сигналов:

$$DI_x = \frac{\sum_{i=1}^n |x(i)|}{n}, \quad DI_y = \frac{\sum_{i=1}^n |y(i)|}{n}.$$

Если у сигналов совпадают средне интегральные значения, то у них совпадает и средняя амплитуда, что означает возможность применения метрики PSNR.

4) отношение сигнал / шум:

$$SNR = 10 \cdot \log_{10} \left(\frac{\sum_{i=1}^n (x(i))^2}{\sum_{i=1}^n (x(i) - y(i))^2} \right).$$

Перечисленные метрики вычисляются во временной области;

5) частотно-временная метрика основана на сравнении спектрограмм, построенных по сигналам. Для получения данной метрики входные сигналы сначала последовательно покрываются небольшими интервалами с некоторым шагом T по времени. В каждом из этих интервалов сигнала вычисляется спектр Фурье (без учета фаз частотных составляющих). Полученные спектры записываются в двумерный массив (время, частота) – спектрограмму. Амплитуды значений спектрограммы в каждой конкретной области представляются в логарифмической шкале.

Для сравнения полученных спектрограмм можно использовать следующую метрику, которая будет сравнивать сигналы, согласно данным о восприятии человеком той или иной частотной составляющей:

$$D = \frac{\sqrt{\sum_{i=1}^F \frac{1}{\alpha_i} \sum_{j=1}^n (Sx(i, j) - Sy(i, j))^2}}{F \cdot T},$$

где Sx , Sy – двумерные массивы амплитуд спектрограмм сигналов x и y , коэффициент α_i зависит от чувствительности уха в i -й частотной полосе.

Для использования данной оценки необходимо, чтобы сигналы содержали одинаковую энергию, то есть среднеквадратичное отклонение в спектрах для всего звукового сигнала должно быть минимальным. По сравнению с обыкновенной PSNR метрикой в данной мере практически решается проблема сравнения сигналов с различными амплитудами и учитывается неравномерная чувствительность уха к частотным составляющим.

Рассмотрим предлагаемый алгоритм. При оценке расстояния между сигналами не требуется формирования всего дерева разложения: достаточно построения одного уровня разложения. В этом случае результат применения U -преобразования к исходному сигналу $g(t)$ ($D = U[g(t)]$) представляет собой набор спектров, $D = \{D_i\}$, где D_i – i -ый спектр разложения. Каждый спектр состоит из L коэффициентов $D_i = \{d_0, \dots, d_{L-1}\}$, вычисленных по сегменту сигнала длиной M отсчётов, L – число фильтров Уолша системы Хармута, используемых при формировании спектра, D_{ij} – j -ый коэффициент i -го спектра.

Анализируя модельные сигналы (синусы с разными частотами) можно выделить следующие свойства U -преобразования:

1) инвариантность к растяжению/сжатию: синус частотой 33 Гц – масштабированная во времени (растянутая) версия синуса частотой 66 Гц; подобие разложений указанных синусов (одинаковые позиции максимумов в спектре, а также одинаковые относительные значения откликов), включающих один период, указывает на инвариантность описания сигнала к его растяжению / сжатию. Указанная особенность разложения может быть использована при решении задач обработки речи, в частности, она позволит указать на эквивалентность сигналов, один из которых получен при замедлении / ускорении темпа речи;

2) инвариантность к масштабированию амплитуды сигнала: если выполнить сравнение разложений исходного сигнала и его масштабированных копий, то максимальные отклики в их спектральных представлениях будут отличаться по значению (по абсолютной величине), но сохраняют своё положение вне зависимости от коэффициента масштабирования.

В табл. 13 приведены данные, подтверждающие указанные свойства U -преобразования.

Таблица 13. Разложения модельных сигналов

33 Гц, 1 период			
1: -154.32 , 1.84, 0.42	2: -0.29, 32.52 , 0.12	-2.16, -32.51 , 0.89	
3: 15.90 , 3.04, -7.64	-15.41 , 3.14, 7.41	-15.99 , -3.03, 7.68	15.32 , -3.16, -7.36
66 Гц, 1 период			
1: -78.16 , 0.42, 0.21	2: 1.09, 15.55 , -0.45	-1.30, -15.55 , 0.54	
3: 8.00 , 1.51, -3.85	-7.55 , 1.60, 3.63	-8.24 , -1.46, 3.96	7.29 , -1.64, -3.50
33 Гц, 4 периода			
1: -0.50, 0.50, 0.50	2: 0.00, -0.03, 308.61	0.00, -0.04, 308.55	
3: -154.32 , 1.84, 0.42	-154.29 , 3.54, 0.42	-154.24 , 5.23, 0.42	-154.18 , 6.92, 0.42
33 Гц, 4 периода, коэффициент масштаба амплитуды = 0.3			
1: -0.17, 0.17, 0.17	2: 0.00, -0.01, 102.87	0.00, -0.01, 102.87	
3: -51.44 , 0.61, 0.14	-51.43 , 1.18, 0.14	-51.41 , 1.74, 0.14	-51.39 , 2.31, 0.14
33 Гц, 4 периода, коэффициент масштаба амплитуды = 3			
1: -1.66, 1.66, 1.66	2: 0.02, -0.10, 1028.71	0.02, -0.15, 1028.50	
3: -514.40 , 6.14, 1.41	-514.30 , 11.78, 1.41	-514.14 , 17.42, 1.41	-513.92 , 23.06, 1.41

Сравнивая U -разложения двух сигналов, один из которых получен из другого с помощью некоторого оператора искажения, можно отметить подобие данных разложений (табл. 14). Следовательно, сравнивая указанные разложения, можно оценить степень отклонения искажённого сигнала от исходного.

Таблица 14. Разложения искажённых сигналов

Синус 33 Гц, 4 периода, искажённый гауссовым шумом (ОСШ = -4 Дб)			
1: 95.10, -1.25, -27.82	2: 13.52, 8.29, 307.87	12.36, -11.09, 315.32	
3: -149.79 , -17.68, 12.81	-158.08 , 8.52, -29.49	-161.65 , 3.92, 62.15	-149.34 , -46.52, 5.12
Синус 33 Гц, 4 периода, искажённый равномерным шумом (ОСШ = -4 Дб)			
1: 20.20, 20.33, 31.44	2: -5.56, -17.29, 328.55	-23.47, 5.40, 320.75	
3: -172.92 , -12.52, 25.05	-155.63 , -14.07, -8.40	-158.07 , 21.33, 6.26	-162.28 , 13.31, -1.18

Предлагаемый алгоритм сравнения сигналов основан на описанных свойствах разложения сигнала, полученного в результате применения U -преобразования, и состоит из следующих шагов:

1) построить спектральные разложения сигналов x и y с использованием U -преобразования для одномерных сигналов:

$$X = U[x], Y = U[y];$$

2) выполнить округление разложений X и Y .

Рассмотрим алгоритм формирования округленного представления.

Входные данные для алгоритма – спектральное разложение X и степень округления k :

1) генерация множества чисел K на отрезке $[0; 1]$ с шагом $1/k$. Например, для $k = 3$ будет получен следующий набор: $K = [0.33, 0.66, 1]$;

2) выполнить для каждого спектра разложения X следующие действия (отклик нулевого фильтра не учитывается):

а) взять максимальный по абсолютной величине элемент спектра (M);

б) умножить каждый элемент множества K на M : допустим, $M = 23$, тогда $K = [7.66, 15.33, 23]$;

в) вычислить к какому из полуинтервалов $[0; 7.66)$, $[7.66, 15.33)$, $[15.33, 23]$ (с учётом знака) нужно отнести i -й элемент спектра, в зависимости от полученного ответа рассматриваемому i -му элементу спектра присваивается одно из значений: $-3, -2, -1, 0, 1, 2, 3$.

Разложение, полученное в результате такого округления, может быть использовано при решении различных задач, например, при сравнении двух сигналов.

Ранее было указано, что при масштабировании сигнала (по времени и по амплитуде) положение максимумов в спектрах сохраняется. Разложения масштабированных версий от разложения исходного сигнала отличаются только амплитудой максимумов. Округление указанных разложений позволит доказать, что сигналы, по которым они были построены, являются масштабированными версиями одного и того же сигнала.

На рис. 37 показан пример округления спектрального разложения синуса на 4-м уровне (частота – 33 Гц, длина сигнала – 4-го периода, $M = 488$).

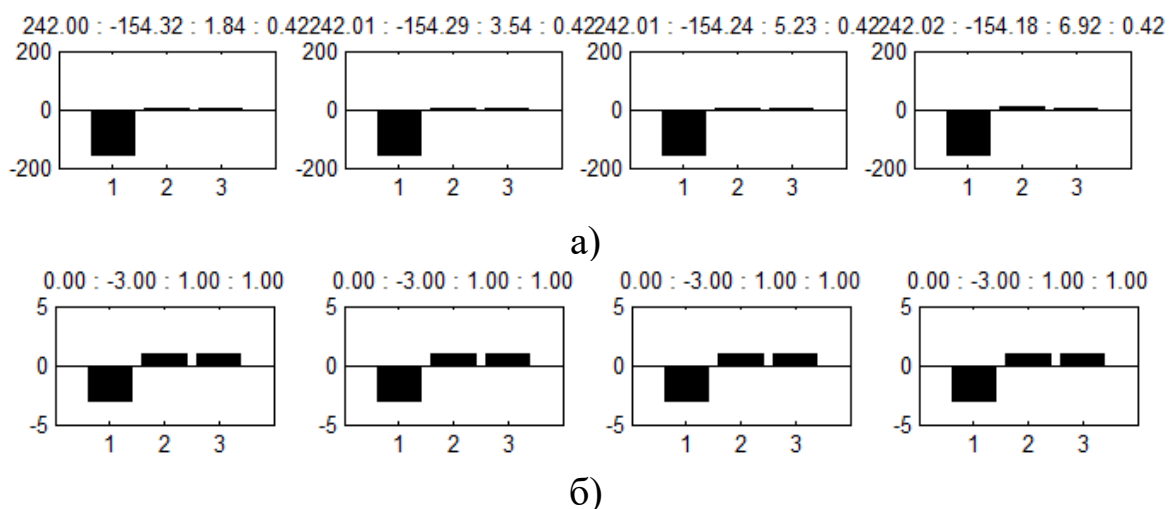


Рис. 37. Округление спектрального представления сигнала:

а – исходное разложение; *б* – округлённое разложение

3) Сравнить разложения X и Y :

$\forall i \in [1; M]$:

$\forall j \in [1; N - 1]$:

если $X_{ij} = Y_{ij}$

$$E = E + k \cdot 2$$

если $S_{D,i}(j) \neq Y_{ij}$

$$E = E - |X_{ij} - Y_{ij}|.$$

Нормируем полученное значение E к отрезку $[0; 1]$:

$$E = E / (M \cdot k \cdot N \cdot 2).$$

Примечание:

1) максимальное расстояние между элементами спектров составляет $k \cdot 2$;

2) если спектральные коэффициенты X_{ij} и Y_{ij} эквивалентны, то к оценке E прибавляется $k \cdot 2$;

3) чем больше отличаются спектральные коэффициенты X_{ij} и Y_{ij} , тем большее значение вычитается из E .

После окончания работы алгоритма значение E будет содержать величину, характеризующую расстояние между сигналами x и y .

Разработанный алгоритм инвариантен к изменению частоты и амплитуды сигналов. При сравнении двух сигналов одинаковой формы, но разной частоты и амплитуды предлагаемый алгоритм укажет на их эквивалентность.

8.9. Поиск сигналов в звуковых базах данных

Быстрорастущий объём коллекций цифровых аудио материалов создаёт потребность в эффективных механизмах поиска по содержимому в данных коллекциях. Традиционный подход к поиску основан на использовании текстовых метаданных, описывающих содержимое звуковых файлов. Часто такие метаданные не доступны или не содержат достаточное для поиска количество нужной информации. В этом случае используется стратегия, основанная на анализе содержимого звуковых файлов. Рассмотрим стратегию поиска, использующая метод поиска по образцу: при задании аудио-запроса задача состоит в том, чтобы извлечь из музыкальной коллекции все сигналы, которые похожи на запрос.

Для поставленной задачи известны различные решения. Например, Philips для описания фрагментов звукового сигнала предлагает компактное, 32-битное, представление разностей энергией поддиапазонов вместе с поиском точного совпадения в хэш-таблице. Поиск в системе Shazam основан на использовании множества ключевых сигнатур, представляющих пары пиков в спектрограммах Фурье. Компанией Fraunhofer предложена технология AudioID. Данная технология построена на классическом фреймворке классификации паттернов, использующем стандартный метод ближайшего

соседа на дескрипторах MPEG-7, закодированных с помощью векторного квантования. Академические исследования включают в себя улучшения метода Shazam с использованием Constant-Q преобразования, методов компьютерного зрения на спектрограммах и исследование стратегии поиска для бинарных звуковых отпечатков.

Концептуально систему поиска звукового сигнала в базе данных можно представить в виде схемы, показанной на рис. 38.



Рис. 38. Концептуальная модель системы

Рассмотрим описание первых трёх модулей, отвечающих за декодирование и нормирование, генерацию отпечатков сигнала и поиск в базе данных. В качестве базы данных предполагается использование реляционных СУБД общего назначения, поскольку их технологическое развитие обеспечивает производительность и масштабируемость.

8.9.1. Предварительная обработка и признаковое описание сигнала

Этап предварительной обработки сигнала s включает его ресемплинг (с целью приведения всех сигналов в базе данных к одинаковой частоте дискретизации) и нормирование по приведённой ниже формуле к диапазону $[0, 1]$:

$$s_{dsc} = s_{dsc} - \min(s_{dsc}), \quad s_{dsc} = s_{dsc} / \max(s_{dsc}).$$

Необходимость нормирования объясняется требованиями U -преобразования к динамическому диапазону сигнала.

