

Методы цифровой обработки звуковых сигналов на основе теории активного восприятия

Гай В. Е.

Элементы теории активного восприятия

Теория активного восприятия (ТАВ, В. А. Утробин)

Основные публикации, описывающие суть теории:

- 1. Утробин В. А. Физические интерпретации элементов алгебры изображения // Успехи физических наук, Т. 174, № 10, 2004, С. 1089–1104.**
- 2. Утробин В. А. Информационные модели системы зрительного восприятия для задач компьютерной обработки изображений: НГТУ, Н. Новгород, 2001, 234 с.**
- 3. Утробин В. А. Компьютерная обработка изображений. Принятие решений в пространстве эталонов: НГТУ, Н. Новгород, 2004, 221 с.**
- 4. Утробин, В. А. Компьютерная обработка изображений. Анализ и синтез: НГТУ, Н. Новгород, 2003, 228 с.**
- 5. Утробин, В. А. Компьютерная обработка изображений. Информационные модели этапа понимания: НГТУ, Н. Новгород, 2006, 247 с.**

Элементы теории активного восприятия

U-преобразование

Звуковой сигнал в ТАВ рассматривается как системное образование. Для обнаружения системных элементов используется интегральное преобразование, а для выявления связей между элементами – пространственное дифференцирование. Результатом выявления дифференциальной структуры является спектральное описание сигнала.

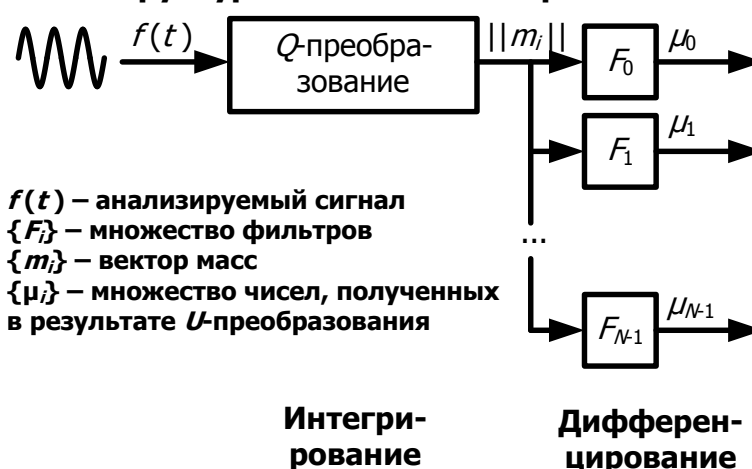


Рис. 1. U -преобразование

Алгебра групп

Алгебра групп отвечает за анализ зависимостей между спектральными коэффициентами разложения.

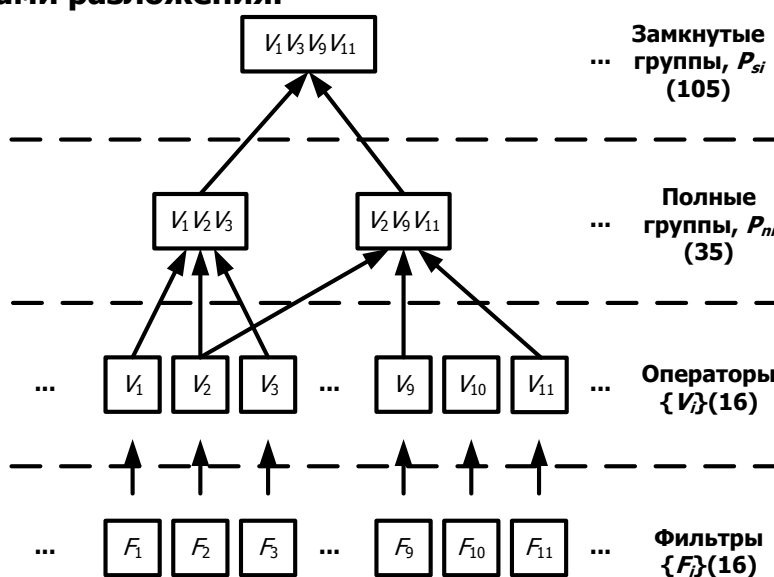


Рис. 2. Иерархия операторов и групп



Методы обработки сигналов

1. Метод вычисления прямого и обратного U -преобразования (для одномерных сигналов)

является базовым для всех других методов обработки звуковых сигналов на основе ТАВ.

