

ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Акимов Алексей Викторович



**АЛГОРИТМЫ ОБРАБОТКИ ИНФОРМАЦИИ В ЗАДАЧАХ  
РАСПОЗНАВАНИЯ И АУГМЕНТАЦИИ СИГНАЛОВ И ИЗОБРАЖЕНИЙ НА  
ОСНОВЕ МОДЕЛЕЙ ДЕФОРМИРУЮЩИХ ИСКАЖЕНИЙ**

Специальность 05.13.01 –

Системный анализ, управление и обработка информации (радиотехника,  
автоматика, связь)

Диссертация

на соискание ученой степени кандидата

физико-математических наук

Научный руководитель  
доктор технических наук  
профессор А. А. Сирота

Воронеж – 2019

## Содержание

Введение.....	5
1 Анализ используемых моделей деформирующих искажений в задачах обработки сигналов и изображений.....	18
1.1 Задачи сопоставления временных рядов.....	21
1.1.1 Методы, основанные на сопоставлении исходных временных рядов.....	21
1.1.2 Методы, основанные на предварительном преобразовании временных рядов.....	24
1.2 Исправление искажений (дисторсии), возникающих при оптической съемке сцены .....	25
1.3 Задача оценки оптического потока.....	28
1.4 Выполнение морфинга.....	32
1.5 Учет деформирующих искажений в моделях распознавания образов .....	36
1.5.1 Эластичные модели распознавания образов .....	36
1.5.2 Статистические модели распознавания деформируемых изображений ...	38
1.6 Использование деформирующих искажений для аугментации (искусственного размножения) обучающих данных.....	39
1.7 Постановка задачи и общая схема проведения исследований в интересах построения алгоритмов обработки информации с использованием моделей деформирующих искажений.....	40
Выводы по главе.....	46
2 Модели и алгоритмы внесения деформирующих искажений в сигналы и изображения.....	50
2.1 Математическая модель внесения деформирующих искажений .....	50
2.2 Внесение деформирующих искажений в решетчатые функции дискретных аргументов .....	52
2.2.1 Интерполяция с помощью радиально-базисных функций .....	55

2.2.2	Экспериментальные исследования точности разных видов интерполяции при внесении деформирующих искажений .....	58
2.3	Эквивалентная статистическая модель влияния деформирующих искажений, основанная на приближенном представлении нелинейного оператора деформации как оператора перестановки.....	63
2.3.1	Анализ свойств оператора перестановки.....	65
2.3.2	Анализ свойств дополнительной помеховой составляющей.....	69
	Выводы по главе.....	75
3	Статистический синтез и анализ алгоритмов распознавания в условиях деформирующих искажений и аддитивного шума.....	77
3.1	Синтез и анализ оптимальных и квазиоптимальных алгоритмов распознавания цифровых сигналов в условиях деформирующих искажений .....	77
3.1.1	Постановка задачи распознавания детерминированных сигналов в присутствии шумов и деформирующих искажений .....	77
3.1.2	Основные соотношения для проведения синтеза и анализа алгоритмов распознавания.....	80
3.1.3	Результаты экспериментальных исследований алгоритмов распознавания цифровых сигналов в условиях деформирующих искажений .....	83
3.2	Синтез и анализ алгоритмов распознавания цифровых изображений в условиях деформирующих и аддитивных искажений .....	94
3.2.1	Исходная статистическая модель и оптимальный алгоритм распознавания .....	94
3.2.2	Использование непараметрических смешанных ядерных оценок.....	98
3.2.3	Анализ свойств смешанных оценок .....	102
3.2.4	Результаты экспериментальных исследований алгоритмов распознавания цифровых изображений в условиях деформирующих искажений.....	107
	Выводы по главе.....	118
4	Модели и алгоритмы искусственного размножения данных путем внесения деформирующих искажений для обучения алгоритмов распознавания лиц методом Виолы-Джонса .....	121

4.1	Обоснование и исследование алгоритмов распознавания изображений на основе метода Виолы-Джонса с использованием технологии вычислений на графических процессорах CUDA .....	122
4.1.1	Реализация последовательной версии алгоритма на основе метода Виолы-Джонса.....	122
4.1.2	Реализация параллельных версий отдельных этапов работы алгоритма на основе метода Виолы-Джонса и экспериментальные исследования по сравнению скорости их работы с последовательными версиями .....	125
4.2	Модели и алгоритмы внесения деформирующих искажений с целью аугментации обучающей выборки при распознавании лиц по методу Виолы-Джонса.....	130
4.2.1	Алгоритм внесения искусственной деформации на основе гармонических функций.....	130
4.2.2	Алгоритм внесения искусственной деформации на основе анализа смещений контрольных точек .....	132
4.2.3	Алгоритм внесения искусственной деформации на основе анализа оптического потока с использованием фильтра энтропии .....	135
4.2.4	Экспериментальные исследования влияния искусственного размножения элементов обучающей выборки на качество распознавания лиц на основе метода Виолы-Джонса.....	137
	Выводы по главе.....	140
	Заключение .....	142
	Список использованных источников .....	147

## Введение

**Актуальность темы диссертации.** Развитие современных систем управления и принятия решений в настоящее время невозможно без использования алгоритмов обработки информации, основанных на применении методов машинного обучения и элементов искусственного интеллекта [1-6].

Важным классом подобных алгоритмов являются алгоритмы поиска, обнаружения и распознавания объектов, представленных порождаемыми ими цифровыми сигналами, изображениями или многомерными признаковыми описаниями. При этом во многих реальных ситуациях в ходе цифровой обработки сигналов и изображений, наряду с наличием аддитивных искажений различной природы (шумов измерения, индивидуальных флюктуаций формы), возникает специфический вид искажающих воздействий нелинейного характера — деформирующие искажения (ДИ). В частности, воздействие ДИ случайного характера при распознавании цифровых видеосигналов приводит к тому, что их форма нелинейным образом искажается с локальным сжатием или растяжением относительно временной оси. Подобная ситуация встречается в задачах кластеризации и извлечения данных (data mining) [7-13], распознавания речи [14], жестов и движений [15-17], исследования биомедицинских данных [18-20], а также при анализе потоков объектов в системах управления процессом сепарации реального времени [21].

Не вызывает сомнений также полезность использования моделей деформирующих искажений в задачах анализа цифровых изображений, например, задачах поиска и обнаружения, а также распознавания лиц или других объектов в условиях их естественной изменчивости [22-28]. В этом свете ДИ можно воспринимать как особый вид случайных помеховых воздействий, оказывающих влияние на форму распознаваемых сигналов или изображений и имеющих высокий

уровень корреляции по времени или по пространству [29-50] (в случае анализа изображений).

В целом следует отметить, что специфика проявления ДИ в задачах обработки информации, отражающей те или иные признаковые описания объектов, связана с необходимостью построения и использования моделей деформирующих искажений при проведении синтеза и анализа алгоритмов обработки такой информации и, в частности, алгоритмов принятия решений в задачах поиска, обнаружения и распознавания объектов на основе современных методов статистической теории решений и машинного обучения.

**Степень разработанности темы диссертации.** Применение ДИ в их разных аспектах к задачам распознавания сигналов и изображений рассмотрено в работах С.Г. Антипова, Д.И. Пономарева, В.Б. Берикова, И.А. Пестунова, М.К. Герасимова, М.А. Дрюченко, А.А. Сироты, А.А. Тропченко, А.Ю. Тропченко, N. Begum, E. Keogh, C.A. Ratanamahatana, T.W. Liao, H. Ding, S. Aghabozorgi, A. Le Guennec, L. Rabiner, B.H. Juang, A. Corradini, R. Vemulapalli, A. Kampouraki, J. Wang, D. Keysers, P. Felzenswalb, L. Wiskott, X. Zhu, D. Ramanan, Ch. Nastar, M. Bichsel. Тем не менее, строгого решения данная задача в известной литературе не получила: предложенные в [7-28] алгоритмические подходы носят преимущественно эвристический характер. В связи с этим представляет интерес постановка и решение задачи распознавания объектов, представленных подверженными ДИ цифровыми сигналами и изображениями случайной формы, и получение на этой основе оптимальных и квазиоптимальных алгоритмов обработки информации в рамках статистической теории решений с использованием моделей аддитивных и деформирующих искажений.

Другой подход, позволяющий осуществлять синтез алгоритмов обработки информации в указанных условиях, основан на использовании современных методов машинного обучения. При его реализации одной из важных задач, которые приходится решать при построении таких алгоритмов, является формирование

обучающих данных в виде выборки, содержащей множество образов анализируемых объектов. Их поиск и подготовка к использованию часто требуют значительных затрат времени и ресурсов. В ряде случаев сбор необходимого числа таких образов оказывается сложен или даже невозможен из-за специфического характера предметной области.

Решение данной задачи особенно актуально для средств поиска и распознавания объектов на изображениях, например, поиска и распознавания лиц в системах видеонаблюдения, базирующихся на реализации современных методов построения алгоритмов машинного обучения. С учетом необходимости обеспечения представительности обучающих данных одним из возможных подходов здесь является *аугментация* данных, под которой далее будем понимать искусственное размножение некоторых «опорных» образов на основе стохастических или детерминистских моделей преобразования. Исследования, посвященные этой проблеме можно встретить в работах С.Г. Небабы, А.А. Захаровой, А.Е. Жуковского, Н.А. Тарасовой, P.Y. Simard, D.C. Ciresan, S.C. Wong, J. Wang, L. Perez, D. Crispell, A.L. Guennec [51-61].

В ряде исследований, например, [53, 59-61], метод искусственного размножения применяется при работе с несбалансированными обучающими выборками с малым числом примеров. В данных выборках образы одного класса (необязательно речь идет об изображениях) значительно преобладают над образами другого. При этом в [59-61] классы уравниваются за счет искусственной генерации новых значений признаков на основе результата оценки статистических параметров признаков исходной обучающей выборки. Также, известным подходом в задачах распознавания изображений, связанных с обнаружением объектов в естественной среде, является размножение обучающей выборки за счет изменения освещенности или искажения перспективы, осуществляемых для моделирования различных условий съемки и положений детектируемых объектов в кадре [56, 58].

В приложении к использованию деформирующих искажений при размножении обучающей выборки, решение задачи таким способом имеет смысл, когда эти воздействия можно рассматривать как реализации некоторой случайной функции, которая описывает различия между элементами размножаемой обучающей выборки, и которые нужно сгенерировать при известной статистической модели ДИ. Можно ожидать, что, подобный способ построения описаний классов хорошо встраивается в задачу машинного обучения алгоритмов на базе обучающих данных, полученных на основе набора реальных объектов.

Таким образом, тема диссертационной работы, посвященная алгоритмам обработки информации в задачах распознавания и аугментации сигналов и изображений на основе моделей деформирующих искажений, является актуальной.

Тема входит в план научно-исследовательских работ по кафедре технологий обработки и защиты информации ФГБОУ ВО «ВГУ» и непосредственно связана с научными направлениями Воронежского государственного университета «Проблемы передачи, получения, обработки и хранения информации» и «Математическое моделирование, программное и информационное обеспечение, методы вычислительной и прикладной математики и их применение к фундаментальным и прикладным исследованиям в естественных науках».

**Цель и задачи исследования.** Целью работы является совершенствование алгоритмов обработки информации в задачах распознавания сигналов и изображений в условиях деформирующих искажений на основе эквивалентных статистических моделей наблюдений, а также адекватных моделей аугментации обучающих данных

Для достижения цели в работе рассматриваются и решаются следующие задачи.

1. Анализ известных подходов к использованию моделей деформирующих искажений в задачах обработки сигналов и изображений в системах управления.



2. Обоснование и исследование моделей и алгоритмов формирования деформирующих искажений для сигналов и изображений.

3. Синтез и анализ алгоритмов распознавания сигналов и изображений в условиях деформирующих искажений и аддитивного шума.

4. Обоснование и исследование моделей и алгоритмов искусственного размножения данных на основе внесения в них деформирующих искажений в задачах машинного обучения на примере алгоритмов поиска и обнаружения лиц методом Виолы-Джонса.

**Объектом исследования** являются системы обработки информации и принятия решений и реализуемые в них процессы анализа цифровых сигналов и изображений.

**Предметом исследования** являются алгоритмы обработки информации в задачах распознавания и аугментации сигналов и изображений на основе моделей деформирующих искажений.

**Научная новизна.** В работе получены следующие результаты, отличающиеся научной новизной.

1. Модель прямого внесения деформирующих искажений в решетчатые функции многих переменных с использованием радиально базисных функций, а также эквивалентная статистическая модель, отличающаяся представлением случайной функции деформации с заданными статистическими характеристиками на основе нелинейного оператора перестановки (размещения с повторениями) элементов исходного дискретного сигнала (изображения) с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации, которая позволяет осуществить проведение статистического синтеза алгоритмов распознавания сигналов и изображений и построение алгоритмов искусственного размножения данных в задачах машинного обучения. Анализ свойств нелинейного оператора и дополнительной аддитивной помеховой

составляющей с учетом ограничений, накладываемых на исходную функцию внесения деформирующих искажений.

2. Оптимальные и квазиоптимальные алгоритмы распознавания цифровых сигналов в условиях деформирующих и аддитивных искажений, отличающиеся использованием точных выражений для функций правдоподобия классов и выражений для функций правдоподобия в гауссовском приближении и базирующиеся на использовании предложенной эквивалентной статистической модели деформирующих искажений и полученных соотношениях для моментов условных многомерных распределений в виде взвешенных сумм по числу возможных комбинаций размещения элементов цифровых сигналов при внесении деформации.