Анализируемый сигнал представляется в виде множества сегментов s , где под сегментом понимается некоторая ненулевая последовательность значений из исходного сигнала S : $s = \{s_k\}$, $k = 1:K$, где K – число сегментов, длина сегмента составляет M отсчётов, сегменты выбираются из сигнала с равномерным шагом в N отсчётов.

Признаковое описание сегмента сигнала формируется в результате применения к i -му сегменту U -преобразования. Первый этап U -преобразования – деление i -го сегмента на M равных по количеству отсчётов частей и применение к каждой из частей операции суммирования (данный этап называется Q -преобразованием):

$$q_k[i] = \sum_{l \in T_i} s_k[l],$$

где T_i – индексы отсчётов сегмента s_k , принадлежащие i -й части сегмента.

Второй этап – применение к полученному вектору частичных сумм q_k фильтров Уолша системы Хармута, в результате чего формируется вектор спектральных коэффициентов μ_k :

$$\mu_k[i] = \sum_{j=0}^{M-1} q_k[j] F_i[j], i=1, M, k=1, K,$$

где M – число используемых фильтров Уолша системы Хармута.

8.9.2. Информационная модель системы хранения и поиска

Эффективность разрабатываемой системы зависит от способа хранения и поиска фрагмента звукового сигнала. В реальных условиях поиск может происходить среди миллионов записей. Для решения данной задачи разработана двухуровневая модель хранения и поиска отпечатков сигнала, представленного в виде множества сегментов. На этапе сохранения на верхнем уровне (рис. 39) сегменты анализируемого сигнала, с помощью метода рекурсивного спуска, размещаются в терминальных узлах двоичного дерева.

Это позволяет заложить в дизайн системы верхнеуровневый параллелизм, например, деление по различным серверам поиска. Затем, по каждому сегменту формируется ключ, который используется при поиске сигнала, но не по всей базе, а только среди сегментов, относящихся к данному терминальному узлу.

8.9.3. Размещение сегментов в двоичном дереве

Размещение элемента (сегмента сигнала) в дереве происходит согласно переходам при рекурсивном спуске по сегменту:

1) инициализируется N -уровневое двоичное дерево, в качестве текущего уровня анализа устанавливается первый: $i = 1$ ($i \leq N$);

2) вычисляется U -преобразование сегмента сигнала s_k , результатом вычисления является вектор μ_k из M значений (в общем виде – M значений, в данной работе M принято равным 16), к которым применяется функция сигнум:

$$\mu'_k = \text{sgn}(\mu_k);$$

3) вычисляется расстояние Хэмминга между μ'_k и бинарными векторами: $T_1 = \{1, 1, 1, 1, 1, 1, 1, 1, -1, -1, -1, -1, -1, -1, -1, -1\}$ и $T_2 = \{-1, -1, -1, -1, -1, -1, -1, -1, 1, 1, 1, 1, 1, 1, 1, 1\}$:

$$d_1 = DIST [\mu'_k, T_1], d_2 = DIST [\mu'_k, T_2],$$

где $DIST []$ – оператор вычисления расстояния Хэмминга. Если $d_1 < d_2$ для дальнейшего анализа выбирается левая половина анализируемого сегмента, а информация о сегменте помещается в левую дочернюю вершину (относительной текущей вершины), в противном случае – для анализа выбирается правая половина сегмента, а информация о сегменте помещается в правую дочернюю вершину.

4) выполняется переход на следующий уровень анализа сегмента s_k : в качестве сегмента s_k теперь рассматривается его левая или правая часть (в зависимости от решения, принятого на предыдущем шаге алгоритма), текущий уровень дерева устанавливается в $i = i + 1$, выполняется переход на шаг 2 алгоритма. В случае если $i > N$, сегмент считается обработанным. Работа алгоритма выполняется до тех пор, пока информация о всех сегментах исходного сигнала не будет помещена в двоичное дерево.

Таким образом, в результате работы алгоритма, информация о сегменте сигнала помещается в один из терминальных узлов дерева.

8.9.4. Формирование ключа по сегменту сигнала

На нижнем уровне выполняется формирование ключа по сегменту s_k , который затем будет использоваться при поиске сигналов в базе данных (в рамках заданного терминального узла). Ключ K_k для сегмента s_k предлагается формировать по самому сегменту s_k , а также по его соседям:

$$K_k = KEY [s_k, s_{k+1}, s_{k+2}, \dots, s_{k+L}],$$

где $KEY []$ – оператор формирования ключа, L – количество соседей, участвующих в формировании ключа. Таким образом, на данном уровне обработка сегментов выполняется с учётом их временной связи друг с другом. Это позволяет значительно повысить точность поиска звуковых сигналов в базе данных.

Для генерации ключа объединяются вектора значений функции сигнум после первого U -преобразования сегмента каждого из сегментов, входящих в ключ. Далее полученный вектор переводится в набор битов, где, к примеру, положительный знак преобразуется в ноль, а отрицательный становится единицей.

Использование при формировании ключа последовательности сегментов, даже состоящей всего из двух элементов (например, s_k и s_{k+1}), по сравнению с одинарным сегментом, позволит решить задачу быстрого отсечения одинарных сегментов, одинаковых для двух разных сигналов, хранящихся в базе данных.

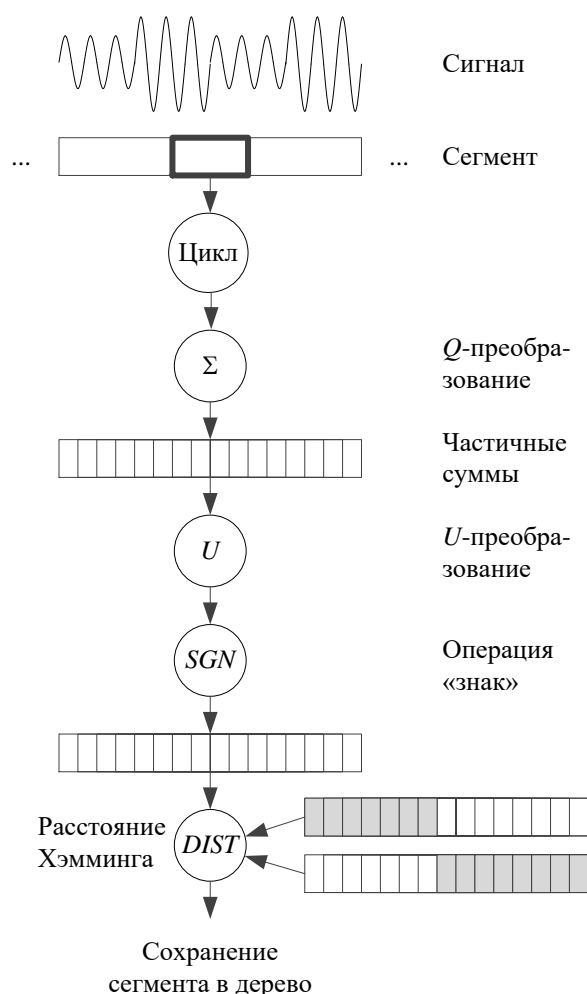


Рис. 39. Генерация отпечатка сегмента

Другими словами, полагается, что вероятность возникновения коротких одинаковых последовательностей сегментов между двумя сигналами, хранящимися в базе данных, меньше, чем вероятность принадлежности одного сегмента двум разным сигналам.

Таким образом получается последовательность, которую можно интерпретировать как беззнаковое целое число. Это число и будет использоваться как ключ при поиске.

8.9.5. Анализ возможных коллизий ключей поиска

Обоснование использования второго уровня анализа сигналов следующее. Проведём небольшой мысленный эксперимент. Пусть данные имеют достаточную энтропию и все случайные события равновероятны.

Пусть имеется 10^9 звуковых сегментов, которые необходимо разместить в базе данных. Данное число может показаться большим, но стоит упомянуть что размер такого сегмента составляет доли секунды.

Пусть глубина разбиения равняется шести. Тогда на верхнем уровне будет построено дерево глубиной в шесть уровней. Следовательно, с

помощью данного дерева можно разделить пространство данных на 2^6 частей, в каждой из которых окажется в среднем около 16 миллионов сегментов.

Предположим, что для сегмента сгенерирован ключ на основе метода, описанного выше, причём для формирования ключа использовался только сам сегмент, без соседей ($L = 0$). Диапазон значений ключа – от 0 до 32767, так как размерность двоичного вектора после (3) составляет 15 элементов. Пусть каждому ключу на втором уровне соответствует порядка 240 сегментов. Число кажется небольшим, но поскольку выполняется поиск последовательности сегментов, необходима выборка каждого следующего для всех 240 элементов сегмента, что увеличит количество операций ввода-вывода и негативно повлияет на производительность.

Если же ключ будет рассчитываться хотя бы для двух последовательных сегментов, то его диапазон значений будет от нуля до $2^{30}-1$. Тогда коэффициент заполнения будет много меньше единицы.

Несмотря на то, что в мысленном эксперименте сделаны некоторые допущения, в реальности данные будут иметь куда более меньшую энтропию, поэтому дисбаланс в количестве записей в пару порядков не создаст большого количества коллизий в отдельных ключах.

8.9.6. Поиск звукового сигнала в базе данных

Алгоритм поиска произвольного сигнала включает те же шаги, что и алгоритм формирования ключа.

Сначала по входному сигналу формируется множество сегментов с заданным шагом сдвига. Затем для каждого сегмента определяется терминальный узел двоичного дерева, к которому он относится и по сегменту формируется ключ. После формирования ключа выполняется его поиск в базе звуковых сигналов. В результате формируется ответ: найден или не найден сигнал в базе данных.

При поиске произвольного сигнала в базе данных одним из основных требований является его длительность. Минимальная длительность сигнала должна быть на один сегмент больше, чем количество сегментов, участвующих на нижнем уровне в формировании ключа для поиска. Это требование является следствием механизма поиска, который двигаясь с первого отсчета входного сигнала с некоторым шагом, ищет возможную последовательность сегментов. Этот процесс может показаться вычислительно сложным, но при хорошем отсечении во время поиска процесс должен проходить достаточно быстро.

8.10. Прогнозирование времени отказа подшипника качения

Механизмы с течением времени изнашиваются и, соответственно, возрастает риск их поломки. Это создает необходимость в своевременном обслуживании изношенных деталей механизмов. Одним из самых важных элементов практически любого механизма является подшипник качения. Поэтому подшипник может рассматриваться как критический узел, так как выход его из строя снижает безопасность и работоспособность машины. Известны различные стратегии планирования мер обслуживания и ремонта технического оборудования. Наиболее эффективная стратегия – стратегия обслуживания оборудования по состоянию. Данная стратегия позволяет оптимизировать обслуживание за счёт предсказания возникновения поломок.

Методы прогнозирования остаточного ресурса устройства можно разделить на три группы: на основе физической модели устройства, на основе данных мониторинга устройства и гибридные.

Методы прогнозирования на основе физической модели, обычно представляются в виде набора алгебраических или дифференциальных уравнений и описывают поведение исследуемой физической системы во времени, включая её разрушение. Преимущество этого подхода в том, что он обеспечивает точные результаты, недостаток – в сложности создания модели вследствие нелинейности процессов, протекающих в устройствах.

Методы прогнозирования на основе данных выполняют преобразование данных мониторинга устройства в информацию, используемую в дальнейшем для прогнозирования. Модель разрушения устройства создаётся только на основе данных, предоставляемых системой мониторинга без использования аналитической модели системы или её физических параметров. Формирование прогноза выполняется на основе методов искусственного интеллекта или статистических методов. Оценка остаточного ресурса на основе данных выполняется в два этапа: создание модели поведения устройства во времени и использование данной модели для предсказания остаточного ресурса устройства.

Гибридный метод прогнозирования объединяет подходы на основе физической модели и данных, и, в результате, обладает достоинствами и недостатками обоих подходов.

В настоящее время наиболее активно развиваются методы прогнозирования на основе данных мониторинга. Например, известны конкурсы, посвящённые разработке указанных методов прогнозирования.

Основные трудности при разработке методов прогнозирования на основе данных связаны с наличием искажений в показаниях датчиков и со сложностью отбора признаков обладающих свойством монотонности и

выделения тренда. Данные свойства необходимы для обеспечения точной оценки остаточного ресурса устройства.

Известно несколько подходов к решению поставленной проблемы. В первом подходе анализ вибросигнала осуществляется при помощи метода главных компонент, использующегося для уменьшения размерности исходных вибрационных признаков, путем удаления избыточных признаков. Такой метод позволяет уменьшить конечный объем вычислений, однако такой способ не реализует полный потенциал имеющихся данных. Для вычисления оставшегося времени работы подшипника используется метод описания данных с помощью опорных векторов, основанный на принципах минимизации структурных рисков. Он используется для соотнесения данных обучения в область гиперсферы, таким образом граница полученного радиуса гиперсферы может использоваться в качестве индикатора состояния подшипника. Такой способ показывает хорошие результаты в данном виде вычислений и было решено использовать его на одном из этапов.

Известен подход, в котором анализ вибросигнала осуществляется при помощи метода Байесовского моделирования Монте-Карло, а этап вычисления оставшегося времени работы подшипника реализован с помощью метода скользящего среднего спектрального эксцесса и метод регрессии опорных векторов.

Информационная модель предсказания состоит из части, ответственной за анализ вибросигнала и части вычисления оставшегося времени работы подшипника.