2. Метод оценки отношения сигнал / шум

позволяет оценить степень искажения наблюдаемого сигнала, используется при адаптации методов обработки к свойствам наблюдаемого сигнала.

3. Метод сравнения сигналов

вычисляет степень схожести между исходным сигналом и сигналом, полученным из исходного в результате некоторого преобразования, может использоваться при оценке качества сжатия и восстановления сигнала, отличается инвариантностью к растяжению / сжатию сигналов.

4. Метод исследования стабильности описания сигнала

позволяет определить участки сигнала, наиболее устойчивые к искажениям.

5. Метод выделения полезного сигнала в условиях помех

Полезный сигнал – информация, которая требуется наблюдателю для принятия решения в рамках поставленной задачи, а помеха – вся остальная информация. В рамках ТАВ полезный сигнал рассматривается как системное образование. В этом случае он должен содержать структурные элементы и связи. Поэтому, задача обнаружения полезного сигнала сводится к задаче выделения структурных элементов сигнала и связей между ними.

Предлагаемый метод выделения полезного сигнала основан на вычислении среднеквадратического отклонения (СКО) коэффициентов U -преобразования (большее значение СКО указывает на большую вероятность присутствия полезного сигнала) и дальнейшей обработке полученной информации с помощью алгоритма кластеризации (для принятия окончательного решения о наличии полезного сигнала).

Методы обработки сигналов

5.1. Особенности

Метод обладает низкой вычислительной сложностью и высокой устойчивостью к помехам. По сравнению с корреляционным приёмником, позволяет выполнять обнаружение сигнала в условиях априорной неопределённости (необходимости знания полезного сигнала не требуется).

Предложенный алгоритм выделения активности голоса выполняет обработку сигнала во временной области и отличается от существующих алгоритмов тем, что выполняет грубо-точное выделение речевой активности. Анализируемый сигнал последовательно разбивается сначала на 2 части, потом на 4, 8 и так далее. Затем выполняется анализ каждой из частей и объединение результатов обработки. В существующих алгоритмах часто придерживаются концепции деления сигнала на более мелкие сегменты, использования предварительной фильтрации сигнала и т.д. Используемый подход позволяет повысить помехоустойчивость алгоритма.

5.2. Применение

Выделение полезного сигнала на фоне шума, выделение активности голоса, сегментация звукового сигнала по признаку тон / шум.

6. Методы сжатия и фильтрации звуковых сигналов

Сжатие и фильтрация звукового сигнала – стандартные задачи цифровой обработки сигналов. Методы, предлагаемые для решения данных задач, используют U-преобразование.

7. Метод оценки частоты основного тона

Предлагаемый метод основан на создании описаний сегментов анализируемого сигнала с помощью замкнутых групп. Далее описания сегментов попарно сравниваются и, по результатам сравнения, определяется информация о периодичности анализируемого сигнала.

Разработанный метод подобен методу оценки основного тона по автокорреляционной функции. Однако в предлагаемом решении выполняется сравнение не самих сегментов сигнала, а их описаний, что позволяет повысить его устойчивость к наличию искажений в сигнале.

Методы обработки сигналов

7.1. Особенности

Метод обладает возможностью снижения / повышения помехоустойчивости (за счёт изменения числа используемых сегментов), что одновременно приводит к повышению / снижению его быстродействия. Возможно эффективное распараллеливание алгоритма, реализующего метод.