3. Алгоритмы машинного обучения в интересах решения задачи распознавания изображений, основанные на использовании непараметрических ядерных оценок функций правдоподобия классов, получаемых в режиме генерации ансамбля реализаций случайной функции деформации и формировании на этой основе необходимых наборов обучающих данных, что позволило обеспечить применение предложенной модели прямого внесения деформирующих искажений без ограничений вычислительной сложности, возникающих при использовании эквивалентной статистической модели. Отличительной особенностью алгоритмов является использование предложенных и исследованных с точки зрения сходимости смешанных ядерных оценок функций правдоподобия, имеющих вид свертки стандартной ядерной оценки, формируемой на основе обучающей выборки, получаемой путем внесения в эталонное изображение случайных деформирующих искажений, и плотности распределения заданного вида, порождаемой аддитивной помеховой составляющей.

4. Алгоритмическая реализация метода Виолы-Джонса поиска и обнаружения лиц на изображениях с использованием технологии параллельных вычислений на графических процессорах на всех этапах обработки информации и ее исследование

при использовании для обучения предложенных моделей аугментации реальных изображений на основе внесения деформирующих искажений. Модели и алгоритмы аугментации обучающих данных на основе базового набора реальных изображений, отличающиеся применением при внесении случайной деформации следующих подходов: разложение изображения с использованием гармонических функций; использование статистических оценок смещений контрольных точек изображений; поиск оптического потока по методу Фарнебака с наложением фильтра энтропии, применение которых позволило сократить объем обучающих данных примерно на порядок при достоверности обнаружения лиц, сопоставимой с достоверностью, получаемой на основе полной обучающей выборки.

**Теоретическая и практическая значимость.** Основные теоретические и экспериментальные результаты работы отвечают потребностям важного направления развития систем управления, принятия решений и обработки информации в части решения задач статистического синтеза и анализа алгоритмов поиска, обнаружения и распознавания образов в условиях деформирующих и аддитивных искажений, а также задач аугментации — искусственного размножения данных в интересах повышения полноты и представительности обучающих выборок при реализации современных методов машинного обучения.

Синтезированные модели и алгоритмы обработки информации в условиях деформирующих искажений и аддитивного шума, а также результаты их анализа применимы при решении задач распознавания объектов на основе анализа порождаемых ими сигналов, для которых характерны нелинейные искажения относительно временной оси, а также при решении задач поиска, обнаружения и распознавания изображений в условиях их естественной изменчивости.

Предложенные и исследованные алгоритмы распознавания сигналов и изображений могут быть использованы при реализации контура адаптивного управления в системах фотосепарации реального времени, ориентированных на анализ интенсивных неоднородных потоков объектов, а также в системах

экспресс-диагностики биологических объектов на основе классификации спектров отражения и пропускания.

Предложенные и исследованные модели деформирующих искажений при реализации алгоритмов машинного обучения применимы для искусственного размножения обучающих данных с целью сокращения затрат на их поиск и подготовку, особенно, в условиях наличия несбалансированных обучающих выборок.

**Реализация результатов работы.** Результаты в части разработки алгоритмов обработки информации в задачах распознавания сигналов и изображений использованы при постановке и выполнении СЧ НИР «Туманность-ВГУ» (2016-2018 гг., заказчик ИНМЭ РАН РФ) и при выполнении государственного задания Минобрнауки РФ по проекту № 8.3844.2017/4.6 «Разработка средств экспресс-анализа и классификации элементов неоднородного потока зерновых смесей с патологиями на основе интеграции методов спектрального анализа и машинного обучения», а также в учебном процессе ФГБОУ ВО «ВГУ» (2017-2018 гг.).

**Методы проведения исследования.** При решении поставленных задач использовались методы математического анализа, теории вероятностей и математической статистики, теории случайных полей, линейной алгебры, теории цифровой обработки сигналов и изображений, а также методы оптимизации, машинного обучения и технологии статистического имитационного моделирования.

**Область исследования.** Работа соответствует специальности 05.13.01 – Системный анализ, управление и обработка информации (радиотехника, автоматика, связь) по следующим пунктам паспорта специальности:

- разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации (п. 4 паспорта специальности 05.13.01);

- разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации (п. 5 паспорта специальности 05.13.01);

- визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации (п. 12 паспорта специальности 05.13.01).

**Положения, выносимые на защиту:**

1. Статистическая модель влияния деформирующих искажений на основе использования эквивалентного представления в виде нелинейного оператора перестановки элементов сигналов и изображений с добавлением дополнительной аддитивной помеховой составляющей.

2. Статистически оптимальный и квазиоптимальный алгоритмы распознавания цифровых сигналов в условиях деформирующих и аддитивных искажений и результаты их сравнительных исследований.

3. Алгоритмы распознавания изображений в условиях деформирующих и аддитивных искажений, основанные на использовании непараметрических смешанных ядерных оценок, и результаты их сравнительных исследований.

4. Модели и алгоритмы искусственного размножения обучающих данных на основе внесения деформирующих искажений в реальные изображения и результаты исследования их применимости в задачах машинного обучения на примере предложенной алгоритмической реализации метода Виолы-Джонса.

**Степень достоверности результатов работы.** Результаты исследований, сформулированные в диссертации, получены на основе корректного использования взаимно дополняющих друг друга теоретических и экспериментальных (имитационное моделирование, обработка реальных изображений) методов исследований, показывают совпадение результатов, полученных различными методами, между собой, а также, в ряде частных случаев, с известными результатами, полученными другими авторами, допускают ясное физическое истолкование, имеют понятную содержательную трактовку.

Корректность использованных моделей и алгоритмов обработки сигналов и изображений в условиях деформирующих искажений, а также моделей и алгоритмов искусственного размножения данных путем внесения деформирующих искажений

подтверждается совпадением результатов проведенных экспериментов с теоретическими оценками.

Таким образом, полученные результаты и выводы, изложенные в диссертации, являются достаточно обоснованными и достоверными.

**Апробация результатов работы.** Приведенные в работе выводы и рекомендации были сформулированы в ходе всестороннего обсуждения результатов исследований. Основные положения диссертационной работы докладывались и обсуждались: на XV Международной научно-технической конференции «Кибернетика и высокие технологии XXI века» (г. Воронеж) в 2014 году; на XIII, XIV и XVI Международных конференциях «Информатика: проблемы, методология, технологии» (г. Воронеж) в 2013, 2014 и 2016 годах; на I Международной научной конференции «Математическое моделирование и информационные технологии в инженерных и бизнес-приложениях» (г. Воронеж) в 2018 г.; на VII Международной научно-технической конференции «Информационные технологии в науке, образовании и производстве» (г. Белгород) в 2018 г.; на ежегодных научных сессиях и семинарах факультета компьютерных наук ВГУ.

**Публикации.** По теме диссертации опубликовано 11 [62-72] работ, в том числе 5 статей в изданиях, рекомендованных ВАК, в том числе 2 статьи в журналах, индексируемых в БД Scopus и Web of Science.

**Личный вклад автора.** Все основные результаты по теме диссертации, выносимые на защиту, получены лично автором и опубликованы в соавторстве с научным руководителем. В совместных с научным руководителем работах последнему принадлежит постановка задачи и определение направлений исследований. Автору принадлежит проведение рассуждений, обоснование используемых математических моделей и вывод аналитических соотношений, необходимых для решения поставленных задач, разработка алгоритмов обработки информации, организация экспериментов для их исследования, анализ и интерпретация полученных результатов, а также разработка программных модулей

для средств анализа сигналов и изображений. В работах [66-68], выполненных совместно с другими соавторами, автору принадлежат результаты обоснования и исследования алгоритмов распознавания изображений в условиях деформирующих искажений на основе смешанных ядерных оценок. В работах [62-64, 71] Акимову А.В. принадлежат обоснование моделей внесения деформирующих искажений и алгоритмы искусственного размножения данных на основе внесения деформирующих искажений.

**Объем и структура диссертации.** Работа состоит из введения, четырех разделов, заключения и списка литературы из 90 наименований. Объем работы составляет 157 страниц текста, включая 41 рисунок и 7 таблиц.

В первом разделе диссертации производится анализ известных подходов к использованию моделей деформирующих искажений в задачах обработки сигналов и изображений: сопоставление временных рядов для различного рода сигналов; исправление возникающих при съемке оптических искажений; расчет оптического потока при последовательной регистрации изображений; выполнение «морфинга» (morphing) — процесса плавного замещения одного изображения другим; учет деформирующих искажений в моделях распознавания образов, а также использование деформирующих искажений при аугментации (искусственном размножении) обучающих данных. Предлагается общая схема проведения исследований в интересах построения алгоритмов обработки информации для различных задач принятия решений с использованием моделей деформирующих искажений.

Во втором разделе вводится математическая модель прямого внесения деформирующих искажений для произвольной решетчатой функции от нескольких переменных. Описываются особенности и алгоритмы внесения ДИ в решетчатые функции дискретных аргументов с использованием различных видов интерполяции, в том числе на основе применения радиально-базисных функций (РБФ). Для задач синтеза алгоритмов распознавания цифровых сигналов в условиях деформирующих

искажений также вводится эквивалентная статистическая модель влияния ДИ, основанная на приближенном представлении нелинейного оператора деформации как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. Проводится анализ свойств оператора и выводятся формулы расчета получающегося количества размещений в зависимости от ограничений, накладываемых на функцию внесения ДИ.

В третьем разделе диссертации вводится математическая модель деформирующих искажений в задачах распознавания детерминированных сигналов в присутствии шумов и осуществляется синтез оптимальных алгоритмов распознавания цифровых сигналов в условиях деформирующих искажений на основе их представления как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. Синтезируются и исследуются оптимальный, основанный на точном вычислении функций правдоподобия, и квазиоптимальный, основанный на использовании гауссовского приближения функций правдоподобия, алгоритмы распознавания. Также вводится математическая модель деформирующих искажений в задачах распознавания изображений в присутствии шумов и осуществляется синтез и анализ алгоритмов их распознавания: оптимальных по аналогии с детерминированными сигналами, а также с использованием стандартных и смешанных непараметрических ядерных оценок функций правдоподобия классов. Представляется обобщенная структура алгоритма обработки информации в интересах синтеза решающих правил при распознавании сигналов и изображений в случае, когда имеются ограничения вычислительной сложности или отсутствуют необходимые исходные данные.

В четвертом разделе работы описывается и исследуется реализация метода Виолы-Джонса с использованием технологии параллельных вычислений на графических процессорах CUDA, а также рассматриваются три модели внесения



деформирующих искажений при размножении обучающей выборки в интересах построения алгоритмов распознавания лиц: на основе применения при деформации гармонических функций, на основе статистического анализа положений контрольных точек и на основе поиска оптического потока по методу Фарнебака при наложении фильтра энтропии.

Автор выражает глубокую признательность научному руководителю проф. А.А. Сироте за постоянное внимание и руководство, а также коллективам кафедры информационных систем и кафедры технологий обработки и защиты информации Воронежского государственного университета за оказанную поддержку.

## 1 Анализ используемых моделей деформирующих искажений в задачах обработки сигналов и изображений

Во многих практически важных ситуациях в ходе цифровой обработки сигналов и изображений наряду с наличием аддитивного шума возникает специфический вид помеховых воздействий нелинейного характера — деформирующие искажения (ДИ).

Математическое представление ДИ для любой исходной функции  $f(x_1, \dots, x_n)$  от  $n$  переменных можно определить следующим соотношением:

$$\begin{aligned} f(x_1 + r_1(x_1, \dots, x_n), \dots, x_n + r_n(x_1, \dots, x_n)) = \\ = f(u_1(x_1, \dots, x_n), \dots, u_n(x_1, \dots, x_n)) = g(x_1, \dots, x_n), \end{aligned}$$

где  $g(x_1, \dots, x_n)$  — результирующая деформированная функция;  $r_i(x_1, \dots, x_n)$ ,  $i = \overline{1, n}$  — функции вносимых деформирующих искажений по каждой координате, которые могут носить детерминированный или стохастический характер.

ДИ можно воспринимать как особый вид случайных помеховых воздействий, оказывающих влияние на форму распознаваемых сигналов и изображений и имеющих высокий уровень корреляции по времени или по пространству. На рисунке 1.1 изображен пример деформации одномерной непрерывной функции вида  $f(x) = \cos(x^2 / 100)$  с помощью периодической функции  $r(x) = 10 \sin(\pi(x + 13) / 18)$  — параметры  $r(x)$  подобраны для наглядности представления.

Как уже отмечалось во введении, при обработке сигналов подобная ситуация встречается в задачах кластеризации и извлечения данных [7-13], распознавания речи [14], жестов и движений [15-17], исследования биомедицинских данных [18-20], а также при анализе движущегося потока объектов, например, элементов зерновых смесей в системах сепарации реального времени [21].