Этап классификации предлагается реализовать с помощью метода машинного обучения One-class SVM, который хорошо себя показывает при поиске отклонений в задачах с большим количеством стандартных данных. Накопление отклонений регистрируется и при превышении определенного значения отображает границу начала т. н. зоны аномальных данных. Это позволяет сократить количество используемых измерений на этапе прогнозирования, беря в расчет только данные из аномальной зоны.

Далее по полученному на предыдущем этапе массиву аномальных данных рассчитывается кумулятивное среднеквадратичное отклонение. На полученных данных производится аппроксимация конечной части кривой с помощью параболы, начало аппроксимации рассчитывается с помощью базисной функции. Один из коэффициентов полученной параболы, в результате сигнализирует о состоянии подшипника (скорости его износа за измеренный отрезок времени). На основе этого производится вычисление оставшегося времени работы подшипника (рис. 40).

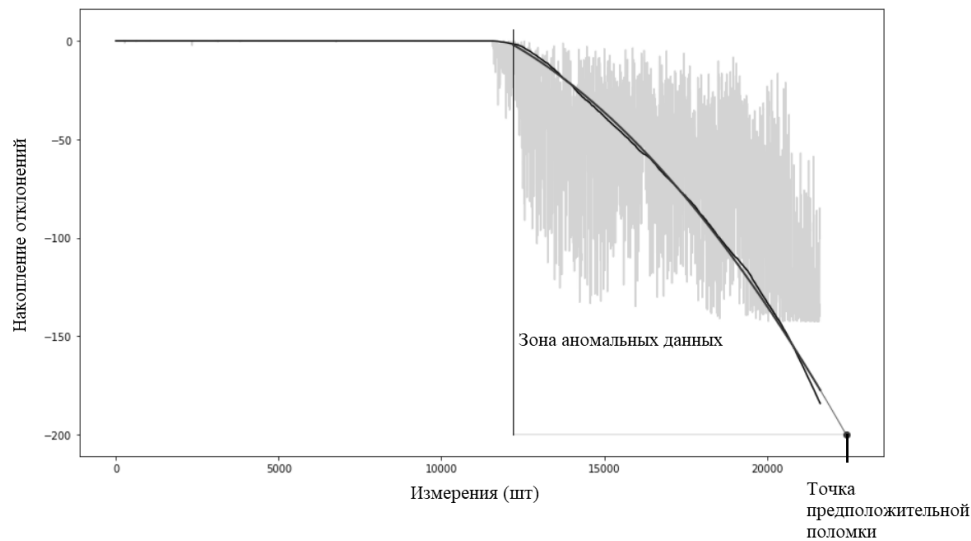


Рис. 40. Вычисление оставшегося времени работы подшипника

Установка для записи сигналов от подшипников показана на рис. 41.

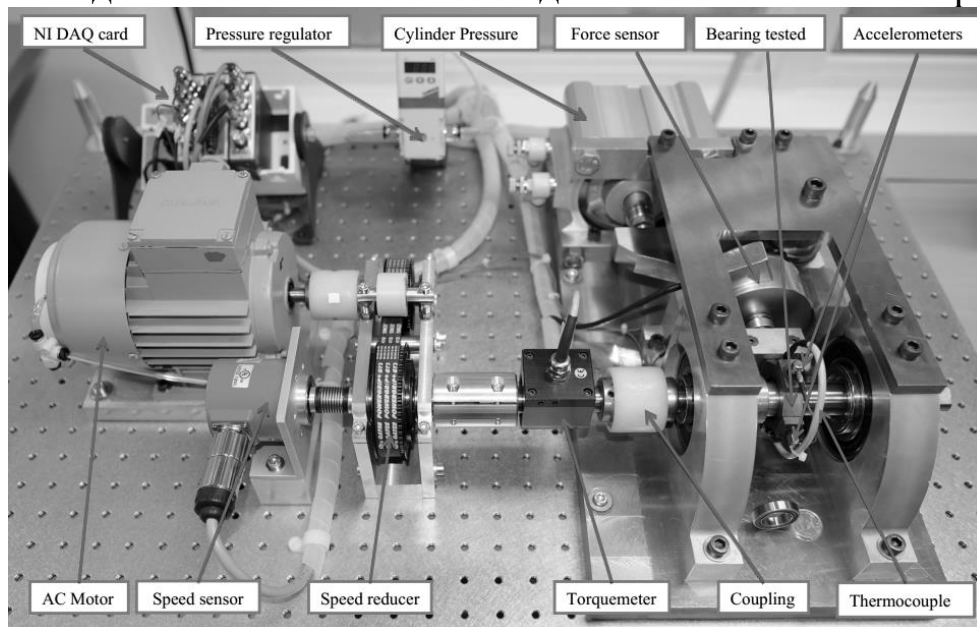


Рис. 41. Установка для тестирования подшипников

Контрольные вопросы

1. Опишите теоретическую основу метода выбора длины сегмента сигнала.
2. Подход к выбору ключевых сегментов сигнала.
3. Метод оценки частоты основного тона голоса.
4. Метод огрубления одномерного сигнала.
5. Известные подходы к оценке отношения сигнал / шум.
6. Метод оценки отношения сигнал / шум.

7. Свойства полезного сигнала и шума.
8. Опишите классическую задачу выделения информационного сообщения из принятого сигнала.
9. Опишите затруднения, связанные с использованием корреляционного приёмника при выделении информационного сообщения из принятого сигнала.
10. Поход к выделению активности голоса с позиций теории активного восприятия.
11. Метод оценки стабильности описания одномерного сигнала.
12. Подходы к фильтрации одномерного сигнала.
13. Подходы к сжатию одномерного сигнала.
14. Опишите метод сравнения одномерных сигналов.
15. Метод прогнозирования состояния подшипника качения.

9. РЕШЕНИЕ ЗАДАЧ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Решение задач распознавания изображений основано на использовании двумерных фильтров Уолша системы Хармута. Правила вычисления полных и замкнутых групп, а также замкнутых множеств для двумерного случая аналогичны правилам для одномерного случая. Один из вариантов набора фильтров (16 фильтров) показан на рис. 42. Серый цвет обозначает «-1», белый – «+1».

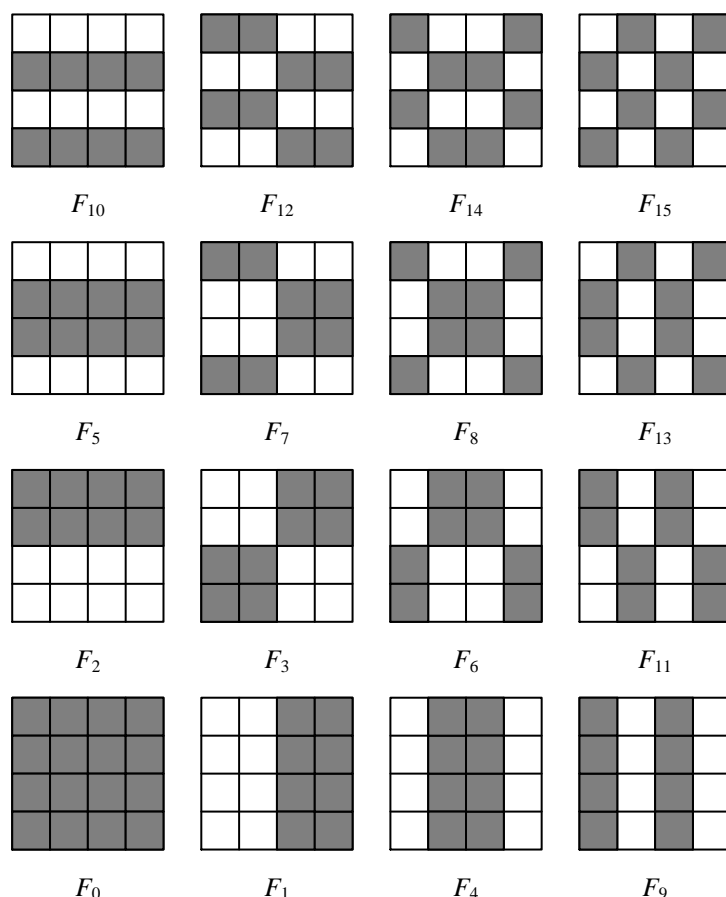


Рис. 42. Двумерные фильтры Уолша системы Хармута

Исходными данными при решении задачи распознавание является цифровое изображение I размером $N \times M$ отсчетов.

9.1. Обнаружение объектов на изображении с использованием локального признакового описания

Предварительная обработка изображения при решении задачи обнаружения объектов заключается в обработке его скользящим окном размера $n \times n$ с определенным шагом по вертикали и по горизонтали. В результате изображение I представляется в виде множества областей:

$$\mathbf{P}_I = \{P_i, i = \overline{1, Q}\},$$

где P_i – область изображения, Q – количество областей на изображении I .

Алгоритм формирования множества областей \mathbf{P}_I выглядит следующим образом:

$$\forall x = \overline{1, (N - n)}$$

$$\forall y = \overline{1, (M - n)}$$

$$P_i = I[x + n, y + n]$$

$$\mathbf{P}_I = \mathbf{P}_I \cup P_i$$

$$i = i + 1$$

$$y = y + s_h$$

$$x = x + s_w$$

где $N \times M$ – размер изображения; $n \times n$ – размер скользящего окна; s_h – шаг по вертикали; s_w – шаг по горизонтали; I – исследуемое изображение. В условиях априорной неопределенности шаг по горизонтали и шаг по вертикали принимаются равными $s_h = s_w$.

Процесс отбора областей, содержащих ключевые точки, начинается с вычисления для каждой области изображения среднеквадратического отклонения (СКО) спектральных коэффициентов её признаков описания:

$$s_i = \sigma[D_i] = \sqrt{\frac{1}{L-1} \sum_{k=1}^{L-1} (d_k - \bar{d})^2},$$

где σ – оператор, вычисляющий СКО; L – количество коэффициентов спектрального представления ($L = 16$); d_k – k -й спектральный коэффициент; \bar{d} – среднее арифметическое спектральных коэффициентов рассматриваемой области без учета нулевого коэффициента.

Предлагается следующий критерий, который позволяет определить, содержит ли область изображения P_i ключевую точку искомого объекта, и сформировать множество описаний ключевых точек \mathbf{K} :

$$1. \mathbf{S} = \{S_i, i = \overline{1, Q}\};$$

$$2. \mathbf{K} = \emptyset;$$

$$3. \forall i, i = \overline{1, Q}$$

$$\text{если } s_i > (k \cdot \max(\mathbf{S})) \rightarrow \mathbf{K} = \mathbf{K} \cup D_i$$

где k – пороговая величина.

Таким образом, считается, что область изображения содержит ключевую точку объекта, если СКО спектральных коэффициентов этой области больше, чем максимальное СКО всех областей изображения, умноженное на порог k . На рис. 43 приведены примеры выявления областей с

ключевыми точками объекта при разных значениях порога k (N – количество обнаруженных ключевых точек).

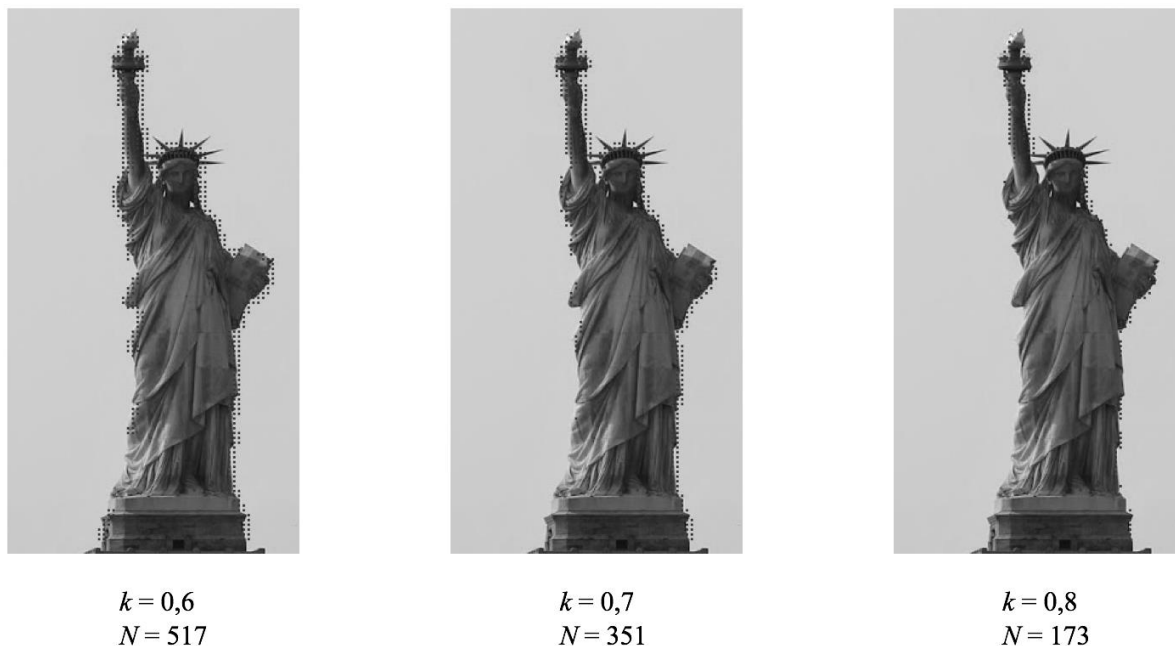


Рис. 43. Примеры выявления областей с ключевыми точками

Формирование признакового описания объекта начинается с построения признакового описания всех областей изображения этого объекта. В качестве описания каждой области используются коэффициенты спектрального представления изображения, получаемые в результате применения к данной области U -преобразования:

$$D_i = U[P_i] = \{d_k, k = \overline{1, L}\},$$

где U – оператор, выполняющий U -преобразование, P_i – область изображения, для которой формируется описание, d_k – k -ый коэффициент спектрального представления области изображения; L – количество фильтров, применяемых при вычислении U -преобразования ($L = 16$).