По сравнению с другими алгоритмами, работающими во временной области, предлагаемый алгоритм позволяет обрабатывать сложные периодические сигналы (с большим количеством гармоник) без предварительного сглаживания сигнала фильтрами нижних частот.

Разработанный алгоритм оценки частоты основного тона (ЧОТ) похож на алгоритм на основе автокорреляционной функции (АКФ). В предлагаемом алгоритме оценки ЧОТ на основе АКФ сигнал сравнивается сам с собой, а в предлагаемом алгоритме сравниваются друг с другом описания элементов сигнала, которые устойчивы к различным искажениям, поскольку выявляют связи между структурными элементами сигнала. Алгоритм обладает потенциалом распараллеливания.

7.2. Применение

Известно, что превышение частоты основного тона, по сравнению с нормой, указывает, что человек находится в активном эмоциональном состоянии (испытывает радость или гнев), а снижение частоты – на то, что человек находится в пассивном эмоциональном состоянии (подавлен или в депрессии). В связи с этим, предложенный метод может использоваться для оценки эмоционального состояния абонентов call-центров, сотрудников оперативных служб. Возможно использование метода при оценке ритма музыкального произведения.

Публикации

1. Гай В. Е. Алгоритмы формирования спектрального представления звукового сигнала на основе U-преобразования // Бизнес информатика. № 1 (23), 2013, С. 44-49.
2. Gai V. E. Information approach to signal-to-noise ratio estimation of the speech signal // Communications in Computer and Information Science, V. 487, P. 137-144, 2014.
3. Gai V. E. Signal comparison algorithm in terms of a priori uncertainty // Pattern Recognition and Image Analysis, 2013, Vol. 23, No. 3, pp. 348–351.



Методы обработки сигналов

- 4. Gai V. E. A study of stability of sound signal description // Pattern Recognition and Image Analysis, 2014, Vol. 24, No. 4, pp. 463-466.**
- 5. Гай В. Е. Алгоритмы фильтрации звуковых сигналов на основе U-преобразования // Инфокоммуникационные технологии, том 10, № 3, 2012, С. 15-20**
- 6. Гай В. Е., Утробин В. А. Алгоритмы сжатия звуковых сигналов на основе U-преобразования // Вестник Нижегородского университета им. Н.И. Лобачевского, 2012, № 5(2). С. 63-66**
- 7. Гай В. Е. Метод оценки частоты основного тона в условиях помех // Цифровая обработка сигналов. 2013. № 4. С. 65-71.**

Реализация этапов системы распознавания

Реализация этапов системы распознавания образов

Процесс распознавания образов, с позиций системного анализа, можно разделить на три этапа: формирование исходного описания, нахождение системы признаков и построение решающего правила (см. рис. 3).

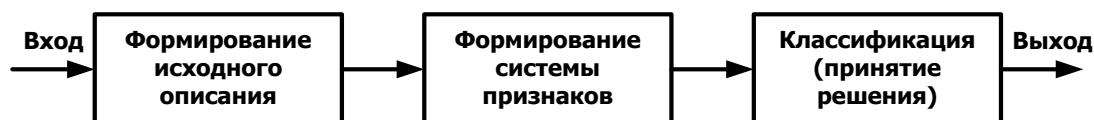


Рис. 3. Структура системы распознавания

1. Формирование исходного описания

Предварительная обработка сигнала заключается в выполнении операции интегрирования (см. рис 4). Допускается использование операции дихотомии для разбиении области определения сигнала g на подобласти.