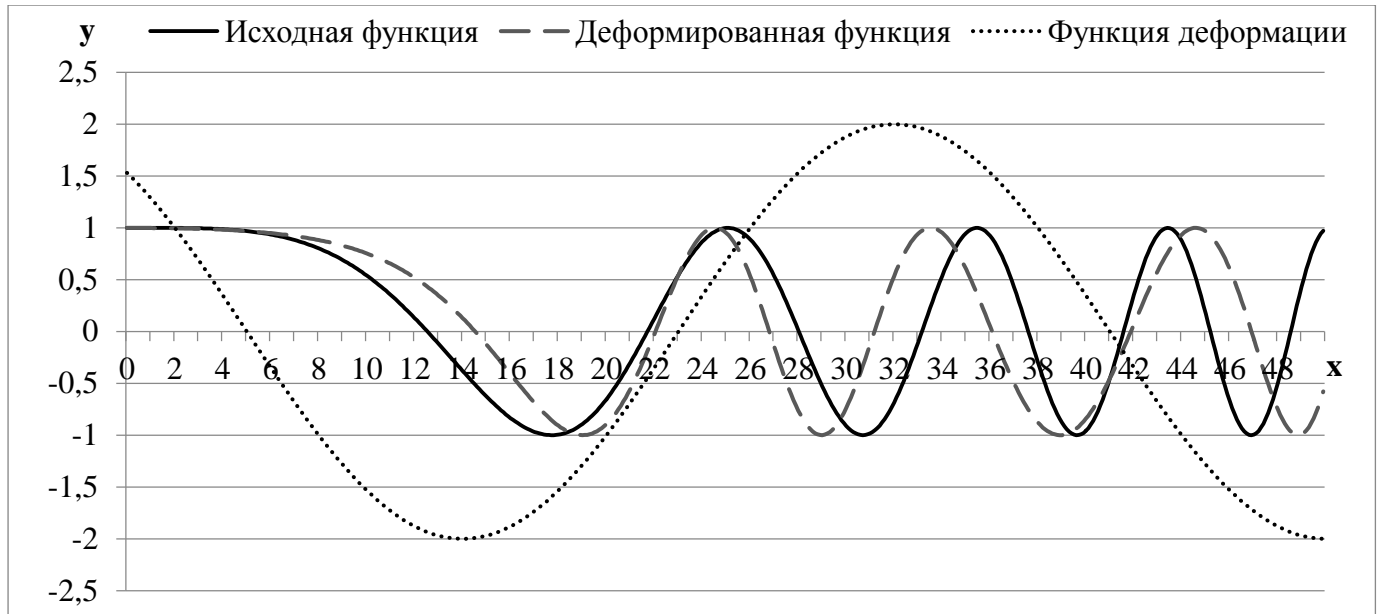


Рисунок 1.1 – Пример деформации одномерной непрерывной функции

Применительно к обработке изображений, влияние ДИ соответствует естественной изменчивости содержимого сцены при нелинейном относительно друг друга перемещении отдельных ее элементов [22-28, 37-44]. Такие ситуации возникают, например, при выражении различных эмоций на лице человека, смене ракурса при съемке или в условиях наличия aberrаций оптической системы, используемой при регистрации изображений [29-35]. Кроме того, как деформирующие искажения можно трактовать естественные изменения формы внутри одного класса объектов (например, в биологии — медицине [36] или ботанике [21]).

С учетом этих факторов подавляющее большинство предшествующих работ посвящено решению разнообразных практических задач, имеющих дело с сигналами или изображениями, подверженных, в том или ином виде, деформирующим искажениям. К ним относятся работы, в которых решались следующие задачи:

- сопоставление временных рядов для различного рода сигналов [7-20];
- исправление возникающих при съемке оптических искажений [32-35];

- расчет оптического потока при последовательной регистрации изображений в процессе изменения положений и ориентации объектов сцены в разные моменты времени, например, при их перемещениях или при смене ракурса съемки [37-44];

- выполнение «морфинга» (morphing) — процесса постепенного замещения одного изображения другим путем плавных изменений интенсивностей пикселей, положений и формы схожих в некоем смысле частей этих изображений [45-50];

- учет различий в форме анализируемых объектов в разнообразных моделях распознавания образов [22-28, 36], в том числе при искусственном размножении обучающих выборок во время построения современных алгоритмов машинного обучения [51-57].

Действительно, использование моделей ДИ является важным для построения решающих правил, основанных на концепциях машинного обучения. Как уже было указано во введении, поиск и подготовка обучающих данных к использованию часто требуют значительных затрат времени и ресурсов. В ряде случаев сбор необходимого числа образов оказывается сложен или даже невозможен из-за специфического характера предметной области. Таким образом, с учетом необходимости обеспечения представительности обучающих данных одним из возможных подходов здесь представляется искусственное размножение ограниченного набора «опорных» образов на основе стохастических или детерминистских моделей преобразования, реализующих необходимые деформации [51-57]. В подобной постановке ДИ могут являться существенным инструментом формирования обучающей выборки, содержащей множество образов анализируемых объектов.

Далее будут более подробно рассмотрены представленные в современной литературе различные аспекты проявления деформирующих искажений и их учета в задачах обработки сигналов и изображений, в том числе примеры применения ДИ при размножении элементов обучающих выборок в задачах обучения алгоритмов распознавания образов.

## **1.1 Задачи сопоставления временных рядов**

Временным рядом называется временная последовательность вещественных чисел, т.е. последовательность, значения которой рассматриваются как функция времени  $f(t)$ . Конкретный временной ряд можно представить либо как реализацию случайного процесса с заданными характеристиками, либо как детерминированную функцию времени. С задачами сопоставления временных рядов, в частности, их классификации, индексирования и кластеризации сталкиваются в самых разных областях науки и техники, экономики и медицины — везде, где требуется в том или ином виде производить сравнение процессов, развивающихся во времени [7-20].

Отличительной особенностью временных рядов является то, что помимо различий в абсолютных значениях у сравниваемых между собой элементов ряда, в том числе под воздействием шумов, для процессов, из которых они получены, часто характерна различная скорость протекания, которая, более того, сама может нелинейно изменяться по мере их развития, что во многих ситуациях требует отдельного рассмотрения или модификации имеющихся методов работы со статичными данными [9, 11].

Методы работы с временными данными условно можно разделить на два основных типа: методы, имеющие дело с исходными временными рядами при проведении сравнения, и методы, основанные на их предварительном преобразовании, например, извлечении примитивов или представлении данных временной последовательности в рамках какой-либо известной модели и последующей оценке и сравнении ее параметров [9, 13].

### **1.1.1 Методы, основанные на сопоставлении исходных временных рядов**

Данные методы основаны на поточечном расчете расстояний между исходными значениями временных рядов. Рассчитанные расстояния далее

используются как базовые значения при кластеризации, классификации или индексировании в зависимости от исходной задачи. Два часто используемых метода сопоставления из этой категории, получивших наиболее широкое освещение в литературе [9-11, 13]:

- расчет эвклидова расстояния (Euclidean Distance в англоязычной литературе — есть также модификации, использующие и другие нормы) — основан на прямом сопоставлении точек временных рядов без учета их возможных сдвигов вдоль временной оси;

- алгоритм динамической трансформации временной шкалы (Distance Time Warping или DTW) — метод, основанный на динамическом программировании и позволяющий находить соответствия между запаздывающими или опережающими друг друга участками временных рядов [8].

На рисунке 1.2 показаны в сравнении результаты сопоставления двух одинаковой формы, но смещенных относительно друг друга вдоль временной оси сигналов методами Euclidean Distance и DTW. Видно, что для метода, основанного на расчете эвклидова расстояния, полученное значение меры близости (часто это сумма расстояний между сопоставленными точками) не учитывает временного сдвига и потому показывает меньшую похожесть сравниваемых временных рядов, чем DTW. Отдельно стоит отметить, что проблема приведения друг к другу отсчетов двух неравномерно дискретизованных временных рядов легко решается при помощи интерполяции данных [8, 9, 19].

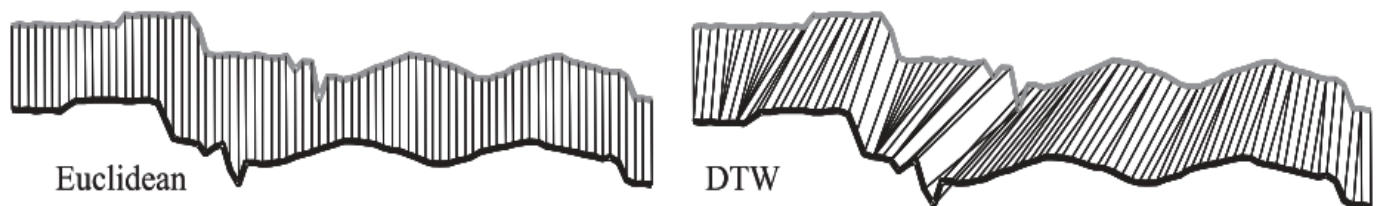


Рисунок 1.2 – Результаты сопоставления 2-х временных рядов методами Euclidian Distance (слева) и DTW (справа)

Принцип работы DTW (рисунок 1.3) в базовой своей вариации основан на построении матрицы расстояний между элементами двух временных рядов  $\mathbf{Q}$  и  $\mathbf{C}$  и поиске кратчайшего пути динамической трансформации  $\mathbf{W}$  в ней

$$dtw(\mathbf{Q}, \mathbf{C}) = \min(\sqrt{\sum_{k=1}^K \mathbf{W}_k}),$$

где  $K$  — длина пути динамической трансформации, который должен начинаться и заканчиваться в крайних точках этих двух временных рядов (граничные условия), быть неразрывным и всегда двигаться в матрице расстояний вправо, вверх или вправо-вверх [8]. Такой кратчайший путь часто ищется при помощи динамического программирования: решение о том, какой элемент пути  $\mathbf{W}$  будет следующим, принимается каждый раз на основе локального сравнения расстояний между текущим элементом пути  $\mathbf{W}$  и ближайшими к нему элементами матрицы сопоставления.

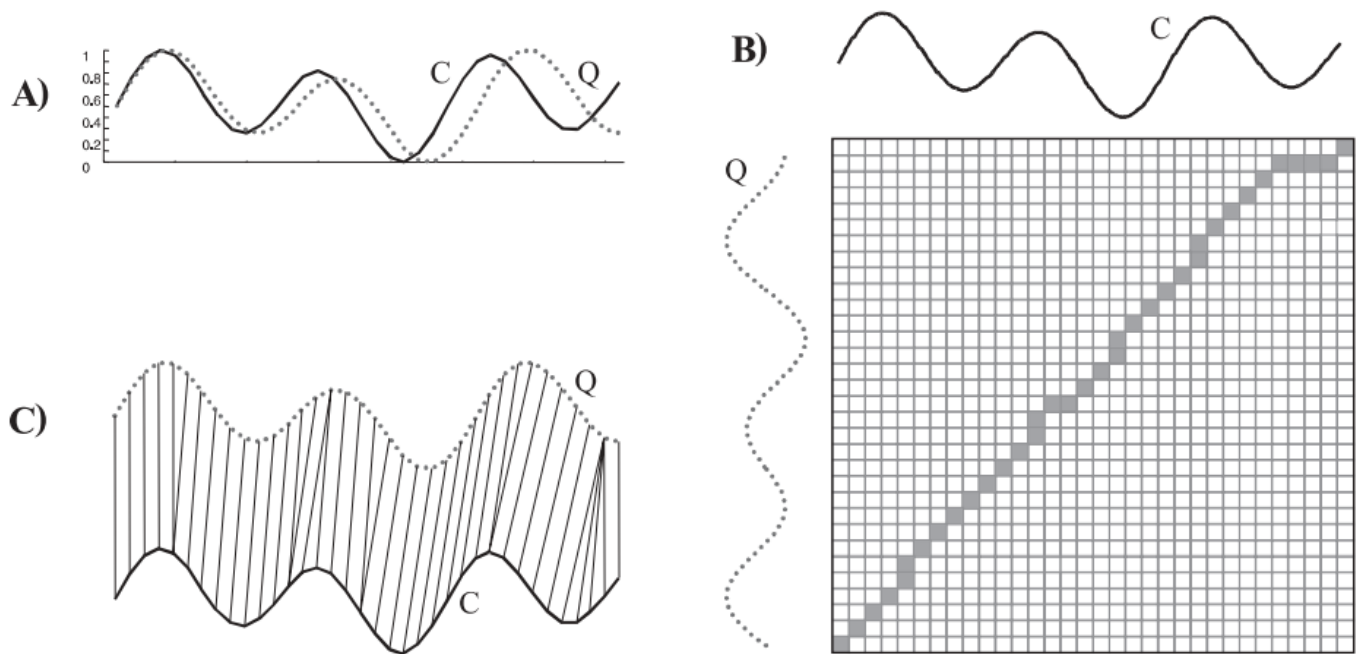


Рисунок 1.3 – Иллюстрация принципа работы метода DTW: А) два исходных временных ряда, В) матрица сопоставления расстояний с выделенным в ней найденным кратчайшим путем динамической трансформации, С) полученный результат сопоставления

Существуют разные модификации данного метода, а также другие подходы, основанные на похожих принципах динамического программирования, но исследования показывают, что лучше всех остальных себя показывают исходный метод динамической трансформации временной шкалы, а также, для больших временных рядов, метод, основанный на простом расчете эвклидова расстояния, который вдобавок имеет гораздо меньшую вычислительную сложность, чем DTW [10]. К недостаткам обоих методов можно отнести их чувствительность к шуму [11].

### **1.1.2 Методы, основанные на предварительном преобразовании временных рядов**

В основе таких методов лежит преобразование данных исходных временных рядов с целью уменьшить размерность (например, метод главных компонент), выделить примитивы (вейвлет-преобразование, масштабно инвариантное преобразование SIFT) или произвести переход в частотную область [9-11, 13]. Также к таким методам можно отнести представление временных рядов в рамках каких-либо известных моделей, при использовании которых вместо самих исходных данных производится оценка и сравнение параметров полученных моделей (например, скрытые марковские или другие статистические модели) [11, 13, 20]. Возможны сочетания описанных подходов.

Распространенными также являются методы, осуществляющие после выделения примитивов из подучастков временных рядов построение некоего их «алфавита» на основе результатов кластеризации и/или показателей их частотного распределения и производящие дальнейшую работу уже с «символами» получившегося «алфавита» вместо исходных временных рядов [10-13, 19]. К таким методам относится, например, SAX — метод символьной аппроксимации временного ряда.



Сравнительные исследования и обзоры [10, 11] показывают, что выбор конкретного подхода к использованию зависит от многих факторов, в том числе от конкретной задачи (кластеризация, индексирование или классификация), самих данных, а также желаемого соотношения скорости и качества работы получившихся алгоритмов, и, в целом, в среднем многие из представленных методов выдают примерно равные показатели точности своей работы.

Тем не менее, существующие публикации, посвященные работе с временными рядами, в большинстве своем обходят стороной сам вопрос анализа влияния деформирующих искажений на процесс их распознавания и носят, как правило, сугубо эвристический характер.