Исходя из этого, признаковое описание изображения объекта в целом представляется в виде набора признаковых описаний всех его областей:

$$\mathbf{D}_1 = \{D_i, i = \overline{1, Q}\}.$$

Следующим шагом формирования признакового описания объекта является выделение его ключевых точек. Точка объекта считается ключевой, если в области, в которой находится эта точка, наблюдается резкий перепад яркости. Перепады яркости соответствуют контурам объектов.

Этап принятия решения заключается в сравнении множества описаний точек исследуемого изображения \mathbf{D}_H и множества описаний областей с ключевыми точками объекта \mathbf{K} и выявлении в \mathbf{D}_H областей, эквивалентных

по своему признаковому описанию областям из \mathbf{K} . Далее из таких точек формируется множество \mathbf{E} .

Пусть две точки P_1 и P_2 имеют признаковые описания D_1 и D_2 . Тогда критерий эквивалентности точек P_1 и P_2 имеет следующий вид:

$$P_1 = P_2, \text{ если } \text{sgn}(D_{1k}) = \text{sgn}(D_{2k}), k = \overline{1, L},$$

где $\text{sgn}(x)$ – оператор вычисления знака выражения x ; D_{1k}, D_{2k} – k -й элемент вектора спектральных коэффициентов точек P_1 и P_2 ; L – количество спектральных коэффициентов в признаковых описаниях точек.

Таким образом, две точки изображения считаются эквивалентными тогда и только тогда, когда соответствующие спектральные коэффициенты их признаковых описаний имеют одинаковые знаки. Тогда алгоритм формирования множества \mathbf{E} имеет следующий вид:

$$1. \mathbf{E} = \emptyset;$$

$$\forall i, i = \overline{1, Q}$$

$$2. \quad \forall j, j = \overline{1, L}$$

$$\text{если } D_{H,i} = K_j \rightarrow \mathbf{E} = \mathbf{E} \cup D_{H,i}.$$

После обнаружения на изображении точек, эквивалентных ключевым точкам объекта, необходимо принять решение о локализации данного объекта в месте наибольшего скопления найденных точек. Предлагается метод принятия такого решения на основе разбиения точек на кластеры по координатному принципу. В таком случае делается предположение, что наибольший по количеству точек кластер располагается на изображении в то же месте, где и искомый объект.

Известны различные методы, применяемые для решения задачи кластеризации объектов. Все они могут давать различные результаты на одних и тех же входных данных, однако сравнить качество работы методов сложно, так как не существует однозначно наилучшего критерия оценки качества кластеризации. Для решения задачи кластеризации предлагается использовать метод на основе алгоритма кластеризации K-Means и метод кластеризации Mean Shift.

9.2. Поиск похожих людей по фотографии лица, идентификация человека по фотографии лица

Задача распознавания лиц относится к числу приоритетных исследований в области систем компьютерного зрения. Она включает на два этапа: выделение лица человека в естественной или искусственной обстановке (детектирование) и последующее установление степени сходства с лицами из известной системе распознавания базы (идентификация). Актуальность этой задачи, а также её предпочтительность по сравнению с другими

средствами идентификации личности (идентификация по отпечаткам пальцев или по сетчатке глаза) заключается в том, что нет необходимости непосредственного контакта системы и человека. Лицо человека само по себе содержит уникальную информацию для проведения безошибочной идентификации. Кроме того, задача распознавания актуальна в системах безопасности:

- 1) сравнение фотографий на документах, удостоверяющих личность;
- 2) контроль доступа к безопасным компьютерным сетям;
- 3) верификация пользователя при проведении транзакций;
- 4) наблюдение в публичных местах для предотвращения террористических актов и др.

Несмотря на известность задачи идентификации человека по изображению лица, нахождением похожих лиц по фотографии занимается лишь относительно небольшое число разработчиков, хотя в индустрии развлечений существует большая потребность и актуальность нахождения группы похожих лиц из больших баз данных фотографий. Так, большое распространение находят сервисы, занимающиеся поиском похожих лиц в социальных сетях. Нахождение похожих лиц является актуальным в системе безопасности. На основании накопленных баз данных лиц, попавших в поле зрения камер, производится поиск похожих лиц по этой базе. Система поиска похожих лиц даёт возможность значительно ускорить процесс поиска конкретного человека в видеоархиве и получить статистику его появлений в поле зрения различных камер.

Область информационного поиска не ограничивается задачей кластеризации фотографий людей, с помощью данной программной системы так же можно выполнять поиск похожих объектов. Система кластеризации фотографий должна максимально быстро и точно подбирать совокупность фотографий, похожих на заданную. Но стоит учесть, что степень похожести людей для каждого человека разная и точность решения данной задачи оценивается экспертом.

Этап предварительной обработки эквивалентен для этапов обучения и применения методов поиска похожих людей по фотографии лица и идентификации человека по фотографии лица.

Предварительная обработка заключается в следующем. Сначала необходимо для каждого изображения задать положение глаз на фотографии. Исходя из предоставленных из координат глаз, изображение поворачивается, масштабируется и обрезается. В результате данных преобразований входное изображение приобретает формат, понижающий воздействие фоновых цветовых и яркостных составляющих входного изображения на процесс распознавания.

Таким образом, глаза человека на входном изображении располагаются на одной горизонтальной прямой, а расстояние между ними равно $1/3$ ширины получившегося изображения. Само изображение приводится к размеру $N \times M$ пикселей. Исходя из того, что цветовые характеристики изображения не оказывают значительного влияния на результат по сравнению с яркостными, для упрощения дальнейших расчётов фотография преобразуется в градации серого. В процессе данного преобразования происходит окрашивание целевого пикселя на основе цветов пикселей исходного изображения. Такой порядок окрашивания выбран для того, чтобы избежать возникновения пробельных линий на целевом изображении.

Затем изображение представляется в виде Q блоков одинаковых размеров $h_L \times w_L$:

$$\mathbf{P}_I = \{P_i, i = \overline{1, Q}\},$$

после чего к каждому из блоков (P_i) применяется Q -преобразование:

$$\mathbf{P}_{QI} = \{P_{Q,i}, i = \overline{1, Q}\}, P_{Q,i} = Q[P_i],$$

где $Q[\bullet]$ – оператор вычисления Q -преобразования.

Этап вычисления признаков заключается в применении к результату Q -преобразования множества фильтров, показанных на рис. 38, т. е. вычислению U -преобразования:

$$\mathbf{P}_{UI} = \{P_{U,i}, i = \overline{1, Q}\}, P_{U,i} = U[P_{Q,i}],$$

где $U[\bullet]$ – оператор вычисления U -преобразования.

Алгоритм принятия решения реализуется только для этапа применения метода поиска похожих людей по изображению лица. Алгоритм принятия решения заключается в кластеризации базы данных изображений. В качестве реализации алгоритма кластеризации, например, можно использовать метод k -средних.

На этапе кластеризации происходит отбор группы фотографий по признакам, которыми являются элементы вектора спектральных коэффициентов, полученные в результате U -преобразования. Сначала сравниваются знаки значений векторов признаков, вычисленных по анализируемой фотографии и фотографиям, хранящимся в базе данных (обучающей выборке). При рассмотрении отбрасываются значения, которые по абсолютному значению менее, чем 20% от максимального значения элемента в текущем векторе спектральных коэффициентов. Эта операция производится для уменьшения влияния нестабильных (маленьких) значений на результат. Далее рассчитывается процент совпадений и в результат включаются те фотографии, чей процент оказался больше выбранного порогового значения.

Для задачи идентификация человека по изображению лица этап предварительной обработки и вычисления признаков описания эквивалентен указанным этапам для задачи поиска похожих людей по фотографии лица.

Для идентификации человека по изображению лица необходимо провести классификацию признаков описаний. Для каждого человека в базе данных присутствует несколько фотографий формирующие признаковое описание одного человека. С помощью классификатора, все признаки классифицируются, таким образом, признаки описания исследуемого изображения будут отнесены к одному из классов. На основе этого будет получен результат. Рассмотрим алгоритм по принятию решения об идентификации человека.

9.3. Обнаружение вредоносного воздействия на систему видеоконтроля

Задача детектирования вредоносного воздействия на систему видеоконтроля является одной из важных задач в области информационной безопасности на сегодняшний день. В современном мире существует большое количество различных мест скопления людей, таких как торговые центры, аэропорты, общественный транспорт, которые оснащены камерами видеонаблюдения. Одновременное наблюдение за изображениями со всех камер в течение длительного времени является достаточно сложной задачей, даже для опытного персонала службы безопасности. Поэтому, очевидно, что введение автоматизации в данный процесс может значительно улучшить эффективность системы наблюдения. Отсюда можно сделать вывод о том, что разработка метода детектирования негативного воздействия на камеру является достаточно актуальной.

Различные исследования проводились в отношении выявления вредоносного воздействия на камеру. Известны работы, посвященные технике «водяных знаков» для обнаружения искажений на изображениях. Техника «водяных знаков» заключается в выявлении искажающих изменений, проявившихся при обработке изображений, с помощью сравнения коррелированных значений из областей изображения. Данные работы рассматривают важные аспекты обнаружения дефектов на изображениях, однако они не являются напрямую связанными с проблемой фальсификации данных камеры, так как целью исследования является создания системы, контролирующей данные с камеры в реальном времени.

Некоторые работы посвящены оценке движения на видео. Например, подход, использующий метод факторизации применяется для оценки структуры движения. В другой работе для этой же цели используется метод наименьших квадратов.

Данные методы и аналогичные им, полезны при выявлении фальсификации данных в тех случаях, когда человек изменяет направление камеры на другой угол обзора. Однако обе работы не рассчитаны на применение в условиях изменчивости среды. Также в задаче детектирования негативного влияния не так важна оценка движения на видео, сколько фиксирование события вмешательства в работу камер.

Для того чтобы выявить иные способы влияния на достоверность данных видеонаблюдения, такие как, например, попытка чем-либо закрыть объектив камеры или подмена изображения существует несколько реализаций алгоритмов.

Известна работа, в которой особое внимание уделяется рассмотрению переходов в видеоизображении с помощью интерполяцией В-сплайнами. Однако здесь подход рассчитан на медленные переходы в видеоряде, что не всегда будет иметь отношение к фальсификации.

Известен подход, в котором моделируется каждый снимок как набор граней на изображении и отслеживают их изменения, делая снимок при резком изменении граней. Этот метод также недостаточно точен в определении вредоносного влияния, так как, например, затемнение/загораживание объектива может происходить постепенно, без резких смен границ.

Известен алгоритм, в котором авторы пытаются определить изменения в изображениях видеоряда с камеры с помощью использования алгоритма, основанного на анализе цветовых гистограмм для выявления значительных изменений картины.

В аналогичном алгоритме используются цветовые гистограммы вместе со значениями пикселей изображения для нахождения различий между изображениями. Методы, основанные на использовании гистограмм, хорошо подходят для обнаружения фальсификации данных с камер, так как они предлагают глобальные измерения содержимого изображения и чувствительны к большинству типов изменений в изображении независимо от их характера или природы.

Таким образом, алгоритмы обнаружения негативного влияния на камеру должны быть чувствительны к любым значительным движениям на камере. Количество ложных обнаружений фальсификации должно быть сведено к минимуму, для того чтобы не отвлекать службы безопасности каждый раз, когда событие обнаружено. Важно отметить, что негативное воздействие на камеру является устойчивым событием, которое может происходить как быстро, так и плавно, и необходимо

Метод обнаружения вредоносного воздействия основан на использовании понятия ключевой точки, описанного в п. 9.1:

1) на вход метода подаётся эталонное изображение I_E о котором известно, что на нём отсутствует вредоносное воздействие на камеру;

2) для изображения I_E вычисляется набор ключевых точек K_E на основе алгоритма, описанного в п. 9.1;

3) для каждого поступающего изображения I_C с видеокамеры вычисляется набор ключевых точек K_C ;

4) выполняется сравнение наборов K_E и K_I : если в K_I большинство ключевых точек из K_E сохранили свои позиции, значит, воздействия не было, иначе – детектируется вредоносное воздействие на камеру.