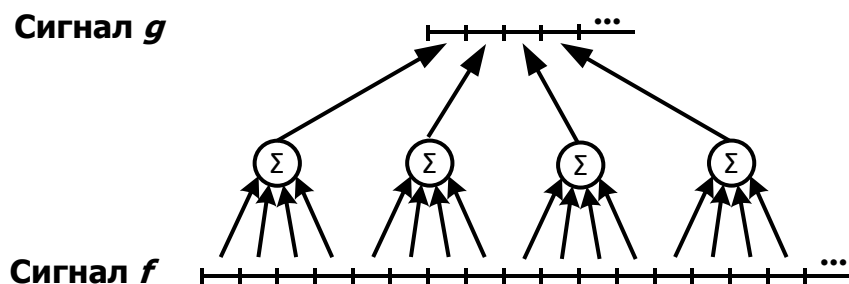


Рис. 4. Предварительная обработка сигнала

2. Формирование системы признаков

2.1) отсчёты сигнала g разбиваются на множество сегментов длиной 16 отсчётов, к каждому сегменту применяется преобразование дифференцирования, реализованное с помощью системы фильтров;

2.3) по вычисленному спектральному представлению сегмента формируется описание с помощью одной или нескольких структур, входящих в алгебру групп. Например, с помощью операторов, полных или замкнутых групп;

2.3) для объединения данных, полученных от разных сегментов сигнала, вычисляется гистограмма элементов.

3. Классификация

Этап классификации может быть реализован с помощью различных классификаторов. Например, с помощью метрического метода классификации k ближайших соседей или линейного метода опорных векторов.

Примеры решения задач распознавания

Предложенная реализация системы распознавания образов использовалась при решении следующих задач

1. Распознавание речевых сигналов

Точность распознавания на словаре из 30 слов достигает 96 %. На этапе предварительной обработки используется операция дихотомии.

2. Вибродиагностика подшипника качения

Экспериментальные исследования на базе вибросигналов показали, что предложенный метод позволяет достигнуть сто процентной точности распознавания состояния подшипника. База данных вибросигналов доступна по адресу <http://csegroups.case.edu/bearingdatacenter>. В базе данных хранятся виброакустические сигналы, записанные от подшипника без дефектов и от подшипников с различными дефектами (21 класс сигналов).

3. Распознавание радиолокационных сигналов

Исследование системы распознавания радиолокационных сигналов выполнялось на базе данных из 453 сигналов, число классов сигналов – 8. Точность классификации составляет 93 %. База данных доступна по адресу <http://cid-3aaf3e18829259c0.skydrive.live.com/home.aspx>.

4. Идентификация диктора по голосу

Экспериментальные исследования на базе речевых сигналов, в которой хранятся записи голосов 60 дикторов, показали, что предложенный метод позволяет достигнуть 100 % точности идентификации.

Примеры решения задач распознавания

5. Распознавание эмоционального состояния человека

Предложенная система признаков использовалась при распознавании эмоционального состояния человека. Тестирование выполнялось на базе данных Berlin Database of Emotional Speech (<http://emodb.bilderbar.info/index-1280.html>). В случае дикторозависимой классификации достигается точность порядка 93 %, в случае дикторонезависимой классификации – порядка 83 %.

6. Идентификация звукового сигнала

Метод идентификации звукового сигнала включает два этапа:

1) Создание цифрового отпечатка звукового сигнала:

а) построение на основе U-преобразования спектрального представления звукового сигнала;

б) выбор ключевых сегментов сигнала (ключевой сегмент соответствует участку сигнала, на котором максимален отклик небольшого количества фильтров; это означает, что сегмент сигнала подходит под образ операторов, давших максимальные отклики);

г) отбор сегментов сигнала, устойчивых к искажениям;

д) создание описания каждого сегмента с помощью замкнутых групп (совокупность описаний сегментов составляет отпечаток звукового сигнала).

2) Поиск отрывка звукового сигнала в базе данных:

а) формирование цифрового отпечатка искомого сигнала;

б) сравнение отпечатка искомого сигнала с отпечатками сигналов, хранящихся в базе данных.

Метод может применяться при мониторинге радио/телеэфира, с целью обнаружения выходов/невыходов рекламных роликов, клипов; поиск в базах данных музыкальных произведений (например, с целью идентификации неизвестной записи).