## **1.2 Исправление искажений (дисторсии), возникающих при оптической съемке сцены**

Из-за особенностей построения приборов съемки с линзовой оптикой, направления пучков света, проходящих через них, часто подвержены различным отклонениям (абберациям), и результаты съемки, таким образом, могут содержать в себе разнообразные геометрические искажения (дисторсию), которые в определенных ситуациях достигают большой величины и в ряде случаев и прикладных задач требуют учета и исправлений [29, 33, 34]. Среди прочих часто выделяют две крупных категории таких искажений: радиальные (рисунок 1.4а) и тангенциальные (рисунок 1.4б) искажения [30-32, 34, 35].

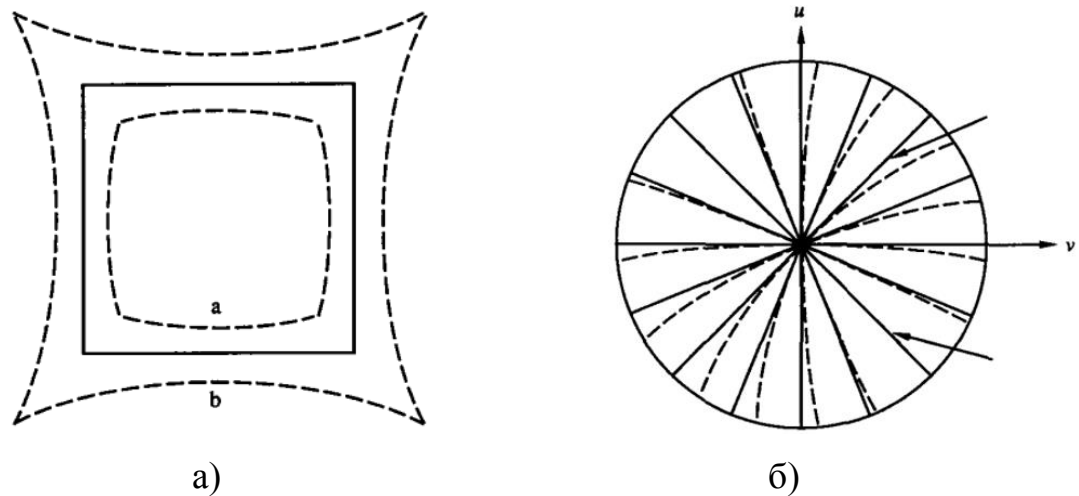


Рисунок 1.4 – Примеры возможных эффектов, вносимых аберрациями оптических систем съемки с линзами, для радиальных (а) и тангенциальных (б) типов искажений (прерывистые линии показывают возможные смещенные позиции соответствующих сплошных линий в зависимости от типа дисторсии)

Причиной радиальных искажений (рисунок 1.4а) является искривление световых лучей на краях выпуклых линз в зависимости от их расстояния до оптического центра. В большей степени данный эффект проявляется для широкоугольных, а также линз малого размера. Это искривление приводит к сучиванию или рассеиванию световых лучей в зависимости от расстояния до оптического центра и, таким образом, к стягиванию или растяжению изображения на его углах (внутренний и внешний контур соответственно на рисунке 1.4а). Тангенциальные искажения (рисунок 1.4б) в дополнение к радиальным возникают в случае неколлинеарности плоскостей расположения линз друг относительно друга или линз и оптического сенсора.

Параметры данных искажений наряду с прочими особенностями оптического прибора оцениваются во время процедуры калибровки. Принцип ее заключается в использовании специальных изображений различного вида паттернов (рисунок 1.5а) при оценке параметров отображения координат исходной трехмерной сцены в двумерные координаты зафиксированного прибором изображения. Делается это,

например, на основе проставления соответствий между исходными и полученными координатами контрольных точек паттерна (рисунок 1.5б).

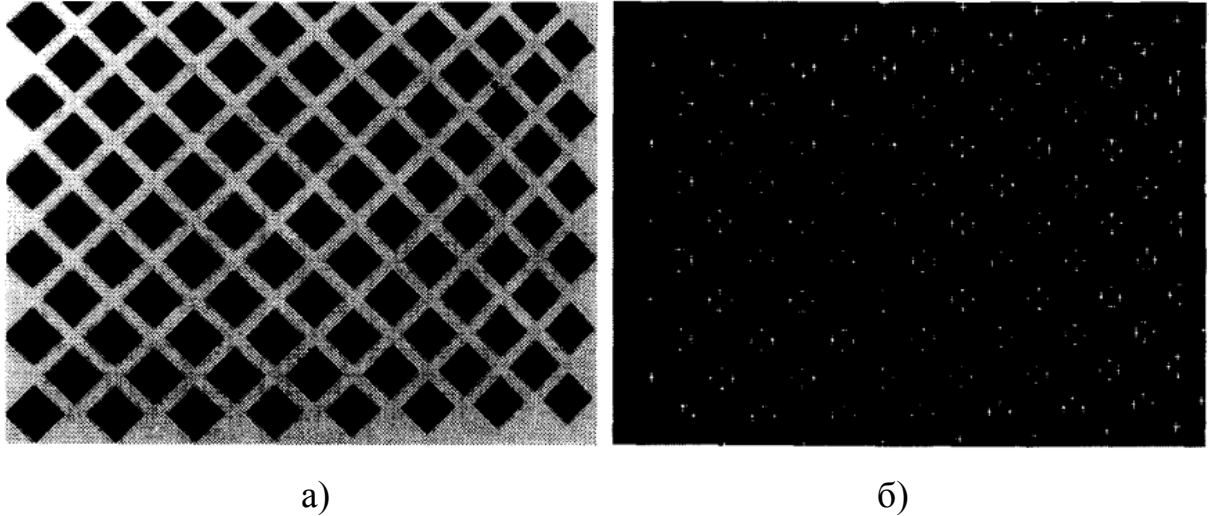


Рисунок 1.5 – Пример паттерна, используемого во время процедуры калибровки (а) и выбранные на его основе контрольные точки (б)

Для оценки параметров дисторсии часто используют представленные в виде полиномов разнообразные модели искажений, одной из которых является модель Брауна-Конради [30-32, 34, 35]:

$$\begin{aligned} x_u &= x_d + \hat{x}(k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) + [p_1(r^2 + 2\hat{x}^2) + 2p_2\hat{x}\hat{y}](1 + p_3 r^2 + \dots), \\ y_u &= y_d + \hat{y}(k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) + [p_2(r^2 + 2\hat{y}^2) + 2p_1\hat{x}\hat{y}](1 + p_3 r^2 + \dots), \end{aligned} \quad (1.1)$$

где  $r = \sqrt{\hat{x}^2 + \hat{y}^2}$ ,  $\hat{x} = x_d - x_c$ ,  $\hat{y} = y_d - y_c$ ;  $(x_d, y_d)$  — координаты точки искаженного оптическим прибором изображения;  $(x_u, y_u)$  — координаты соответствующей точки в исправленном изображении;  $(x_c, y_c)$  — координаты центра искажений (может быть смещен относительно оптического);  $k_i$  и  $p_j$  — полиномиальные коэффициенты радиальных и тангенциальных искажений соответственно, число которых бесконечно, но на практике, как правило, используют первые три или четыре.

Для определения значений этих полиномиальных коэффициентов часто используют различные методы оптимизации. На практике, в отличие от (1.1), более

полезным является обратное отображение  $(x_u, y_u) \rightarrow (x_d, y_d)$ , позволяющее для каждой точки выстраиваемого исправленного изображения узнать, какая точка искривленного изображения изначально ей соответствовала. Ряд работ, например [32, 35], посвящены именно поиску и исследованию моделей таких обратных преобразований.

Таким образом, отличительной чертой возникновения деформирующих искажений, вносимых оптическими приборами съемки, базирующимися на использовании линз, является известность моделей вносимых искажений, что основано на самой природе их проявления, связанной с особенностями построения данных приборов. Это используют при их учете и исправлении еще на самых ранних этапах получения и обработки изображений, и, следовательно, данные искажения в последующем, как правило, не играют существенной роли.

### 1.3 Задача оценки оптического потока

Задача оценки оптического потока состоит в расчете полей  $w = (w_x, w_y)$  смещений координат пикселей вдоль осей  $x$  и  $y$  при сравнении двух или более изображений и нацелена на определение изменений положений и ориентации объектов сцены в разные моменты времени, например, при их перемещениях или при смене ракурса съемки:

$$I(x + w_x(x, y), y + w_y(x, y), t + \Delta t) - I(x, y, t) = 0, \quad (1.2)$$

где  $I(x + w_x(x, y), y + w_y(x, y), t + \Delta t)$  и  $I(x, y, t)$  — изображения, взятые в разные моменты времени. Визуально оптический поток часто представляют либо в виде направленных стрелок (рисунок 1.6), либо при помощи цветовой кодировки, где интенсивность цвета отвечает за амплитуду перемещения, а сам цвет — за направление перемещения [42].

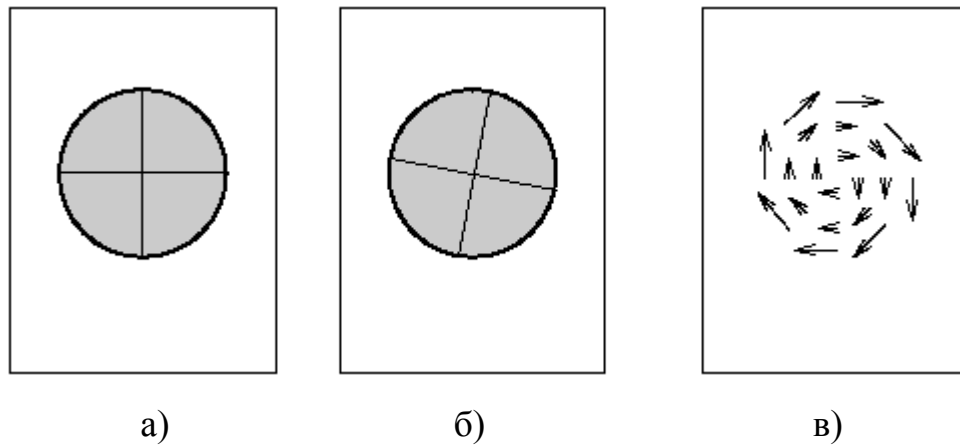


Рисунок 1.6 – Пример визуализации оптического потока между двумя изображениями меняющего свою ориентацию объекта: положение 1 (а), положение 2 (б), оптический поток между (а) и (б), визуализированный при помощи направленных стрелок (в)

Оценка оптического потока имеет множество применений: при обработке видео и регистрации изображений [37] (в том числе, в медицине), в робототехнике и для автоматической навигации [43], в системах видеонаблюдения и при оценке деформаций различных объектов — во всех областях, где требуется сравнение изменений, происходящих со временем с объектом или объектами наблюдения [38, 42, 44].

Сама постановка задачи подразумевает при этом постоянство яркости между сравниваемыми изображениями:  $\frac{dI(x, y, t)}{dt} = 0$ , что в большинстве случаев не выполняется при работе с реальными снимками. Существует множество техник, нацеленных на борьбу с данным феноменом: от учета моделей изменения освещенности в самих моделях, используемых при оценке оптического потока, и обращения с изменениями яркости как с шумом до предварительной обработки анализируемых изображений с целью нивелировать это влияние.

Другой проблемой является некорректность самой задачи: требуется оценить две компоненты смещения координат пикселей на основе всего одного ограничения.

В физическом смысле данную проблему можно охарактеризовать как проблему апертуры, заключающуюся в том, что определить правильное направление смещения группы пикселей без учета контекста их перемещения невозможно [38-40, 42, 43]. Таким образом, для решения задачи оценки оптического потока требуются дополнительные ограничения, нацеленные на увеличение его пространственной согласованности или, другими словами, — обеспечивающие гладкость изменений значений  $w(x, y)$ .

Это в свою очередь приводит к невозможности отлавливать крупные изменения в положениях и ориентации объектов сцены. Данная проблема в большинстве случаев решается путем применения методов расчета оптического потока в разных масштабах от большего к меньшему с последующим уточнением получаемых значений на более мелком уровне [40, 42]. Нежелательным эффектом, возникающим при таком подходе, является невозможность учета больших перемещений маленьких объектов внутри сцены, борьба с чем является предметом последних исследований [42].

Сами методы расчета оптического потока можно разделить на две большие группы: параметрические методы и методы регуляризации [42].

Параметрические подходы часто базируются на подборе различными методами оптимизации параметров моделей возможных перемещений положений групп пикселей при сравнении изображений или используют в своей основе разнообразные модели представления самих изображений (в том числе в виде примитивов [41]). Примером такого подхода является метод Фарнебака, основанный на полиномиальной модели [40], в котором сравниваемые изображения (предварительно сглаженные) представляются в виде квадратичных полиномов вида:

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b} \mathbf{x} + c,$$

где  $\mathbf{x} = (x, y)^T$ ;  $\mathbf{A}$ ,  $\mathbf{b}$ ,  $c$  — параметры полинома, значения которых вычисляются во время его построения в соответствии с критерием наименьших квадратов:  $\mathbf{A}$  —

симметричная матрица,  $\mathbf{b}$  — вектор,  $c$  — скаляр. При расчете оптического потока между кадрами  $f_1(\mathbf{x})$  и  $f_2(\mathbf{x})$  ищется смещение  $\mathbf{d}$ , удовлетворяющее уравнениям:

$$f_2(\mathbf{x}) = f_1(\mathbf{x} - \mathbf{d}) = (\mathbf{x} - \mathbf{d})^T \mathbf{A}_1(\mathbf{x} - \mathbf{d}) + \mathbf{b}_1(\mathbf{x} - \mathbf{d}) + c_1 = \mathbf{x}^T \mathbf{A}_2 \mathbf{x} + \mathbf{b}_2 \mathbf{x} + c_2.$$

Построение полинома при этом производится в маленьких локальных областях изображения, и, в целом, все коэффициенты:  $\mathbf{A}$ ,  $\mathbf{b}$ ,  $c$  и смещение  $\mathbf{d}$  — являются зависимыми от координат  $\mathbf{x} = (x, y)^T$ :  $\mathbf{A}(\mathbf{x})$ ,  $\mathbf{b}(\mathbf{x})$ ,  $c(\mathbf{x})$  и  $\mathbf{d}(\mathbf{x})$ .