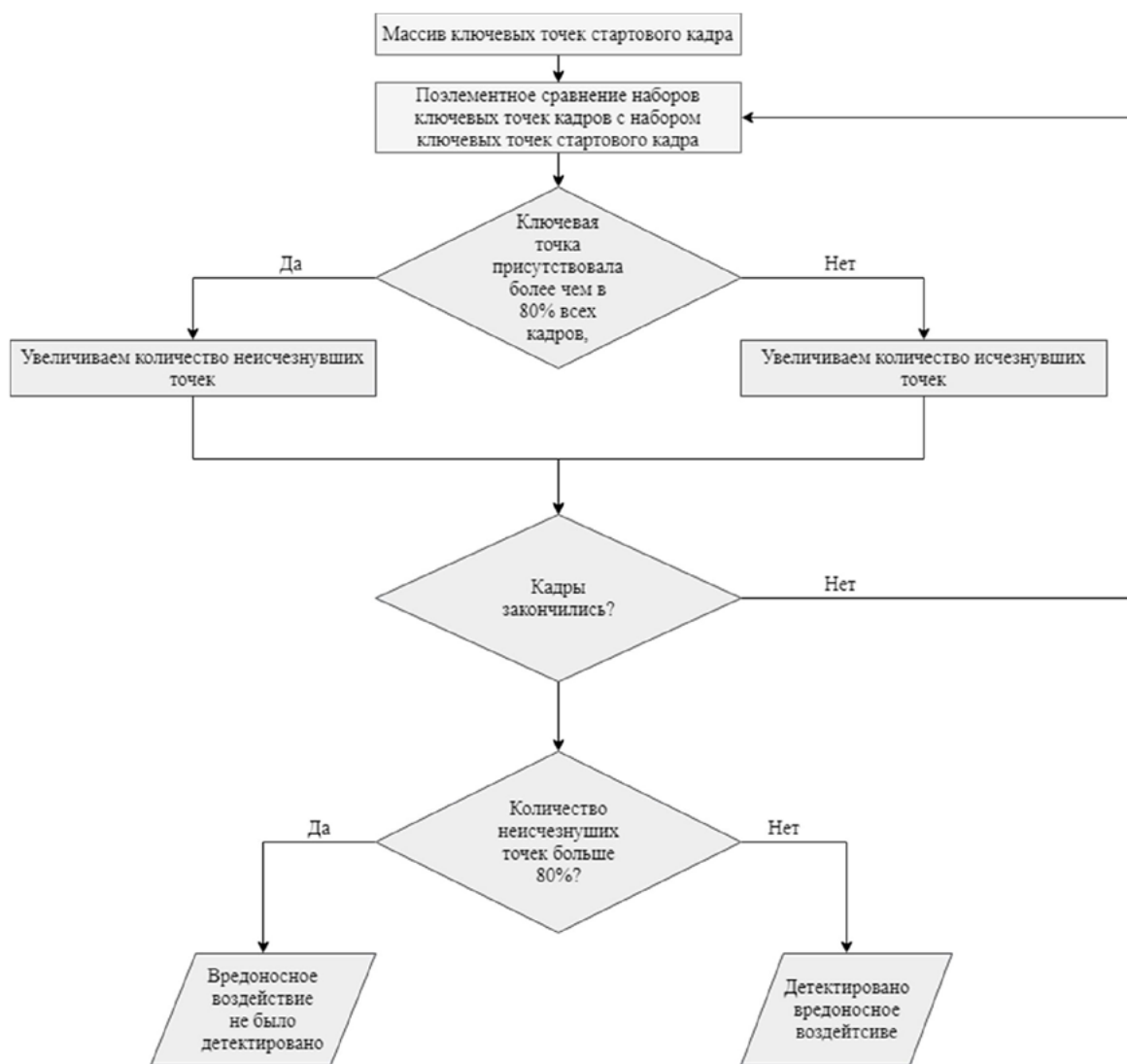


Рис. 44. Схема принятия решения алгоритма обнаружения вредоносного воздействия

9.4. Формирование карты глубины на основе стереопары

Одна из важных задач компьютерного зрения это преобразование стереопары изображений в трехмерную сцену. В результате этого процесса происходит восстановление информации о глубине (расстоянии от камеры до реальной точки сцены) каждой точки изображения. Получение точной

карты глубины – цель восстановления трёхмерности изображения.

Полученная в результате этого процесса информация о глубине может быть использована во множестве других сфер. Например, карты глубины используются для захвата опорных точек объектов в кинопроизводстве при создании спецэффектов, а также в системах компьютерного зрения, используемых на транспортных средствах для предупреждения водителя о возможном столкновении.

9.4.1. Общий принцип работы методов построения карт глубины по стереопаре

Общий алгоритм работы системы построения карты глубины по стереоизображениям включает: калибровку камеры, ректификацию изображения, сегментацию изображения, поиск соответствий между точками пары изображений, преобразование карты несоответствий в карту глубин.

Не рассматриваются первые два этапа, так как эти этапы представляют собой простые геометрические преобразования и в большинстве систем компьютерного зрения решаются на аппаратном уровне.

При анализе алгоритмов, реализующих описанные этапы, выявлены следующие проблемы.

Проблема сегментации. Данная проблема заключается в том, что часть алгоритмов сегментации не обладает достаточной точностью, в связи с чем на этапе поиска соответствий появляются множественные ошибки, связанные с некорректной сегментацией. Другая же часть алгоритмов обеспечивает достаточную точность, но обладает высокой вычислительной сложностью. Также существует проблема с соотнесением сегментов двух изображений.

Проблема поиска соответствий. Состоит в несовершенстве алгоритмов поиска соответствий в результате чего снижается точность построения карты глубины.

Проблема обработки ошибок после поиска соответствий. Заключается в том, что обычно после этапа поиска соответствий карта несоответствий содержит ряд ошибочно определенных точек и необходима их дополнительная обработка.

На основании выявленных этапов и рассмотренных существующих методов решения задачи построения карт пространственной глубины был предложен новый метод решения данной задачи на основе теории активного восприятия. Данный метод предлагает собственную комбинацию и реализацию подходов к решению указанных выше проблем на этапе сегментации и поиска соответствий.

9.4.2. Метод построения карты глубины

Предлагаемый метод решения задачи построения карты глубины применяет теорию активного восприятия (ТАВ) на этапе сегментации и поиска соответствия точек.

Для решения задачи построения карты глубины в данной работе предложен следующий алгоритм:

- 1) ввод изображения – получение изображений с камер или из файлов;
- 2) предварительная обработка – преобразование изображений в функцию яркости. Необходимо для дальнейшей работы с ТАВ;
- 3) сегментация – выделение объектов на первом изображении с целью уменьшения области поиска в дальнейшем;
- 4) поиск соответствия сегментов – поиск сегментов левого изображения на правом изображении;
- 5) формирование карты несоответствий – формирование матрицы, содержащий в себе информацию о том на сколько каждая точка первого изображения отличается по положению в пространстве от этой же точки на втором изображении;
- 6) построение карты глубины – заключительный этап восстановления информации о глубине с последующей визуализацией результатов.

Далее каждый этап будет рассмотрен подробнее.

9.4.3. Сегментация

Следующим шагом обработки является разделение изображения на сегменты. Это необходимо для уменьшения области поиска на этапе поиска соответствий.

В связи с тем, что происходит поиск соответствий точек на одних и тех же объектах, наилучшим решением было бы разделить изображение на набор объектов, тем самым сузив область поиска до внутренней области объектов. Также два исходных изображения являются эпиполярными, что позволяет разделить изображение на горизонтальные сегменты, не опасаясь потери точности.

Исходя из этого было принято решение производить сегментацию в два этапа:

- 1) разделение на горизонтальные сегменты;
- 2) выделение в горизонтальных сегментах, сегментов на основе границ объектов.

9.4.4. Выделение горизонтальных сегментов

Поскольку изображения, входящие в систему, являются эпиполярными, то все горизонтальные линии одного изображения совпадают с горизонтальными линиями другого. Поэтому можно разделить изображение на горизонтальные сегменты без ущерба точности поиска. Поскольку на следующем этапе будет применяться ТАВ, размер фильтров которого 4×4 , выберем высоту сегмента равной четырём пикселям (рис. 45).

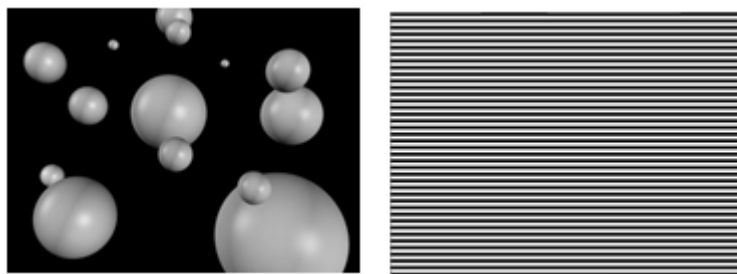


Рис. 45. Горизонтальное сегментирование

Такое разделение позволяет существенно сократить вычислительную сложность последующего алгоритма.

9.4.5. Выделение границ объектов

После того как изображение разделено на горизонтальные сегменты необходимо выделить внутри каждого горизонтального сегмента границы объектов. Для этого процесса будем использовать фильтры, позволяющие найти изменение яркости в разных направлениях (рис. 46).

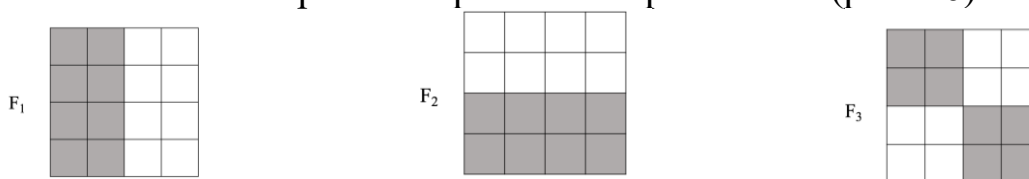


Рис. 46. Фильтры, используемые для выделения границ

Фильтр F_1 используется для выделения вертикальных границ; F_2 – для выделения горизонтальных границ; F_3 – диагональных границ.

Эти фильтры применяются для каждой точки в горизонтальном сегменте. Из полученных значений выбирается наибольшее по модулю, т.е. обозначающая наибольший перепад яркости.

Для каждого горизонтального сегмента выбирается n точек с наибольшим значением фильтра, то есть границы в данном сегменте (рис. 47).

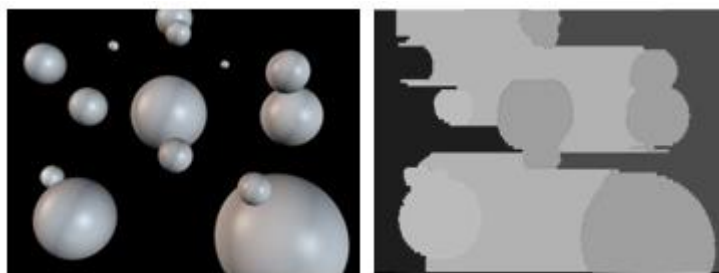


Рис. 47. Сегменты

Для последующего использования необходимо сформировать модель сегмента. Она состоит из следующих элементов:

- 1) точка начала сегмента и ее описание с помощью ТАВ;
- 2) точка окончания сегмента и ее описание с помощью ТАВ.

Для описания точки используются фильтры ТАВ. Описание точек формируется применением к ним всех 16 фильтров ТАВ.

9.4.6. Сопоставление сегментов

На данный момент одно изображение разделено на сегменты. Следующим этапом является поиска сегментов первого изображения на втором. Для этого на втором изображении выполняется поиск наиболее похожих точек для начала и конца сегмента по следующему алгоритму:

- 1) вычисляется отклик для эталонной точки по всем 16 фильтрам;
- 2) осуществляется проход окна размером 4x4 по пикселям горизонтального сегмента второго изображения. Текущий пиксель служит координатой верхнего левого угла окна размером 4x4. При проходе окна по изображению подсчитывается отклик по все 16 фильтрам;
- 3) находится модуль разности ("дельта") каждого отклика с эталонным откликом, который был найден в начале;
- 4) все шестнадцать разностей суммируются и сохраняются вместе с координатами текущего положения окна;
- 5) из всех полученных разностей находится минимальная, которая и определяет минимальное отличие найденной точки от исходной;
- 6) эта точка ставится в соответствие исходной.

Данный алгоритм выполняется для начальной и конечной точки сегмента. Таким образом формируются пары сегментов первого и второго изображения.

9.4.7. Формирование карты несоответствий

Следующим этапом алгоритма является формирование карты несоответствий – матрицы содержащий в себе информацию о том, насколько

каждая точка первого изображения отличается по положению в пространстве от этой же точки на втором изображении.

Для каждой точки сегмента происходит поиск соответствующей точки на втором изображении. Область поиска в данном случае ограничена размерами сегмента.

При нахождении нужной точки вычисляется её несоответствие по формуле

$$D = |X_1 - X_2|,$$

где X_1 – координаты точки на первом изображении; X_2 – координаты точки на втором изображении.

9.4.8. Построение карты глубины

Построение карты глубины – заключительный этап решения задачи. На этом этапе карта несоответствий преобразуется в карту глубины. Так же необходимо решить задачу возможных ошибок, допущенных на этапе поиска соответствий. Поэтому было принято решение применить ко всем точкам карты несоответствий следующую формулу:

$$D_{x,y} = \begin{cases} D_{x,y}, D \leq \mathbf{Max}, \\ \frac{\sum_{i=x-\frac{n}{2}}^{x+\frac{n}{2}} D_{i,y}}{n}, D > \mathbf{Max}, \end{cases}$$

где $D_{x,y}$ – значение карты глубины в точке; \mathbf{Max} – максимально возможное значение карты глубины; n – размер области на которой вычисляется среднее значение

Данная формула представляют собой фильтр. Иными словами, если значение точки больше ожидаемого заменяем ее значение средним значением от соседних точек. На этом восстановление глубины завершено. Для визуализации результатов применяется формула:

$$G_{x,y} = 255 * (D_{x,y} / D_{\max}),$$

где $D_{x,y}$ – значение карты глубины в точке, D_{\max} – максимальное значение карты глубины; $G_{x,y}$ – значение точки в градациях серого.

9.4.9. Вычислительный эксперимент

Для проведения вычислительного эксперимента необходимо сформировать базу данных стереоизображений. Для каждой из пар изображений в базе данных должна присутствовать эталонная карта глубины (рис. 48).



Рис. 48. Пример изображений, используемых в вычислительном эксперименте (левое, правое и карта глубины)

Вычисление точности можно выполнить следующим образом:

$$P = \frac{1}{X \cdot Y} \sum_{i=1}^X \sum_{j=1}^Y (|d_c(i, j) - d_T(i, j)|),$$

где X – ширина изображения; Y – высота изображения; d_c – карта глубины, полученная с помощью описанного метода; d_T – эталонная карта глубины.