Публикации

1. Гай В. Е. Информационный подход к описанию звукового сигнала // Труды МФТИ, 2014, Т. 6, № 2, С. 167-173.

2. Gai V. E. Method of diagnostics of the state of rolling element bearing on the basis of the theory of active perception //

Примеры решения задач распознавания

**Mechanical Engineering, Automation and Control Systems (MEACS),
2014 International Conference on, pp.1-4, 16-18 Oct. 2014.**

- 3. Гай В. Е., Гай Н. В., Яковлев О. А., Сажин В. Г. Информационная модель слухового восприятия // Учёные записки физического факультета Московского Университета, 2014, № 5, 145354.**
- 4. Кондратьев В. В., Утробин В. А., Макаров Н. Н., Гай В. Е. Выявление дефектов подшипника качения с использованием системы признаков на основе теории активного восприятия // Информационно-измерительные и управляющие системы, №3, С. 31-36, 2015.**
- 5. Gai V. E. The Method and Algorithmic Support of the Determining the Presence of Information Message in a Priori an Uncertain Sound Signal // International Journal of Tomography and Simulation, V. 28, N. 2, p. 1-14. 2015.**
- 6. Гай В. Е., Утробин В.А., Лукьянчикова А.В., Поляков И.В. Распознавание изолированных речевых команд с позиций теории активного восприятия // Системы управления и информационные технологии, №3(61), 2015. – С. 75-79.**
- 7. Гай В. Е., Утробин В.А., Родионов П.А., Дербасов М.О. Оценка эмоционального состояния человека по голосу с позиций теории активного восприятия // Системы управления и информационные технологии, №1.1(59), 2015. – С. 118-122.**
- 8. М.О. Дербасов, А.А. Лаптев, А.А. Филяков, Гай В. Е. Речевое управление робототехнической системой с позиций теории активного восприятия // Труды Нижегородского государственного технического университет им. Р.Е. Алексеева. НГТУ им. Р.Е. Алексеева. – Нижний Новгород, 2015, № 3 (110). – С. 70-76.**

Особенности предлагаемого подхода

Оригинальность разрабатываемых методов связана с рассмотрением вычислительных процедур обработки звуковых сигналов с точки зрения теории активного восприятия, которая придерживается концепции грубо-точного анализа сигналов с обеспечением максимального параллелизма при обработке и распознавании. В известных подходах распознавания звуковых сигналов реализуется обработка сигналов на уровне отсчётов.

U-преобразование имеет минимально возможную вычислительную сложность, поскольку при его реализации используются простейшие операции – сложение и вычитание. Стандартные преобразования, требуют реализации свертки, а на уровне весовых коэффициентов – операции арифметического умножения.

Предлагаемый подход реализует грубо-точный (а не точно-грубый, как, например, вейвлет-преобразование) подход к обработке звукового сигнала.

Оригинальный характер разрабатываемой технологии определяется возможностью использования полученных результатов для обработки звуковых сигналов в условиях априорной неопределённости.

Междисциплинарный характер работ связан с тем, что в них проводится аналогия между создаваемыми методами обработки и информационными процессами, протекающими в слуховой системе.

Иерархическая система признаков, предлагаемая в теории активного восприятия, подобна концепции автоассоциатора, используемого в модели глубокого обучения. Однако отличается от автоассоциатора вычислительной эффективностью, обоснованностью выбора системы признаков с точки зрения системного анализа и теории информации, позволяет вычислять признаки без реализации этапа обучения.

Информация об авторе

**Гай Василий Евгеньевич, к.т.н., доцент,
доцент кафедры «Вычислительные системы
и технологии»**

**Нижегородский государственный
технический университет им. Р. Е. Алексеева
(НГТУ им. Р. Е. Алексеева)**

E-mail: vasiliy.gai@gmail.com

Сайт: <http://snip.ps/iagsav>

ORCID: <http://orcid.org/0000-0002-3644-5234>