Методы регуляризации в свою очередь основаны на наложении различных дополнительных ограничений или их комбинаций на уравнение (1.2) и нахождении соответствующего им решения также часто различными оптимизационными методами. К таким подходам относится, например, метод Хорна-Шунка [39], накладывающий ограничения на гладкость значений  $w(x, y)$  и относящийся к группе дифференциальных методов расчета оптического потока, основанных на работе с видоизмененным уравнением (1.2), получаемым путем дифференцирования  $I(x, y, t)$  и использования ряда Тейлора:

$$\frac{\partial I(x, y, t)}{\partial x} w_x(x, y) + \frac{\partial I(x, y, t)}{\partial y} w_y(x, y) + \frac{\partial I(x, y, t)}{\partial t} = 0.$$

Таким образом, существует множество различных методов оценки оптического потока, каждый из которых в большей или меньшей степени нацелен на определенную область применимости, где показывает себя лучше, чем другие, из-за заложенных в него ограничений и допущений, на основе которых и производится его расчет. Особый интерес при этом среди них представляет подгруппа параметрических методов расчета оптического потока, базирующихся на использовании примитивов, так как они позволяют оценивать перемещения и деформации схожих по смыслу элементов сравниваемых изображений, которые далеко не всегда близки в своем исходном пиксельном представлении.

## 1.4 Выполнение морфинга

Морфинг (morphing) — процесс постепенного замещения одного изображения другим путем плавных изменений значений интенсивностей их пикселей одновременно со смещениями положений и модификациями формы схожих в некоем смысле частей этих изображений. Находит свое применение в области создания эффектов анимации в компьютерной графике, а также в задачах визуализации, например, в геоинформационных системах [50] или медицине при построении обобщенных результатов компьютерной томографии, сделанных на основе нескольких разрезов [45-47].

На практике, для создания эффекта плавного перехода в компьютерной анимации процесс морфинга состоит в построении ряда промежуточных изображений от исходного до целевого, каждое из которых является смесью этих двух с большей степенью влияния того, к которому оно ближе расположено (рисунок 1.7).

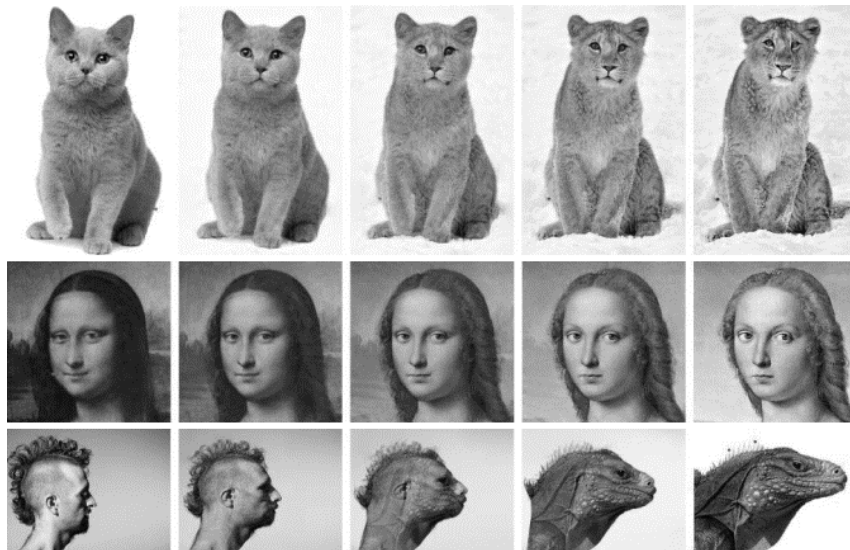


Рисунок 1.7 – Примеры морфинга из [49]

При выполнении морфинга деформация исходных изображений осуществляется для того, чтобы при смешивании значений интенсивностей их



пикселей во время построения изображений промежуточных этапов (рисунок 1.7) каждое из них выглядело естественно, и не возникало эффектов «просвечивания» одного изображения сквозь другое. Такая деформация производится для обоих исходных изображений с целью получить соответствующую текущему промежуточному этапу деформированную версию каждого из них:

$$\begin{aligned} I_{\alpha 1}(x, y) &= I_1(x + \alpha v_x(x, y), y + \alpha v_y(x, y)), \\ I_{\alpha 2}(x, y) &= I_2(x - (1 - \alpha) v_x(x, y), y - (1 - \alpha) v_y(x, y)), \end{aligned} \quad (1.3)$$

где  $I_{\alpha 1}$  и  $I_{\alpha 2}$  — соответствующие текущему промежуточному этапу морфинга деформированные изображения, полученные на основе исходных изображений  $I_1$  и  $I_2$ ;  $v = (v_x, v_y)$  — рассчитанная тем или иным образом функция деформации;  $\alpha \in [0, 1]$  — весовой коэффициент, определяющий близость получаемых деформированных изображений исходным, на которых они основаны.

Здесь представлен простейший случай, при котором деформация производится по линейному закону, определяемому для данного промежуточного этапа значениями  $\alpha v$  и  $(1 - \alpha)v$ . В более сложных случаях вместо линейной модели деформации (1.3) могут использоваться, например, квадратичные модели [49], или осуществляется ручное задание разной скорости деформации отдельных участков изображений [45, 46].

Для получения функции деформации, обеспечивающей плавность и «естественность» процедуры морфинга, сначала необходимо определить соответствия между участками исходных изображений. Делается это при помощи различных методов, которые можно подразделить на ручные и автоматические.

Среди ручных методов широко распространены: разметка контрольных точек, наложение на оконечные изображения сеток, узлы которых соответствуют друг другу, а также начертание на подвергаемых морфингу изображениях отображающихся друг в друга пар прямых или кривых линий и в принципе контуров свободной формы, которые часто при дальнейшей обработке преобразуются в те же

самые группы контрольных точек [45, 46]. Плюсом последнего подхода является гибкость в возможностях получения требуемой последовательности кадров при морфинге в противовес к трудоемкости ручного указания соответствий на исходных изображениях.

Часть автоматических методов в определенных ситуациях (близкие по виду изображения одного и того же или похожих объектов) используют техники поиска соответствий на изображениях на основе расчета оптического потока [38-42] и в целом из области регистрации изображений [37]. Другие — базируются на результатах работы разнообразных алгоритмов сегментации или распознавания и при их помощи определяют позиции и границы соответствующих друг другу участков подлежащих деформации изображений из какой-либо конкретной предметной области, например, при работе с лицами или медицинскими изображениями [45-47]. Тем не менее, сама задача нахождения общего и подходящего в большинстве случаев подхода к автоматическому поиску соответствий на подлежащих морфингу изображениях и затем на этой основе расчета значений функции деформации еще ждет своего решения.

Широко распространены также автоматизированные подходы, которые в дополнение к автоматическому поиску соответствий на изображениях учитывают необязательные ручные указания оператора [48, 49].

Таким образом, после определения тем или иным способом пар соответствий отдельных участков исходных изображений требуется рассчитать остальные значения функции деформации  $v = (v_x, v_y)$ , что часто сводится к задаче интерполяции, и для ее решения применяются разнообразные подходы: от линейной интерполяции и использования радиально базисных функций до применения оптимизационных методов, в которых используются целевые функции, основанные на различных ограничениях, накладываемых на значения  $v = (v_x, v_y)$ . Одним из таких часто используемых ограничений является сплайн тонкой пластины (основан

на поведении тонкой пластины металла при ее сгибании), нацеленный на минимизацию энергии деформации  $E_{\text{TPS}}$  путем обеспечения плавности изменений значений функции ДИ независимо вдоль осей  $x$  и  $y$  за счет накладывания ограничений на вторую производную функции деформации  $\mathbf{v} = (v_x, v_y)$  [45, 46, 49]:

$$E_{\text{TPS}}(\mathbf{v}) = E_{\text{TPS}}(v_x) + E_{\text{TPS}}(v_y),$$

$$E_{\text{TPS}}(v_{x,y}) = \left(\frac{\partial^2 v_{x,y}(x,y)}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 v_{x,y}(x,y)}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 v_{x,y}(x,y)}{\partial y^2}\right)^2.$$

Существуют также подходы, которые как задачу оптимизации (в том числе, в рамках динамического программирования [49]) решают и само нахождение близких друг к другу по какому-либо признаку участков деформируемых изображений и использующие соответствующие ограничения внутри своих целевых функций. Часто в этом случае применяется иерархический подход, и оптимизационная работа, подобно оценке оптического потока [38, 40, 42], производится в разных масштабах от большего к меньшему с постепенным уточнением результатов [48, 49].

После того, как деформированные изображения  $I_{\alpha_1}(x, y)$  и  $I_{\alpha_2}(x, y)$  построены (1.3), в самом простом случае получение итогового изображения  $I_{\alpha}(x, y)$  данного промежуточного этапа морфинга производится путем смешения значений интенсивностей их пикселей на основе того же весового коэффициента  $\alpha$ , что и в (1.3) [45, 46]:

$$I_{\alpha}(x, y) = \alpha I_{\alpha_1}(x, y) + (1 - \alpha) I_{\alpha_2}(x, y).$$

В целом, обзор имеющейся литературы по данной тематике показывает, что существует большое количество различных методов осуществления морфинга, каждый из которых показывает себя лучше или хуже в зависимости от конкретной задачи и природы исходных изображений, и предполагающих разную степень ручного вмешательства оператора.

## 1.5 Учет деформирующих искажений в моделях распознавания образов

Потенциальная польза от учета деформирующих искажений в моделях распознавания образов связана с тем, что объекты реального мира, запечатленные на изображениях, имеют тенденцию к тому, чтобы перемещаться, изменять размеры, поворачиваться и искажаться от кадра к кадру [25-28]. Существующие модели распознавания образов, учитывающие внутри себя или основанные на учете таких искажений, можно разделить на две большие группы:

- эластичные модели распознавания образов, сравнивающие изображения не только по содержимому, но и по близости расположения их отдельных частей к определенным позициям на изображении или относительно друг друга [22-26];

- статистические модели, анализирующие близость изображений на основе нахождения наиболее вероятной деформации, необходимой для преобразования их друг в друга [27, 28].

### 1.5.1 Эластичные модели распознавания образов

Эластичные модели распознавания образов в общем случае основаны на декомпозиции функции принятия решения  $d(l)$ , анализирующей какую-либо часть распознаваемого изображения, в виде суммы

$$d(l) = d_c(l) + d_p(l),$$

где  $d_c(l)$  — функция, осуществляющая сравнение близости по содержимому (это может быть классификатор любого вида или прямое сравнение интенсивностей пикселей);  $d_p(l)$  — функция, осуществляющая сравнение близости по расположению (часто рассчитывается на основе различных моделей деформации, например, сплайна тонкой пластины или других ограничений, но это может быть и просто проверка на вхождение в допустимую область относительно позиции,

принимаемой за эталон [22]), а  $l \in L$  представляет собой собственно анализируемую область из множества подлежащих исследованию областей изображения.

Итоговое решение при этом принимается на основе работы всего набора функций  $d(l)$  по количеству анализируемых частей изображения, которые могут быть организованы в виде линейной комбинации [24], скрытой марковской модели [22], графа [25] или дерева [26]. Для последних трех моделей характерно, что расчет  $d_p(l)$  производится на основе относительных позиций анализируемых частей изображений, что для графа и дерева соответствует параметрам их ребер (рисунок 1.8).

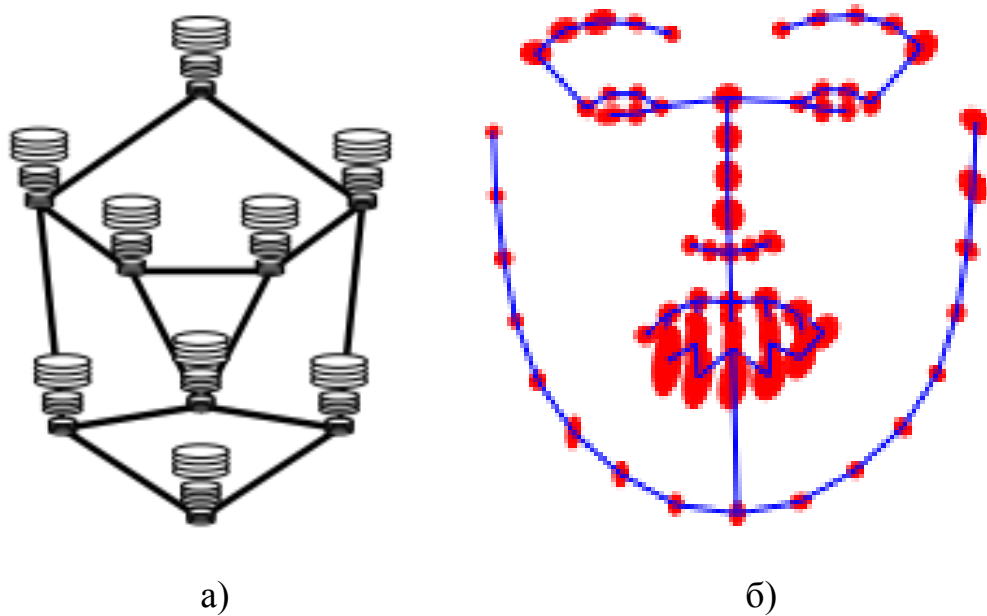


Рисунок 1.8 – Графические представления эластичных моделей распознавания на основе (а) графа (из [25]) и (б) дерева (из [26])

Задачи, которые решаются такими классификаторами можно разделить на сопоставление изображений друг с другом и поиск объектов на изображениях [22]. Для последних часто характерна большая трудоемкость исходной разметки изображений обучающей выборки [25, 26], но есть и модели, которые сами определяют позиции анализируемых подобластей изображений путем поиска их

оптимального в некотором смысле положения, в том числе и на основе методов динамического программирования [24].