9.5. Моделирование ассоциативной памяти

Увеличение объёмов хранимой информации (музыкальных произведений, изображений) приводит к необходимости создания новых и эффективных методов организации памяти и методов доступа к ней. Одно из возможных решений данной задачи – использование ассоциативной памяти. Основное отличие ассоциативной памяти от памяти с произвольным доступом – организация доступа к данным не по адресу, а по содержимому. Данный механизм аналогичен механизмам, присутствующим в человеческом мозгу. Например, услышав несколько нот человек способен воспроизвести в своей памяти мелодию целиком. Ассоциативная память, по сравнению с памятью с произвольным доступом, способна исправлять ошибки и дополнять входные образы, а также выполнять поиск похожих образов.

Наиболее известные модели ассоциативной памяти были созданы на основе нейронных сетей. Одно из самых первых применений ассоциативной памяти связано с созданием методов программного поиска, а также организацией работы некоторых сетевых устройств.

Известны приложения ассоциативной памяти в категоризации текстовых документов, распознавании лиц, жестов и речевых сигналов.

Настоящая работа посвящена созданию моделей ассоциативной памяти с позиций теории активного восприятия для решения задачи поиска в базах данных изображений. Теория активного восприятия позволяет, по сравнению, с нейронными сетями, разбить множество изображений на классы без использования обучения.

Аналогами разрабатываемой модели ассоциативной памяти являются методы, для создания описаний изображений и видео с целью решения задач поиска похожих видео и изображений, а также поиска изображений и видео по содержанию.

Известные методы для создания описаний изображений используют локальные и глобальные характеристики изображений. Методы SIFT и SURF часто используются при создании локальных описаний, локально-чувствительное хэширование – при создании глобальных описаний.

Для хранения описаний изображений могут использоваться деревья, например, *kd*-дерево.

Известно, что понятие "ассоциативный" отражает, в сущности, лишь факт наличия взаимосвязи между данными и не имеет отношения к самому механизму хранения данных. В психологии под ассоциацией понимается закономерная связь двух или нескольких психических процессов (ощущений, представлений, мыслей, чувств, движений и т. п.), выражающаяся в том, что появление одного из них вызывает появление другого или других психических процессов.

В рамках своего образа-покрытия операторы и группы позволяют создать связь между описаниями изображений на заданном уровне разрешения: если пара изображений *A* и *B* на заданном уровне разрешения имеют одинаковые описания D_A и D_B , следовательно, их можно считать ассоциативно связанными.

Например, на рис. 49 показано два ассоциативно связанных изображения: изображение подъемного крана и изображение буквы "Г".

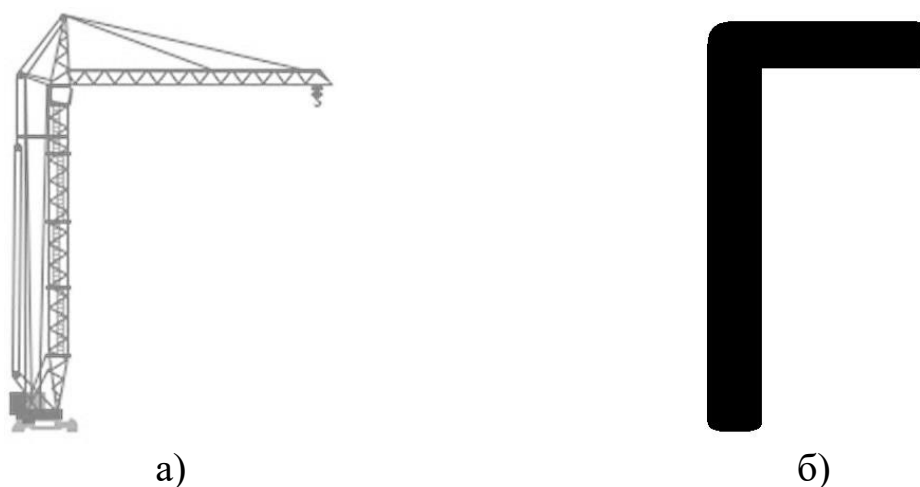


Рис. 49. Ассоциативно связанные изображения:
а – подъемный кран; *б* – буква "Г"

Теория активного восприятия позволяет описать механизм создания связи между парой изображений, однако не описывает архитектуру (способ организации) памяти. В связи с этим для организации модели ассоциативной памяти предлагается использовать иерархическую структуру – *М*-уровневое *N*-арное дерево.

Рассмотрим общую концепцию к реализации механизмов запоминания и выборки изображений, которые будут зависеть от выбранной архитектуры дерева.

Процесс запоминания изображения I включает следующие шаги:

1) формирование многоуровневого представления D изображения I :

$$\mathbf{D} = DEC [I, M],$$

где $\mathbf{D} = \{ D_{ij} \}$, $i = \overline{1, M}$, $j = \overline{1, N^{i-1}}$, M – число уровней разложения (данное значение соответствует числу уровней дерева), N – число ветвей, исходящих из узла, D_{ij} – j -я область изображения на i -м уровне, $DEC [*]$ – оператор формирования грубо-точного представления изображения;

2) формирование признакового описания изображения I :

$$\mathbf{H} = DESC [\mathbf{D}],$$

где $\mathbf{H} = \{ H_{ij} \}$, – многоуровневое (грубо-точное) признаковое описание изображения I , M – число уровней разложения (данное значение соответствует числу уровней дерева), $DESC [*]$ – оператор формирования грубо-точного признакового описания изображения, H_{ij} – описание D_{ij} области изображения;

2) выбор множества узлов **Node** на каждом уровне разрешения, в которых будет сохранено изображение I , выбор узла определяется близостью признакового описания изображения к некоторому обобщающему эталону, который задан для каждого узла дерева:

$$\mathbf{Node} = SELNODE [\mathbf{H}],$$

где $SELNODE [*]$ – оператор выбора множества узлов;

3) сохранение изображения в дерево:

$$\mathbf{T} = PUTIM [I, \mathbf{Node}],$$

где \mathbf{T} – дерево, хранящее базу данных изображений; $PUTIM [*]$ – оператор сохранения изображения в дерево.

В общем виде, процесс поиска реализуется в два шага:

1) "грубый" поиск – используется для сокращения объёмов обрабатываемой информации с учётом структуры дерева и состоит в выдвижении гипотезы об адресе узла в дереве, в котором, предположительно, находится искомое изображение:

$$\mathbf{Node} = ROUGH [\mathbf{T}, I],$$

где $ROUGH [*]$ – оператор «грубого» поиска изображения; при реализации данного оператора выполняется последовательный проход по всем уровням дерева от грубого к точному; **Node** – адрес терминального узла в дереве, например, для узла, заключённого в прямоугольник, адрес, начиная от корня дерева, имеет следующий вид:

$$\mathbf{Node} = \{ 1, 2, 2 \}, \mathbf{Node} = \{ Node_i \}, i = \overline{1, M};$$

2) "точный" поиск – заключается в сравнении описания искомого изображения с описаниями изображений, хранящимися в терминальном

узле T_{ij} , адрес которого хранится в $Node$, и выборе в качестве результирующего изображения, имеющего минимальное расстояние до искомого:

а) выборка изображений из терминального узла:

$$Data = GET [T, Node],$$

где $Data$ – изображения, хранящиеся в терминальном узле с адресом $Node$;

$Get [*]$ – оператор выборки изображений из дерева;

б) сравнение изображений из $Data$ с анализируемым изображением:

$$IND = \arg \min_{k \in \overline{1, |Data|}} FINE [Data_k, I],$$

где $Data_k$ – k -е изображение, хранящееся в узле T_{ij} ; IND – индекс найденного изображения, $FINE [*]$ – оператор «точного» поиска изображения, $| Data |$ – количество изображений, хранящихся в $Data$.

Рассмотрим возможные подходы к выбору области изображения для анализа (выбор области интереса) на i -ом уровне разрешения:

1) анализ описаний всех областей изображения на i -ом уровне разрешения; при использовании такого подхода на i -ом уровне выбирается область интереса, которая не связана с областью интереса на $(i - 1)$ -ом уровне; данный подход предлагается из тех соображений, что операторы и группы с максимальной массой наиболее устойчивы к шуму;

2) на i -ом уровне разрешения ($i > 2$) выбирается подобласть, которая:

- является дочерней для области, выбранной для анализа на $(i - 1)$ -м уровне;

- включает оператор $(V_1, \dots, V_{15}, \overline{V_1}, \dots, \overline{V_{15}})$ или группу (полную или замкнутую) с максимальной массой;

а) на первом уровне разрешения анализируется описание исходного изображения в целом;

б) на втором уровне разрешения область для анализа выбирается только на основе масс операторов или групп (т.е. без учёта родительской области);

3) на i -м уровне разрешения ($i > 2$) выбирается подобласть, которая:

- является дочерней для области, выбранной для анализа на $(i - 1)$ -м уровне;

- включает оператор V_0 с максимальной массой;

а) на первом уровне разрешения анализируется описание исходного изображения в целом;

б) на втором уровне разрешения область для анализа выбирается только с учётом массы оператора V_0 .

Принимая во внимание предложенные алгоритмы выбора области интереса, рассмотрим возможные алгоритмы выбора узла дерева для размещения изображения на i -ом уровне разрешения:

1) N -арное дерево:

а) анализ описания выбранной области изображения (есть два подхода для решения этой задачи, см. выше) и определение оператора или группы (полной или замкнутой) с максимальной массой, выбранный оператор или группа указывают на узел, в который помещается изображение;

б) анализ описаний всех областей изображения: данные (массы операторов или групп) объединяются от всех частей изображения, затем выполняется выбор максимального оператора или группы; выбранный оператор или группа определяют узел, в который помещается изображение;

в) анализ описаний всех областей изображения: среди описаний областей изображения выполняется поиск оператора или группы с максимальной массой; выбранный оператор или группа определяют узел, в который помещается изображение;

2) бинарное дерево:

а) анализ описания выбранной области изображения (есть два подхода для решения этой задачи, см. выше): формируется бинарное описание выбранной области на основе операторов или групп, затем выбирается тот узел (обобщающий эталон) до которого расстояние Хэмминга от полученного описания минимально; обобщающий эталон, представленный в виде бинарного описания, задаётся заранее;

б) анализ описаний всех областей изображения: выполняется поиск оператора или группы с максимальной массой по признаковым описаниям всех областей изображения; затем для выбранной области формируется бинарное описание выбранной области на основе операторов или групп; выбирается тот узел, до которого расстояние Хэмминга от полученного описания минимально; обобщающий эталон, представленный в виде бинарного описания, задаётся заранее.

Описанная архитектура ассоциативной памяти на основе N -арного дерева непосредственно следует из структуры алгебры изображений и алгебры групп. Бинарное дерево обладает меньшей скоростью роста узлов с увеличением числа уровней (по сравнению с N -арным деревом), но также использует элементы ТАВ. Бинарное дерево в качестве архитектуры ассоциативной памяти выбрано вследствие того, что, по Стаффорду Биру, наиболее эффективный метод разбиения множества изображений в условиях неопределённости заключается в использовании дихотомии.

9.6. Распознавание изображений текстур

Задача распознавания текстур – одна из фундаментальных проблем в области компьютерного зрения и обработки изображений. Распознавание текстур используется при автоматическом и автоматизированном анализе медицинских изображений, распознавании объектов, моделировании

окружающей среды, поиске изображений в базах данных. Методы распознавания текстур также находят своё применение при решении задачи автотопилотирования в сложных природно-климатических условиях.

Формирование признакового описания текстурированного изображения можно выполнить следующим образом. Алгоритм формирования признакового описания изображения состоит в объединении описаний отдельных областей исходного изображения, получаемых на основе полных или замкнутых групп, в гистограммы полных и замкнутых групп. Алгоритм формирования признакового описания можно записать так:

```

for  $i = 1:s_h:(N - h)$ 
  for  $j = 1:s_w:(M - w)$ 
     $I_s = I[i : (i + s_h - 1); j : (j + s_w - 1)];$ 
     $I_G = H[I_G, \Gamma];$ 
     $I_D[I_G] = I_D[I_G] + 1.$ 

```

В записи алгоритма используются следующие обозначения I_D – описание изображения (вектор), полученное в ходе работы алгоритма, I_s – область изображения, по которой формируется описание I_G , размер области I_s – $h \times w$ отсчётов, Γ – тип формируемого описания (« P_{nm} » – полные группы на операции умножения, « P_{na} » – полные группы на операции сложения, « P_s » – замкнутые группы). Описание I_G может быть получено в виде полных или замкнутых групп. Значение шага сдвига области I_s по изображению I составляет s_h отсчётов по горизонтали и s_w отсчётов по вертикали.

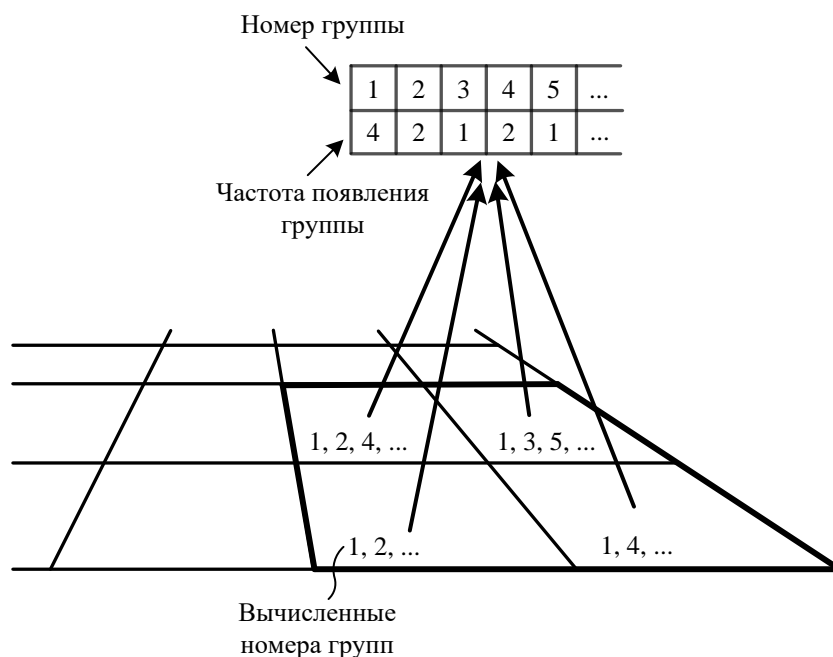


Рис. 50. Схема вычисления гистограммы групп

При использовании для описания области I_s полных групп вектор I_D включает 140 элементов, при использовании замкнутых групп – 840 элементов. На рис. 50 показана схема вычисления гистограммы групп.