### **1.5.2 Статистические модели распознавания деформируемых изображений**

Данные модели основаны на оценке вероятностей деформаций, необходимых для преобразования одних изображений в другие. Делается это в основном параметрическими методами на базе какой-либо модели, в том числе при помощи ограничений, накладываемых на допустимые деформации. Для уменьшения пространства размерности признаков среди прочих используется метод главных компонент [27, 28].

Такие подходы в своей основе предназначены для сопоставления изображений в условиях предположения постоянства яркости, но, тем не менее, возможные изменения в освещенности в них могут быть представлены в виде аддитивного шума [27].

Обучение алгоритмов распознавания здесь производится путем выбора изображения, которое будет использовано в качестве референтного, и оценки функции плотности вероятности внутриклассовых деформаций на основе известных преобразований референтного изображения во все другие изображения данного класса. При распознавании, таким образом, ищется наиболее вероятная деформация среди возможных деформаций распознаваемого изображения в изображения, которые были выбраны в качестве референтных [27, 28].

В целом, общим выводом для всех методов распознавания, основанных на учете возможных деформирующих искажений распознаваемых изображений, является то, что они нацелены на имитацию свойства изменчивости объектов реального мира и, таким образом, на практике позволяют достичь повышения показателей распознавания, а также, в отдельных случаях, понизить требуемый

объем обучающих выборок, однако, при этом, такие основанные на учете ДИ модели являются более трудоемкими для расчета.

## **1.6 Использование деформирующих искажений для аугментации (искусственного размножения) обучающих данных**

Для аугментации — искусственного размножения элементов обучающих выборок — наряду с другими методами, работающими как с исходным представлением распознаваемых данных (афинные трансформации, нарезка изображений, моделирование шума и освещенности) [51-54, 57, 58], так и в пространстве полученных на их основе признаков (Smote, DBSmote и прочие) [59-61], также используются и деформирующие искажения [51-57].

Сама идея искусственного увеличения объема обучающих выборок связана с частыми требованиями большого объема обучающих данных у современных алгоритмов распознавания и сложностью или даже невозможностью в отдельных случаях их набора и подготовки в нужном количестве [53-61]. В частности, при работе с нейронными сетями такой подход позволяет бороться с их переобучением [53-55, 59, 61]. Еще одна проблема, которую можно решить подобным образом — это несбалансированность обучающих выборок [53, 54, 59-61]. Подобные ситуации особенно часто возникают при работе с медицинскими изображениями, где набор подходящих примеров для обучения изначально ограничен в целом или для конкретного класса образов [53, 54, 60, 61]. Данная область исследований в последнее время получила дополнительный толчок вместе с распространением глубоких нейронных сетей, для которых идея аугментации обучающих данных является неотъемлемой частью их процесса обучения [52-54, 56, 57].

Существуют разные подходы к внесению деформирующих искажений в элементы обучающей выборки. В самых простых случаях они основываются на применении в качестве ДИ случайных полей, сгенерированных по некоему закону

распределения, например, при распознавании рукописных символов с целью смоделировать случайные осцилляции мускулов при письме [51-54].

При обучении алгоритмов распознавания лиц для размножения элементов обучающей выборки также применяются методы, базирующиеся на формировании трехмерных моделей головы на основе двумерных изображений лиц и построении проекций полученных моделей с целью смоделировать различные ракурсы съемки или повороты головы в кадре (здесь также применяются деформации самой трехмерной модели с целью получить большее соответствие конкретному лицу) [55, 56].

Деформирующие искажения используются в том числе и при размножении обучающих выборок в задачах классификации временных рядов [57].

Результаты экспериментов в данной области говорят о том, что размножение обучающих выборок на основе внесения в них деформирующих искажений во многих случаях увеличивает точность работы алгоритмов распознавания [51-57]. Особо отмечается, что наилучших результатов можно достичь только в случае наличия информации о параметрах возможных деформирующих искажений, описывающих различия между элементами обучающей выборки [53]. Таким образом, на данный момент перспективными являются исследования, посвященные нахождению параметров таких искажений, для чего представляется полезным использование подходов и разработок из областей морфинга и расчета оптического потока.

### **1.7 Постановка задачи и общая схема проведения исследований в интересах построения алгоритмов обработки информации с использованием моделей деформирующих искажений**

На основе проведенного анализа известных работ в рассматриваемой области необходимо отметить следующее.

Применение деформирующих искажений в их разных аспектах и проявлениях к задачам распознавания сигналов и изображений рассматривается в разнообразных



работах, посвященных сопоставлению временных рядов [7-20], а также распознаванию образов [22-28, 36], где они используются для учета изменений и вариаций в форме анализируемых объектов. Тем не менее, предложенные в данных работах алгоритмические подходы носят преимущественно эвристический характер, связанный с прикладным характером решаемых в них задач. В связи с этим представляет интерес постановка и решение задачи распознавания объектов, представленных подверженными ДИ цифровыми сигналами и изображениями случайной формы, и получение на этой основе оптимальных и квазиоптимальных алгоритмов обработки информации в рамках статистической теории решений с использованием моделей аддитивных и деформирующих искажений.

Другой подход, позволяющий осуществлять синтез алгоритмов обработки информации в указанных условиях, основан на использовании современных методов машинного обучения. При его реализации одной из важных задач, которые приходится решать при построении таких алгоритмов, является формирование обучающих данных в виде выборки, содержащей множество образов анализируемых объектов. В ряде случаев сбор необходимого числа таких образов требует значительных затрат времени и ресурсов или даже оказывается невозможен из-за специфического характера предметной области.

С учетом необходимости обеспечения представительности обучающих данных одним из возможных подходов здесь является аугментация данных или искусственное размножение обучающей выборки. В приложении к использованию моделей деформирующих искажений, решение задачи таким способом имеет смысл, когда их можно рассматривать как реализации некоторой случайной функции, описывающих различия между элементами размножаемой обучающей выборки, и которые нужно сгенерировать при известной статистической модели ДИ [51-57].

В соответствии с поставленными задачами в рамках диссертации была предложена и реализована общая схема проведения исследований в интересах построения алгоритмов обработки информации для решения задач принятия

решений различных типов (поиска и обнаружения объектов, распознавания образов), отражающая логику решения различных частных задач и связи между ними и представленная на рисунке 1.9.

Изначально в качестве основы для проведения дальнейших исследований потребовалось обосновать общую математическая модель внесения деформирующих искажений в решетчатые функции многих переменных, а также исследовать особенности, возникающие при реализации соответствующих алгоритмов обработки информации.

Далее, в зависимости от структуры исходных данных, было реализовано два подхода к проведению исследований:

- синтез алгоритмов в рамках статистической теории решений, если порождаемые наблюдаемыми объектами сигналы или изображения представлены функциями известной формы, а также известны (полностью или частично) статистические описания деформирующих и аддитивных искажений, рассматриваемых как реализации случайных функций (полей);

- построение алгоритмов на основе методов машинного обучения с аугментацией обучающих данных в случае, если порождаемые наблюдаемыми объектами сигналы или изображения представлены только наборами (выборками) обучающих данных в ограниченном количестве и при этом заранее неизвестны статистические описания деформирующих искажений.

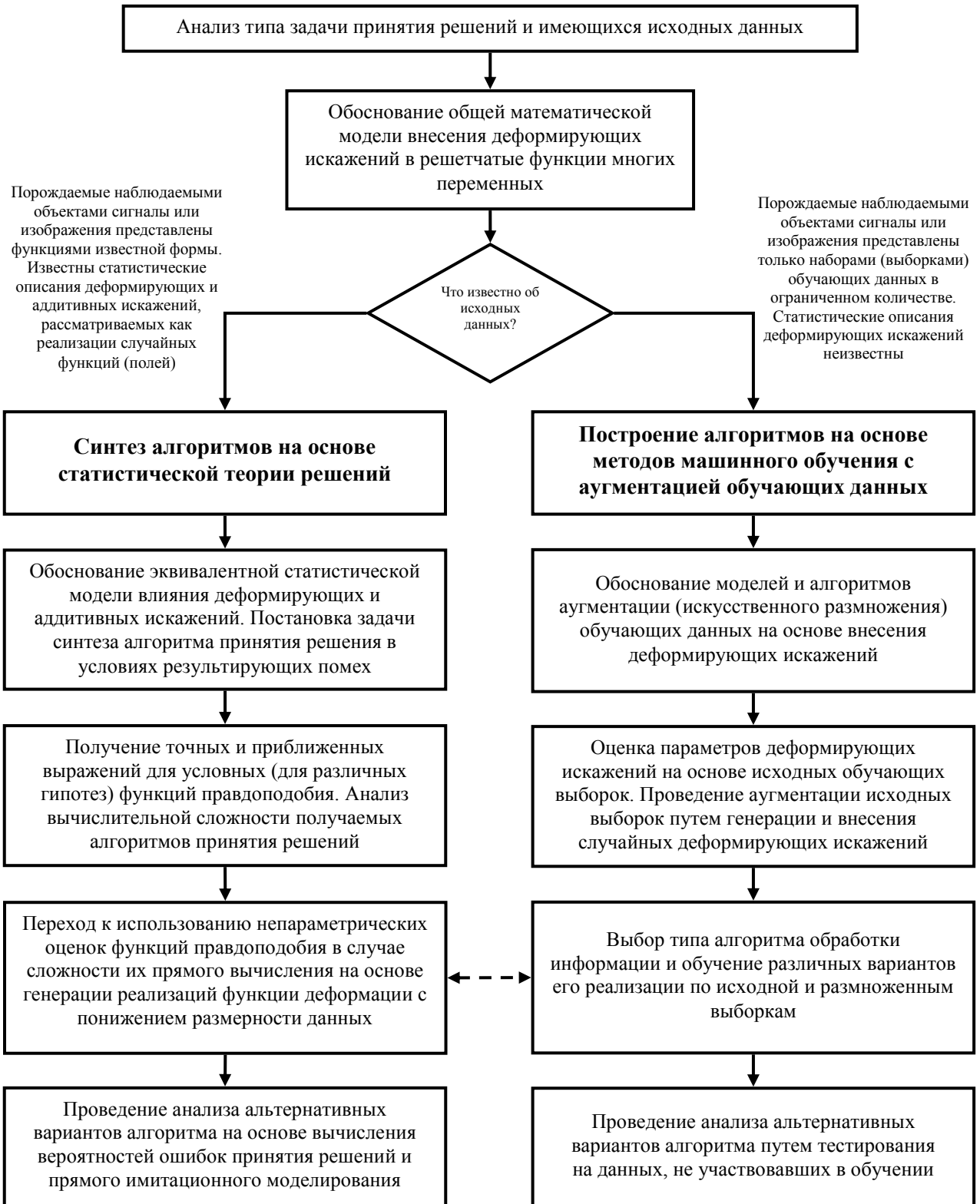


Рисунок 1.9 – Общая схема проведения исследований в интересах построения алгоритмов обработки информации для различных типов задач принятия решений в условиях деформирующих искажений

**При реализации первого подхода** в рамках предложенной схемы с учетом первичного анализа типа задачи принятия решений, имеющихся исходных данных и обзора существующих исследований по заданной тематике в диссертации была представлена и обоснована эквивалентная статистическая модель влияния деформирующих и аддитивных искажений. Она основана на представлении нелинейного оператора деформации как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. На базе этой модели была осуществлена постановка и выполнен синтез алгоритма принятия решения применительно к задаче распознавания объектов различных классов в условиях результирующих помех. В интересах проведения такого синтеза при реализации данного подхода были получены точные и приближенные выражения для условных (для различных гипотез) функций правдоподобия.

Далее был проведен анализ сложности получаемых алгоритмов принятия решений при использовании точных и приближенных выражений для функций правдоподобия. Он показал, что во многих ситуациях применение таких выражений для функций правдоподобия с увеличением размерности анализируемых данных (особенно в задаче обработки изображений) наталкивается на ощутимые вычислительные трудности. В связи с этим, при работе с изображениями был осуществлен переход к использованию непараметрических оценок функций правдоподобия путем генерации нужного числа реализаций функции деформации с понижением размерности данных на основе использования метода главных компонент. Данный подход во многом перекликается с методами, основанными на технологиях машинного обучения, хотя и рассчитанные оценки функций правдоподобия подставляются в решающие правила, имеющие такую же структуру, как и полученные в результате строгого синтеза.

Сравнительный анализ синтезированных альтернативных вариантов алгоритма был проведен на основе вычисления вероятностей ошибок принятия решений (при наличии такой возможности) или путем прямого имитационного компьютерного моделирования.

**При реализации второго подхода** в предложенной схеме в диссертации было осуществлено построение алгоритмов обработки информации применительно к задачам распознавания объектов различных классов, а также поиска и обнаружения объектов на изображениях на основе методов машинного обучения с аугментацией обучающих данных.

В рамках данного подхода потребовалось провести обоснование моделей и алгоритмов аугментации или размножения элементов обучающей выборки на основе внесения в них деформирующих искажений. Для каждого из таких алгоритмов была осуществлена оценка параметров деформирующих искажений на основе статистической обработки различий между элементами исходной обучающей выборки и проведена ее аугментация путем генерации и внесения случайных деформирующих искажений, сформированных в соответствии с оцененными параметрами.

Далее было проведено обучение детекторов фронтальных лиц по исходной и размноженной обучающей выборкам, а также их тестирование на новых данных, не участвовавших в процессе обучения. Это позволило провести оценку эффективности того или иного способа аугментации, а также выбрать предпочтительный вариант алгоритма обработки информации.

В качестве инструмента исследования был выбран алгоритм поиска и обнаружения лиц по методу Виолы-Джонса как один из самых надежных и известных на данный момент детекторов фронтальных лиц. Тем не менее, предложенный общий подход к решению задачи позволяет сравнивать и использовать сформированные обучающие выборки и в рамках любого другого

алгоритма поиска и обнаружения объектов на изображениях, основанного на использовании методов машинного обучения.

Следует также отметить взаимосвязь обоих подходов к решению задач данного типа, которая просматривается в предложенной схеме. Она определяется тем, что использование средств аугментации обучающей выборки допускает построение непараметрических смешанных оценок функций правдоподобия гипотез с переходом в дальнейшем к применению решающих правил, синтезированных в рамках статистической теории. Так в диссертации решалась задача обработки и распознавания реальных изображений биологических объектов (элементов зерновых смесей). Аналогично полученные непараметрические оценки могут быть использованы для обоснования алгоритмов аугментации экспериментальных данных, представленных обучающими выборками.

### **Выводы по главе**

Рассмотрены основные аспекты использования моделей деформирующих искажений в современных задачах обработки сигналов и изображений. При этом могут быть сделаны следующие выводы.

1. Анализ публикаций применительно к задаче сопоставления временных рядов позволяет сделать заключение, что учет возможных искажений порождающих их сигналов вдоль временной оси часто позволяет получить более корректные результаты. Сравнительные исследования показывают, что выбор конкретного подхода к учету влияния деформирующих искажений зависит от многих факторов, в том числе от конкретной задачи (кластеризация, индексирование или классификация), самих данных, а также от желаемого соотношения скорости и качества работы. При этом, в целом, многие из представленных методов и алгоритмов обеспечивают примерно равные показатели точности. Тем не менее, существующие публикации, посвященные работе с временными рядами, в

большинстве своем обходят стороной сам вопрос анализа влияния деформирующих искажений на процесс их распознавания и носят, как правило, сугубо эвристический прикладной характер.

2. Рассмотрены работы, посвященные исправлению оптических искажений (дисторсии), возникающих при съемке основанными на использовании линз приборами. В них отмечается, что отличительной чертой возникающих таким образом деформирующих искажений является известность моделей этих искажений, что основано на самой природе их проявления, связанной с особенностями построения данных приборов.

3. Проведен сравнительный анализ различных методов оценки оптического потока, предназначенных для анализа изменений в положениях и ориентации объектов сцены в разные моменты времени при их перемещениях или при смене ракурса съемки. Каждый из представленных подходов в большей или меньшей степени нацелен на определенную область применимости, в которой показывает себя лучше, чем другие, из-за заложенных в него ограничений и допущений, на основе которых и производится его расчет. Особый интерес при этом среди них представляет подгруппа параметрических методов расчета оптического потока, базирующихся на использовании примитивов, так как они позволяют оценивать перемещения и деформации схожих по смыслу элементов сравниваемых изображений, которые далеко не всегда близки в своем исходном пиксельном представлении.

4. Описана процедура морфинга — процесса постепенного замещения одного изображения другим путем плавных изменений значений интенсивностей пикселей, положений и формы схожих в некоем смысле частей этих изображений — и разные подходы к ее осуществлению. Анализ представленных публикаций, относящихся к данной области, показал, что существует большое количество различных методов осуществления морфинга, каждый из которых показывает себя лучше или хуже в

зависимости от конкретной задачи и природы исходных изображений, и предполагающих разную степень ручного вмешательства оператора.

5. Проанализированы известные подходы к распознаванию изображений, основанные на учете деформирующих искажений. Их можно разделить на две большие группы: эластичные модели распознавания образов, сравнивающие изображения не только по содержимому, но и по близости расположения их отдельных частей к определенным позициям на изображении или относительно друг друга, и статистические модели, анализирующие близость изображений на основе нахождения наиболее вероятной деформации, необходимой для преобразования их друг в друга. Следует отметить вычислительную сложность используемых при этом моделей преобразования изображений. В целом, общим для всех методов и алгоритмов распознавания изображений, основанных на учете возможных деформирующих искажений, является то, что они нацелены на имитацию свойств изменчивости конкретных объектов реального мира, что позволяет достичь лучших показателей распознавания, а также в отдельных случаях понизить требуемый объем обучающих выборок. При этом известные алгоритмы в недостаточной степени учитывают статистический характер деформирующих искажений как реализаций случайных полей и не всегда позволяют получить оценки вероятностных характеристик распознавания объектов различных классов.

6. Проанализированы современные технологии искусственного размножения обучающих выборок на основе внесения в них деформирующих искажений. Результаты экспериментов в данной области говорят о том, что размножение обучающих выборок таким путем во многих случаях увеличивает точность работы алгоритмов распознавания. Особо отмечается, что наилучших результатов можно достичь в случае наличия информации о параметрах возможных деформирующих искажений, описывающих различия между элементами обучающей выборки. Таким образом, на данный момент следует отметить актуальность и перспективность исследований, посвященных нахождению параметров таких искажений, для чего



представляется полезным использование соответствующих подходов и разработок из областей морфинга и расчета оптического потока.

7. Проведенный анализ также показывает, что строгое решение задачи синтеза и анализа алгоритмов распознавания изображений с учетом влияния деформирующих искажений в известных источниках не получила: предложенные в них алгоритмические решения носят преимущественно эвристический характер. В связи с этим представляются важными постановка и решение задач распознавания объектов, представленных цифровыми сигналами и изображениями случайной формы в условиях деформирующих искажений, как в рамках статистической теории решений, так и на основе методов машинного обучения. В соответствии с данными соображениями представлена и описана общая схема проведения исследований в интересах построения алгоритмов обработки информации для различных задач принятия решений с использованием моделей деформирующих искажений.

## 2 Модели и алгоритмы внесения деформирующих искажений в сигналы и изображения

### 2.1 Математическая модель внесения деформирующих искажений

Как уже отмечалось ранее, математическое представление процесса внесения деформирующих искажений (ДИ) для любой исходной функции  $f(x_1, \dots, x_n)$  от  $n$  переменных определяется через аддитивное добавление значения функции деформации к исходным значениям этих переменных:

$$\begin{aligned} f(x_1 + r_1(x_1, \dots, x_n), \dots, x_n + r_n(x_1, \dots, x_n)) = \\ = f(u_1(x_1, \dots, x_n), \dots, u_n(x_1, \dots, x_n)) = g(x_1, \dots, x_n), \end{aligned} \quad (2.1)$$

где  $g(x_1, \dots, x_n)$  — результирующая деформированная функция;  $r_i(x_1, \dots, x_n)$ ,  $i = \overline{1, n}$  — непрерывные функции вносимых деформирующих искажений по каждой координате, которые могут носить детерминированный или стохастический характер.

Используя ранее введенные обозначения, а также, векторы:  $\mathbf{x} = (x_1, \dots, x_n)^T$ ,  $\mathbf{r}(\mathbf{x}) = (r_1(\mathbf{x}), \dots, r_n(\mathbf{x}))^T$  и  $\mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), \dots, u_n(\mathbf{x}))^T$ , можно записать ДИ в векторном виде

$$f(\mathbf{x} + \mathbf{r}(\mathbf{x})) = f(\mathbf{u}(\mathbf{x})) = g(\mathbf{x}). \quad (2.2)$$

Ранее, на рисунке 1.1, был изображен пример деформации одномерной непрерывной функции вида  $f(x) = \cos(x^2 / 100)$ , иллюстрирующий общий характер возникающих искажений. Внесение ДИ в рамках рассмотренной модели может также производиться и обратным образом:

$$f(\mathbf{x}_1) = g(\mathbf{u}^{-1}(\mathbf{x}_1)) = g(\mathbf{x}_1 + \mathbf{r}^{-1}(\mathbf{x}_1)),$$

где  $\mathbf{r}^{-1}(\mathbf{x}_1)$  — функция деформации, обратная к  $\mathbf{r}(\mathbf{x}_2)$ , а  $\mathbf{x}_1$  и  $\mathbf{x}_2$  такие, что

$$f(\mathbf{x}_2 + \mathbf{r}(\mathbf{x}_2)) = f(\mathbf{u}(\mathbf{x}_2)) = g(\mathbf{x}_2),$$

$$\mathbf{x}_1 = \mathbf{x}_2 + \mathbf{r}(\mathbf{x}_2), \quad \mathbf{x}_2 = \mathbf{x}_1 + \mathbf{r}^{-1}(\mathbf{x}_1).$$

При реализации процесса внесения ДИ на основе (2.1) и (2.2) возникает ряд особенностей [62-65]. Первая из них заключается в том, что для того, чтобы

значение деформированной функции  $g(\mathbf{x}) = f(\mathbf{u}(\mathbf{x}))$  существовало, область ее определения не должна выходить за пределы области определения исходной функции  $f(\mathbf{x})$ . Иначе, если для  $f(\mathbf{x})$  вектор  $\mathbf{x} \in \Omega_x$ , то и вектор  $\mathbf{u}(\mathbf{x}) \in \Omega_x$ . Для обеспечения выполнения условия  $\mathbf{u}(\mathbf{x}) \in \Omega_x$  применимы различные подходы.

Один из них состоит в том, чтобы при выходе какой-либо компоненты координат в  $\mathbf{x} + \mathbf{r}(\mathbf{x}) = \mathbf{u}(\mathbf{x})$  из (2.2) за границы исходной области определения  $\Omega_x$  в качестве значения данной компоненты  $\mathbf{u}(\mathbf{x})$  при расчете  $f(\mathbf{u}(\mathbf{x}))$  подставлять граничное значение соответствующей координаты исходной функции  $f(\mathbf{x})$ . Преимущество данного подхода состоит в том, что при его применении не модифицируются значения функции деформации  $\mathbf{r}(\mathbf{x})$  в случае, когда  $\mathbf{x} + \mathbf{r}(\mathbf{x}) = \mathbf{u}(\mathbf{x})$  находится в пределах  $\Omega_x$ .

Другой подход [64] состоит в использовании масочных функций  $\mathbf{h}(\mathbf{x}) = (h(x_1), \dots, h(x_n))^T$ , ограничивающих краевые значения  $\mathbf{u}(\mathbf{x})$ :

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} + \text{diag}[\mathbf{h}(\mathbf{x})]\mathbf{r}(\mathbf{x}),$$

где  $\text{diag}()$  — обозначение диагональной матрицы. Эти функции должны стремиться к нулю на границах области определения  $\Omega_x$ . В качестве них могут быть использованы, например, или трапециевидные оконные функции, или гауссианы вида (представлено для одномерного случая):

$$h(x) = \frac{1}{(2\pi)^{1/2} \xi} \exp\left[-\frac{(x - \mu)^2}{2\xi^2}\right]$$

где  $\mu$  и  $\xi$  — некие вещественные числа, которые всегда можно выбрать так, чтобы обеспечить выполнение указанных условий.

## 2.2 Внесение деформирующих искажений в решетчатые функции дискретных аргументов

В случае задания исходной функции как функции дискретных аргументов (как решетчатой функции) главная особенность [64, 71] процесса внесения ДИ на основе (2.1) и (2.2) состоит в следующем. Если исходная функция  $\hat{f}(\mathbf{x})$  является решетчатой функцией дискретных аргументов, т.е. задана на многомерной дискретной сетке (например, в случае внесения ДИ в цифровые изображения), прямое применение (2.1) и (2.2) невозможно. В этом случае для обеспечения возможности модификации аргументов  $\hat{f}(\mathbf{x})$  при внесении в них ДИ произвольного характера необходимо предварительно выполнить ее интерполяцию и представление в виде функции  $\tilde{f}(\mathbf{x})$  вещественных переменных. При этом общая процедура будет соответствовать схеме, предложенной автором в работе [64] и представленной на рисунке 2.1.

Следует также отметить, что в данной схеме для исходной функции  $\hat{f}(\mathbf{x})$  не учитывается возможный дискретный характер по уровню, т.е. считается, что разрядность представления данных велика и ею можно пренебречь, считая, что значения функции имеют вещественные значения. В некоторых случаях это не так, например, если функция описывает полутонные изображения с равномерным квантованием на 256 уровней. Тогда описанные в схеме преобразования осуществляются по отношению к исходной квантованной по уровню функции, и с учетом того, что при деформации возможно получение промежуточных значений между исходными уровнями, проводится дополнительное квантование деформированной функции и ее приведение к стандартной сетке значений.

**Утверждение 2.1.** Пусть  $f(\mathbf{x})$  — функция, заданная на множестве значений  $n$  непрерывных аргументов  $\Omega_x$ ;  $\hat{f}(\mathbf{x})$  — решетчатая функция, сформированная на многомерной дискретной сетке, покрываемой  $\Omega_x$ ;  $\tilde{f}(\mathbf{x})$  — функция,

интерполированная на основе значений  $\hat{f}(\mathbf{x})$  и определенная на  $\tilde{\Omega}_x \subset \Omega_x$ ;  $g(\mathbf{x})$  — деформированная функция на основе внесения ДИ  $\mathbf{r}(\mathbf{x})$  в  $f(\mathbf{x})$  по формуле (2.1);  $\tilde{g}(\mathbf{x})$  — деформированная аналогичным образом функция на основе внесения ДИ  $\mathbf{r}(\mathbf{x})$  в  $\tilde{f}(\mathbf{x})$ ;  $X$  — некое связное подмножество значений непрерывных аргументов  $\mathbf{x} \in X \subset \tilde{\Omega}_x$ . Пусть также для любого  $\mathbf{x} \in X$  выполняется  $\mathbf{u}(\mathbf{x}) \in X$ .

Тогда ошибка  $|g(\mathbf{x}) - \tilde{g}(\mathbf{x})|$ , возникающая при внесении деформации в интерполированную функцию  $\tilde{f}(\mathbf{x})$ , для  $\mathbf{x} \in X \subset \tilde{\Omega}_x$  не превысит максимальную ошибку интерполяции  $|f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$ .