В качестве исходных данных для выполнения вычислительного эксперимента можно использовать базу изображений текстур Brodatz-32. Размер каждого изображения текстуры составляет 2843×2309 отсчётов.

При решении задачи классификации можно использовать метод опорных векторов с линейным ядром.

9.7. Обнаружение объектов на изображении с использованием глобального признакового описания

Задача обнаружения объектов на изображении – одна из фундаментальных проблем в области компьютерного зрения и обработки изображений. Практическое применение данной задачи находит в системах автопилотирования, индексирования изображений и видео, охранных системах.

Структуру системы обнаружения объектов на изображении можно представить в виде совокупности трёх этапов: предварительная обработка изображения, формирование признакового описания, принятие решения.

Предварительная обработка изображения, обычно, заключается в применении к изображению фильтра, подавляющего шум. Часто реализация данного этапа не выполняется, в этом случае, ответственность за устойчивость к шуму перекладывается на метод формирования признакового описания.

При формировании признакового описания объекта используется гистограмма ориентированных градиентов, каскады Хаара и т. д.

Этап принятия решения заключается в формировании модели на основе признакового описания с использованием нейронных сетей, метода опорных векторов, k -ближайших соседей.

Известно решение задачи детектирования объекта, представляющего собой несколько соединенных сфер разного радиуса. В решении поставленной задачи определен ряд этапов: выделение границ объекта с использованием детектора Канни, поиск структурных элементов на основе преобразования Хафа, определение положения объекта с использованием распознавания по эталону и поиска изменений по кадрам. К недостаткам предложенного метода распознавания по эталону можно отнести частный характер задачи (должна быть известна структура объекта), значительные затраты времени и вычислительных ресурсов на формирование базы данных всевозможных изображений объекта.

В работе также предложено решение задачи классификации на основе нейросети. К достоинствам использования неокогнитрона при распознавании относятся отсутствие необходимости этапов поиска границ и выделения окружностей, а также возможность внесения изменений для каждой новой задачи. К недостаткам следует отнести сложность структуры сети,

порождающую большой объем вычислений и отсутствие возможности самообучения.

Известен подход к распознаванию, относящийся к классу обобщающих методов. Объект рассматривается как совокупность элементов, в качестве которых выбраны линии границ. Для обучения используется эталонное изображение объекта, на основе которого строится эталонный каркас. Исследуемое изображение обрабатывается фильтром Хаара с последующим выделением и соединением точек. Обнаружение заключается в построении каркаса исследуемого изображения, поиск наилучшего совмещения с эталоном, принятие решения об обнаружении путем сравнения количества совпавших линий с пороговым значением. К недостаткам данного метода можно отнести то, что он применим только для ограниченного класса объектов характерной формы.

Известен метод к распознаванию объекта на изображении с использованием алгоритма адаптивного усиления (AdaBoost), в основе которого лежит идея отбора и комбинирования слабых классификаторов. Тестирование реализации предложенного алгоритма показало приемлемое время обработки изображений в тренировочной выборке. Установлено также, что в реальных условиях требуется обучающая выборка большого объема, а на обучение каскада классификаторов необходимо потратить несколько дней.

Известен метод детектирования лиц на изображениях, в основу которого положены идеи алгоритма Viola&Jones. Для повышения скорости и точности детектирования в предобработке изображения предложено использовать серию фильтров, где каждый следующий фильтр обрабатывает только «перспективные» части изображения, полученные от предыдущих фильтров. При этом на основе использования информации о цвете и форме объекта, достигается отделение регионов, не содержащих лиц. Для детектирования лиц использовались совокупности анизотропных гауссовых примитивов, объединенные в каскадную модель (AdaBoost). Приведённые результаты экспериментов показывают, что предложенный гибридный метод превосходит классические по качеству и скорости распознавания, однако, сохраняются недостатки, связанные с трудоемкостью обучения каскада классификаторов.

Известен алгоритм обнаружения объектов, основанный на обучении свёрточной нейронной сети. Особенностью алгоритма является то, что в единой сети решаются сразу две задачи: выделение прямоугольных блоков, содержащих объекты и определение принадлежности объекта некоторому классу. Для распознавания исходное изображение приводится к размеру 448×448 , делится на ячейки 7×7 , для каждой ячейки формируется метка, характеризующая ее принадлежность объекту некоторого класса.

Достоинством предложенного алгоритма является высокая скорость обработки изображений. К недостаткам следует отнести затраты на обучение сети, снижение точности обнаружения по сравнению с аналогичными алгоритмами, ограничения, связанные с размерами ячеек сетки (возникают трудности при обнаружении мелких объектов и объектов, расположенных близко друг к другу). Информационная модель обнаружения объекта на изображении показана на рис. 51.

Этап обучения заключается в выполнении следующих шагов:

1) формируется «положительные» и «отрицательные» образцы объектов для каждого из обнаруживаемых классов объектов; база данных «положительных» образцов включает изображения обнаруживаемых C классов объектов, каждый образец имеет размер $h_c \times w_c$, где h_c – высота образца, w_c – ширина образца, c – класс образца ($c \in \overline{1, C}$), каждый образец содержит изображение детектируемого объекта, база данных «отрицательных» образцов строится аналогично базе для «положительных», каждый «отрицательный» образец не содержит изображение детектируемого объекта;

2) вычисляется признаковое описание для изображений, находящихся в указанных базах данных; формирование признакового описания изображения I выполняется по следующему алгоритму:

а) изображение I разбивается на L^2 равных частей без перекрытия $\mathbf{P} = \{P_i\}$, $i = \overline{1, L^2}$ размером $(N / L) \times (M / L)$ отсчётов (эквивалентность частей изображения выбрана исходя из того, что изображение обрабатывается в условиях априорной неопределённости), где P_i – i -ая область изображения, L – количество разбиений одной стороны изображения;

б) для каждой части вычисляется признаковое описание на основе U -преобразования

$$D = FC [I],$$

где оператор $FC[\bullet]$ вычисляет признаковое описание изображения и реализуется следующим образом:

$$\forall j = \overline{1, L^2} :$$

$$D = D \cup U [P_j],$$

где $U[\bullet]$ – оператор вычисления U -преобразования; D – признаковое описание изображения I . Таким образом, число частей, на которое разбивается изображение определяет размерность признакового описания: $1 \times (15 \cdot L^2)$. В результате выполнения данного шага формируется признаковое описание положительных образцов для каждого класса $F_P = \{F_{P,i}\}$ и признаковое описание отрицательных образцов для каждого класса: $F_N = \{F_{N,i}\}, i = \overline{1, C}$;

3) с использованием метода опорных векторов на основе вычисленных признаковых описаний формируется модель для детектирования

каждого из классов объектов; таким образом, каждый классификатор обучается различать два класса: «положительный» образец некоторого класса и «отрицательный» образец для того же класса:

$$\mathbf{S} = \{S_i\}, S_i = SVM_t [F_{P,i}, F_{N,i}], i = \overline{1, C},$$

где $SVM_t[\bullet]$ – оператор определения параметров модели метода опорных векторов; \mathbf{S} – множество моделей, полученных в результате обучения классификатора для каждого из классов объектов.

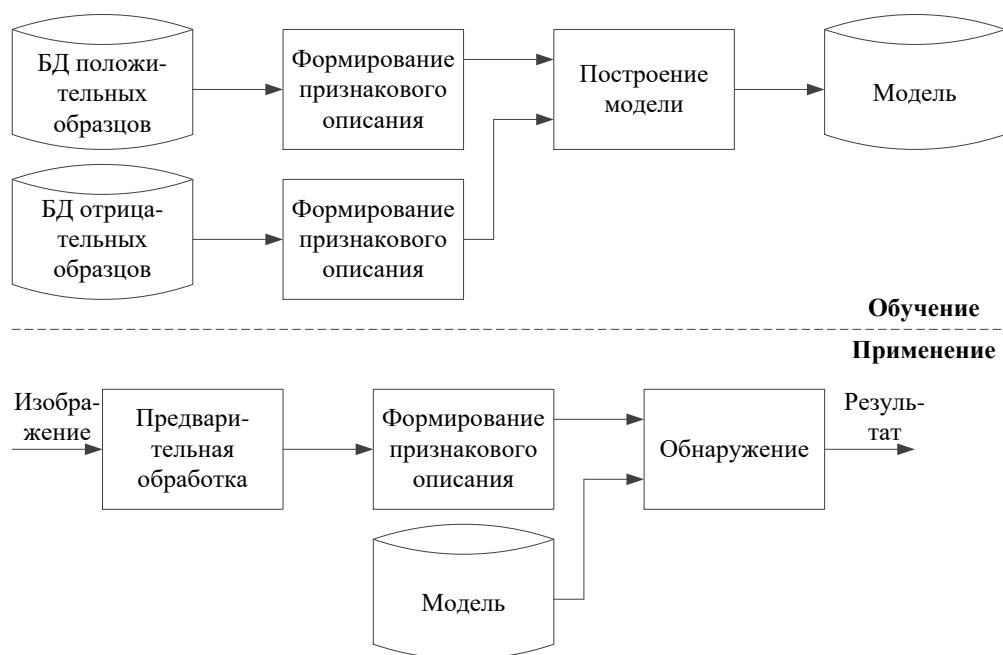


Рис. 51. Информационная модель обнаружения объектов

Этап применения заключается в выполнении следующих шагов:

1) предварительная обработка изображения I – деление изображения I , на котором выполняется детектирование объектов, на области с шагом s_h по горизонтали и s_v по вертикали, размер области для i -го класса объектов известен заранее и установлен на этапе обучения, выделение областей выполняется на нескольких масштабах; это используется для обнаружения объектов, размер которых меньше размера эталона

$$\mathbf{Q} = \{Q_{i, (x, y), s}\}, i = \overline{1, T},$$

где \mathbf{Q} – множество всех областей, полученных по изображению, $Q_{i, (x, y), s}$ – i -ая область, (x, y) – координата верхнего правого угла области, s – масштаб изображения ($0 < s < 1$), T – количество полученных областей;

2) для каждой области Q_i формируется признаковое описание:

$$D_{Q_{ixy}} = FC [Q_{i, (x, y), s}], i = \overline{1, T},$$

где $D_{Q_{ixys}}$ – признаковое описание области $Q_{i, (x, y)}$;

3) при выполнении обнаружения на вход классификатора направляются полученные на предыдущем этапе признаковые описания областей изображения:

$$C_{Qixys} = SVM_u[D_{Qixys}, S],$$

где SVM_u – оператор определения класса по признаковому описанию D_{Qixys} на основе множества моделей S , области Q и модели метода опорных векторов, C_{Qixys} – класс области $Q_{i, (x, y), s}$;

4) учитывая, что на предыдущем шаге для одного объекта тестового изображения генерируется множество близких друг другу вариантов расположения выполняется подавление немаксимумов с использованием алгоритма Soft-NMS; данный метод выбирает область с максимальной оценкой правдоподобия некоторого класса объектов, а все остальные области, перекрывающиеся в определённом процентном соотношении с данной областью, подавляет.

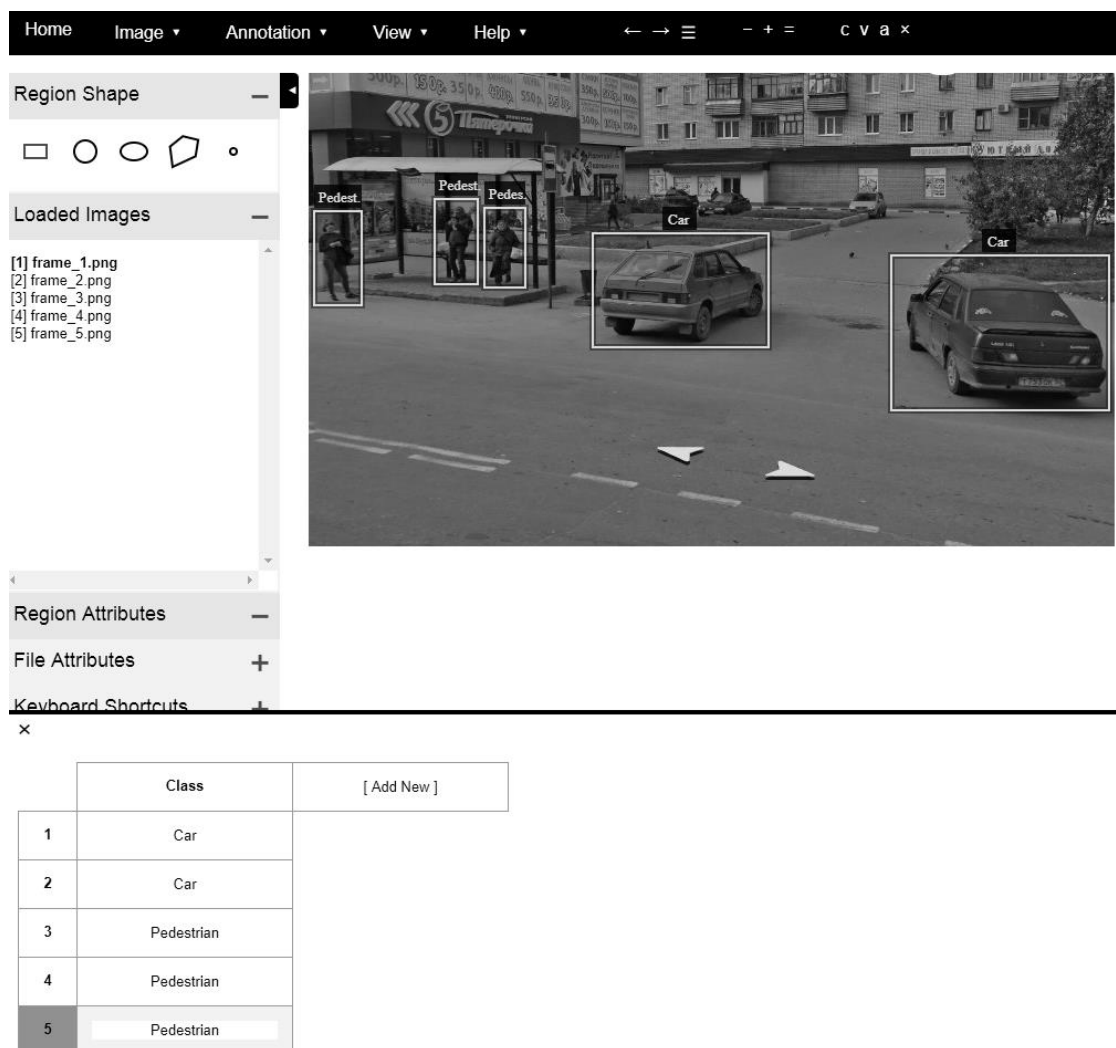


Рис. 52. Разметка тестового изображения

Для проведения вычислительного эксперимента нужно выбрать объекты двух классов, например, «Пешеход» и «Легковой автомобиль».

Сохранить в обучающую и тестовую выборку определённое количество положительных и отрицательных эталонов.

Разметку изображений можно выполнить в программе VGG Image Annotator, разработанной группой визуальной геометрии Оксфордского университета (рис. 52).

На рис. 53 показан результат детектирования объектов класса «Пешеход» и «Легковой автомобиль».

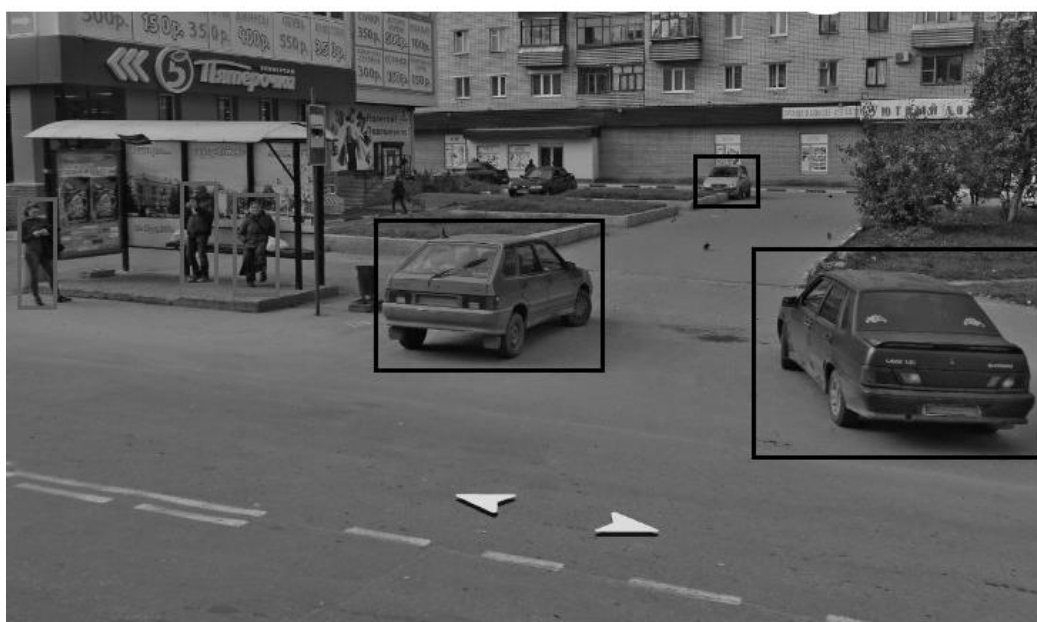


Рис. 53. Результат детектирования объектов

Контрольные вопросы

1. Подход к решению задачи распознавания изображений с помощью фильтров и операторов ТАВ.
2. Алгоритм вычисления Q -преобразования для изображений.
3. Алгоритм вычисления U -преобразования для изображений.
4. Алгоритм поиска ключевых точек на изображении.
5. Алгоритм обнаружения объектов на изображении с использованием локального признакового описания.
6. Алгоритм предварительной обработки изображения при решении задачи идентификации человека по изображению лица и поиска похожих людей по изображению лица.
7. Алгоритм формирования признакового описания изображения при решении задачи идентификации человека по изображению лица и поиска похожих людей по изображению лица.
8. Алгоритм обнаружения вредоносного воздействия на систему видеоконтроля.
9. Алгоритм формирования карты глубины на основе стереопары.

10. Алгоритм поиска похожих изображений с позиций ассоциативной памяти.

11. Алгоритм формирования дерева для поиска изображений.

12. Алгоритм поиска изображения в дереве.

13. Алгоритм обнаружения объектов на изображении с использованием глобального признакового описания.

10. ЗАДАНИЯ ДЛЯ ВЫПОЛНЕНИЯ

10.1. Методика проведения эксперимента

Тестирование точности решения задачи должно быть основано на использовании метода перекрёстной проверки. Перекрёстная проверка (cross-validation) – метод оценки аналитической модели и её поведения на независимых данных. При оценке модели, данные, имеющиеся в наличии, разбиваются на k частей. Затем на $(k - 1)$ частях данных производится обучение модели, а оставшаяся часть данных используется для тестирования. Процедура повторяется k раз; в итоге каждая из k частей данных используется для тестирования. В результате получается оценка эффективности выбранной модели с наиболее равномерным использованием имеющихся данных.

Перекрёстная проверка имеет два основных преимущества перед применением одного множества для обучения и одного для тестирования модели. Во-первых, распределение целевых значений оказывается более равномерным, что улучшает качество обучения. Во-вторых, если при каждом проходе оценить выходную ошибку модели и усреднить ее по всем прохождениям, то полученная её оценка будет более достоверной.

На практике чаще всего выбирается $k = 10$ (10 – проходная перекрёстная проверка), когда модель обучается на 9/10 данных и тестируется на 1/10. Исследования показали, что в этом случае получается наиболее достоверная оценка выходной ошибки модели.

10.2. Варианты заданий

10.2.1. Решение практических задач

При решении приведённых ниже задач необходимо выбрать три различных системы признаков. В качестве классификаторов допустим использовать нейронные сети, метод k -ближайших соседей, метод опорных векторов и т. д.

1. Решить задачу оценки уровня глюкозы в крови по электрокардиографическому сигналу. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки, диаграммы Кларка и коэффициента корреляции Спирмена.

2. Решить задачу оценки уровня давления в крови по электрокардиографическому сигналу. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки, коэффициента корреляции Спирмена и критерий Британского общества артериальной гипертензии:

Оценка	Абсолютная разница между эталоном и предсказанным значением		
	≤ 5 мм. рт. ст.	≤ 10 мм. рт. ст.	≤ 15 мм. рт. ст.
<i>A</i>	60	85	95
<i>B</i>	50	75	90
<i>C</i>	40	65	85
<i>D</i>	Хуже, чем <i>C</i>		

Данные для анализа выдаёт преподаватель.

3. Решить задачу оценки уровня давления в крови по фотоплетизмографическому сигналу. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки, коэффициента корреляции Спирмена и критерий Британского общества артериальной гипертензии. Данные для анализа выдаёт преподаватель.

4. Решить задачу вибродиагностики состояния подшипника качения. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

5. Решить задачу прогнозирования состояния подшипника качения. Данные для анализа выдаёт преподаватель.

6. Решить задачу классификации птиц по голосу. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

7. Решить задачу идентификации человека по изображению лица. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

8. Решить задачу поиска похожих людей по изображению лица. Данные для анализа выдаёт преподаватель.

9. Решить задачу обнаружения объектов заданного класса на изображении. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

10. Решить задачу выделения активности голоса в речевом сигнале. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

11. Решить задачу оценки отношения сигнал / шум для заданного класса одномерных сигналов. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

12. Решить задачу оценки частоты основ тона голоса. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

13. Решить задачу оценки эмоционального состояния диктора по голосу. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

14. Решить задачу фильтрации одномерного сигнала. Данные для анализа выдаёт преподаватель.

15. Решить задачу сжатия одномерного сигнала. Данные для анализа выдаёт преподаватель.

16. Решить задачу монауральной оценки направления на источник звука. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

17. Решить задачу распознавания текстур. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

18. Решить поиска похожих изображений на основе ассоциативной памяти. Оценку точности решения задачи выполнить с использованием метода перекрёстной проверки. Данные для анализа выдаёт преподаватель.

10.2.2. Разработка новых систем признаков

1. Разработайте и выполните тестирование на одной из приведённых выше задач признаковое описание сигнала в виде двумерной гистограммы:

- по одной из осей которых отложены операторы, а по другой – полные группы (на операции сложения, умножения) – две системы признаков;

- по одной из осей которых отложены замкнутые группы, а по другой – полные группы (на операции сложения, умножения) – две системы признаков;

- по одной из осей которых отложены операторы, а по другой – замкнутые группы – одна система признаков;

- по одной из осей которых отложены операторы, а по другой – замкнутые множества (на операции сложения, умножения) – две системы признаков;

- по одной из осей которых отложены замкнутые группы, а по другой – замкнутые множества (на операции сложения, умножения) – две системы признаков;

2. Разработайте и выполните тестирование на одной из приведённых выше задач признаковое описание сигнала в виде матрицы вероятностей переходов:

- между операторами и полными группами (на операции сложения, умножения) – две системы признаков;

- между замкнутыми группами и полными группами (на операции сложения, умножения) – две системы признаков;

- между операторами и замкнутыми группами;
- между операторами и замкнутыми множествами (на операции сложения, умножения) – две системы признаков;
- между замкнутыми группами и замкнутыми множествами (на операции сложения, умножения) – две системы признаков.

ЗАКЛЮЧЕНИЕ

Настоящее учебное пособие содержит описание современных методы распознавания образов и анализа изображений и одномерных сигналов. Приводятся авторские методики формирования признаков описания. Для ряда описанных в пособии задач распознавания образов представлен обзор наиболее современных алгоритмов.

В пособии рассматриваются теоретические основы к решению задач распознавания образов различной прикладной направленности. К задачам анализа и обработки одномерных сигналов относятся: задача вибродиагностики, фильтрации и сжатия сигналов, поиска сигналов звуковых в звуковых базах данных. Рассматриваются также следующие задачи анализа изображений: задача обнаружения объектов с использованием локального и глобального признаков описания, задача распознавания текстур, поиска похожих людей по фотографии.

Таким образом, данное учебное пособие содержит анализ и методы решения широкого круга задач интеллектуального анализа данных.

СПИСОК ЛИТЕРАТУРЫ

1. Алпайдин, Э. Машинное обучение: новый искусственный интеллект / Э. Алпайдин – М.: Издательская группа «Точка», 2017. – 208 с.
2. Баргесян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А. А. Баргесян [и др.] – 2-е изд., перераб. и доп. – СПб.: БХВПетербург, 2007. – 384 с.
3. Бринк, Х. Машинное обучение / Х. Бринк, Дж. Ричардс, М. Феве-ролф. – пер. с англ. И. Рузмайкина – СПб.: Питер, 2017. – 336 с.
4. Вьюгин, В. В. Математические основы теории машинного обучения и прогнозирования / В. В. Вьюгин. – М.: Издательство Московского центра непрерывного математического образования (МЦНМО) 2013. – 305 с.
5. Тархов, Д. А. Нейросетевые модели и алгоритмы / Д. А. Тархов – М.: Радиотехника, 2014. – 352 с.
6. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. – пер. с англ. А. А. Слинкина. – М.: ДМКПресс, 2015. – 400 с.
7. Утробин, В. А. Компьютерная обработка изображений. Основы специальной теории восприятия / В. А. Утробин. – Нижний Новгород: Нижегород. гос. техн. ун-т им. Р. Е. Алексеева, Нижний Новгород. – 2015.
8. Утробин, В. А. Физические интерпретации элементов алгебры изображения // Успехи физических наук, Т. 174. 2004. № 10. С. 1089–1104.

**Гай Василий Евгеньевич
Поляков Игорь Владимирович
Дребасов Максим Олегович
Андреева Ольга Вячеславовна**

**МОДЕЛИ И МЕТОДЫ РАСПОЗНАВАНИЯ
ОДНОМЕРНЫХ СИГНАЛОВ И ИЗОБРАЖЕНИЙ**

Редактор Н. Н. Максимова
Компьютерный набор и вёрстка В. Е. Гай

Подп. к печ. 18.03.2019. Формат 60×84 ¹/₁₆.
Печать офсетная. Бумага офсетная.
Усл. печ. л. 6,75. Тираж 33 экз. Заказ

Нижегородский государственный технический университет им. Р.Е. Алексе-
ева. Типография НГТУ, 603950, Нижний Новгород, ул. Минина, 24.