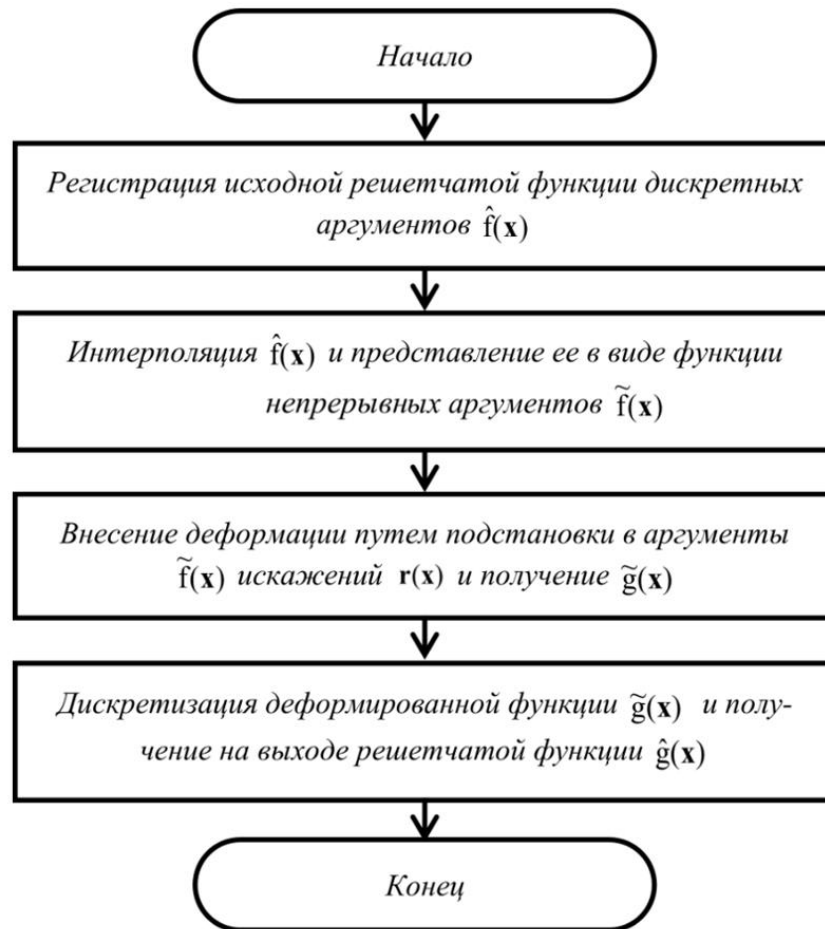


Рисунок 2.1 – Обобщенная схема процесса внесения деформирующих искажений в решетчатую функцию

**Доказательство.** Исходя из сформулированных условий, имеем, что  $g(\mathbf{x}) = f(\mathbf{u}(\mathbf{x}))$  и  $\tilde{g}(\mathbf{x}) = \tilde{f}(\mathbf{u}(\mathbf{x}))$ . Вычтем эти два выражения друг из друга и возьмем абсолютные значения разностей в левой и правой части:

$$\Delta(\mathbf{x}) = |g(\mathbf{x}) - \tilde{g}(\mathbf{x})| = |f(\mathbf{u}(\mathbf{x})) - \tilde{f}(\mathbf{u}(\mathbf{x}))|.$$

Пусть  $U$  — множество всех значений, принимаемых  $\mathbf{u}(\mathbf{x})$  внутри  $X$ , т.е.  $\mathbf{u}(\mathbf{x}) \in U \subset X$ . Таким образом, функции  $g(\mathbf{x}) = f(\mathbf{u}(\mathbf{x}))$  и  $\Delta(\mathbf{x})$  определены на множестве  $U$ . При этом функция  $|f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$  определена на множестве  $X$ . Возможны две ситуации.

Первая из них состоит в том, что функция искажения  $\mathbf{r}(\mathbf{x})$  может оказаться такой, что для некоторого аргумента  $\mathbf{x}^{(j)} \in X$  не найдется такого значения из  $U$ , что  $\mathbf{u}(\mathbf{x}) = \mathbf{x} + \mathbf{r}(\mathbf{x}) = \mathbf{x}^{(j)}$  (данный элемент не будет использован при деформации). Действительно, достаточно, например, взять  $\mathbf{r}(\mathbf{x}) = -\mathbf{x} + \mathbf{x}^{(k)}$ , для некоторого  $\mathbf{x}^{(k)} \in X$  такого, что  $\mathbf{x}^{(k)} \neq \mathbf{x}^{(j)}$ . Тогда

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} + \mathbf{r}(\mathbf{x}) = \mathbf{x} - \mathbf{x} + \mathbf{x}^{(k)} = \mathbf{x}^{(k)} \neq \mathbf{x}^{(j)}.$$

В этом случае множество  $X \setminus U$  не пустое. Среди его элементов могут оказаться такие, что  $|f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$  будет принимать свое максимальное в  $X$  значение. Тогда  $\max |g(\mathbf{x}) - \tilde{g}(\mathbf{x})| < \max |f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$ , иначе  $\max |g(\mathbf{x}) - \tilde{g}(\mathbf{x})| = \max |f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$ .

Во втором случае, если для данного искажения  $\mathbf{r}(\mathbf{x})$  для любого  $\mathbf{x}^{(j)} \in X$  найдется  $\mathbf{x} \in X$  такое, что  $\mathbf{x}^{(j)} = \mathbf{x} + \mathbf{r}(\mathbf{x})$ , то  $U = X$ , так как каждому  $\mathbf{x} \in X$  соответствует свое уникальное значение  $\mathbf{u}(\mathbf{x}) \in U$ , при том, что по определению  $U \subset X$ , и в этом случае  $\max |g(\mathbf{x}) - \tilde{g}(\mathbf{x})| = \max |f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$ .

Аналогичным образом можно доказать, что

$$\min |g(\mathbf{x}) - \tilde{g}(\mathbf{x})| \geq \min |f(\mathbf{x}) - \tilde{f}(\mathbf{x})|.$$

**Следствие.** Ошибка  $|g(\mathbf{x}) - \tilde{g}(\mathbf{x})|$ , возникающая при деформации интерполированной функции  $\tilde{f}(\mathbf{x})$  с областью определения  $\tilde{\Omega}_x$ , не выйдет за пределы максимального и минимального значения ошибки интерполяции  $|f(\mathbf{x}) - \tilde{f}(\mathbf{x})|$  при условии, что  $\mathbf{u}(\mathbf{x}) \in \tilde{\Omega}_x$ . Для подтверждения этого следствия достаточно повторить предыдущее доказательство в случае  $X = \tilde{\Omega}_x$ .

Таким образом, при внесении ДИ достаточно выполнить  $\mathbf{u}(\mathbf{x}) \in \tilde{\Omega}_x$ , что производится при помощи того же подхода, что и при обеспечении условия  $\mathbf{u}(\mathbf{x}) \in \Omega_x$ . При этом  $\mathbf{u}(\mathbf{x}) \in \tilde{\Omega}_x$  автоматически означает, что  $\mathbf{u}(\mathbf{x}) \in \Omega_x$ , так как  $\tilde{\Omega}_x \subset \Omega_x$ .

### 2.2.1 Интерполяция с помощью радиально-базисных функций

В соответствии с вышеописанными соображениями в интересах реализации алгоритма внесения ДИ в решетчатые функции дискретных аргументов было проведено сравнительное исследование различных способов интерполяции, в том числе двух широко известных и часто применяемых на практике: линейной и кубической интерполяции [73], специфика применения которых для решаемой задачи приведена автором в [64]. Кроме того, для решения этой задачи была исследована возможность применения радиально-базисных функций (РБФ). Данный подход потенциально позволяет вносить непрерывные искажения  $\mathbf{r}(\mathbf{x})$  любого вида. Результаты опубликованы автором в работах [62-64].

В одномерном случае при его реализации исходная решетчатая функция  $\hat{f}(\hat{x}_i)$ ,  $i = \overline{1, N}$  подразделяется на отрезки, каждый длиной  $n$ . При этом крайние точки соседствующих друг с другом отрезков являются для них общими. Внутри отрезков используются локальные координаты. После этого для каждого из отрезков ищется аппроксимация непрерывной функцией вида

$$\tilde{f}(x) = \sum_{i=1}^n a_i \exp(-(x - \hat{x}_i)^2 / (2\sigma^2)), \quad (2.3)$$

где  $\sigma$  — параметр влияния РБФ-функций, а  $a_i$  — неизвестные коэффициенты. Элементом суммы соответствуют центры РБФ-функций в точках  $\hat{x}_i$  — по одной функции на точку исходного дискретного набора отсчетов внутри данного отрезка.

Если для  $x$ , соответствующих позициям известных точек исходного интерполируемого отрезка, вместо  $\tilde{f}(x)$  подставлять в (2.3) значения  $\hat{f}(x)$ , то для каждого отрезка получится система из  $n$  линейных уравнений (число неизвестных  $a_i$  также равно  $n$ ):

$$\sum_{i=1}^n \exp(-(\hat{x}_j - \hat{x}_i)^2 / (2\sigma^2)) a_i = \hat{f}(\hat{x}_j), \quad j = \overline{1, n},$$

или в матричном виде:  $\mathbf{E}\mathbf{A} = \mathbf{F}$ , где матрица  $\mathbf{E}$  состоит из элементов  $\exp(-(\hat{x}_j - \hat{x}_i)^2 / (2\sigma^2))$  со столбцами, пронумерованными  $i = \overline{1, n}$ , и строками —  $j = \overline{1, n}$ , вектор  $\mathbf{A} = (a_1, \dots, a_n)^T$  и вектор  $\mathbf{F} = (\hat{f}(\hat{x}_1), \dots, \hat{f}(\hat{x}_n))^T$ .

Длина отрезка и параметр влияния РБФ-функций  $\sigma$  подбираются экспериментальным путем. Получаемые таким образом системы уравнений не всегда могут иметь стандартное решение  $\mathbf{A} = \mathbf{E}^{-1}\mathbf{F}$  из-за плохой обусловленности матрицы  $\mathbf{E}$ . В этих случаях применяют методы регуляризации. Например, вместо  $\mathbf{E}^{-1}$  используют псевдообратную матрицу  $\mathbf{E}^+$  (алгоритм сингулярного разложения SVD) или регуляризацию по А.Н. Тихонову [74]. В последнем случае решение имеет вид

$$\mathbf{A} = (\lambda \mathbf{I} + \mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T (\mathbf{F} - \mathbf{E} \mathbf{A}_0),$$

где  $\mathbf{I}$  — единичная матрица размером  $n \times n$ ;  $\lambda$  — параметр регуляризации;  $\mathbf{A}_0$  — априорное решение. При регуляризации, чем больше  $\lambda$ , тем лучше обусловленность и ближе решение к априорной оценке  $\mathbf{A}_0$ , но дальше от решения исходной некорректной задачи. Исходя из этих соображений, параметр  $\lambda$  должен принимать промежуточные значения, определяемые эмпирическим путем или при помощи



некоторого анализа. Если в качестве  $\mathbf{A}_0$  использовать нулевую матрицу, то регуляризация будет проводиться без учета априорного решения. Для нахождения разумного значения  $\lambda$  применяются различные методы, из которых наиболее часто применяемым является метод невязки с оценкой уровня погрешности при получении исходных данных, которая, собственно, и определяет неустойчивость решения.

Для функций двух аргументов процесс во многом аналогичен одномерному случаю. Исходная матрица значений решетчатой функции  $\hat{f}(\hat{x}_i, \hat{y}_j)$  при  $i = \overline{1, M}$  и  $j = \overline{1, N}$  подразделяется на блоки размерами  $m \times n$ . При этом крайние точки соседствующих друг с другом блоков являются для них общими, и используются локальные координаты вдоль каждой из осей. После чего для каждого из блоков ищется аппроксимация непрерывной функцией вида

$$\tilde{f}(x, y) = \sum_{i=1}^n \sum_{j=1}^m a_{i,j} \exp(-((x - \hat{x}_i)^2 + (y - \hat{y}_j)^2) / (2\sigma^2)). \quad (2.4)$$

Элементом суммы соответствуют центры РБФ-функций с координатами  $\hat{x}_i$  и  $\hat{y}_j$  (по одной функции на каждую точку текущего блока исходной матрицы). Если для  $x$  и  $y$ , соответствующих координатам точек в блоке, вместо  $\tilde{f}(x, y)$  подставлять в (2.4) соответствующие этим точкам значения  $\hat{f}(x, y)$ , то для каждого блока получится система из  $m \times n$  линейных уравнений (число неизвестных также равно  $m \times n$ ):

$$\sum_{i=1}^m \sum_{j=1}^n \exp(-((\hat{x}_k - \hat{x}_i)^2 + (\hat{y}_l - \hat{y}_j)^2) / (2\sigma^2)) a_{i,j} = \hat{f}(\hat{x}_k, \hat{y}_l), \quad k = \overline{1, m}, \quad l = \overline{1, n}.$$

В матричном виде получим уравнение  $\mathbf{E}\mathbf{A} = \mathbf{F}$ , где матрица  $\mathbf{E}$  состоит из элементов  $\exp(-((\hat{x}_k - \hat{x}_i)^2 + (\hat{y}_l - \hat{y}_j)^2) / (2\sigma^2))$  со столбцами, пронумерованными комбинациями значений индексов  $i = \overline{1, m}$  и  $j = \overline{1, n}$ , и строками, пронумерованными комбинациями  $k = \overline{1, m}$  и  $l = \overline{1, n}$ , вектор  $\mathbf{A} = (a_{11}, \dots, a_{mn})^T$  и вектор  $\mathbf{F} = (\hat{f}(\hat{x}_1, \hat{y}_1), \dots, \hat{f}(\hat{x}_m, \hat{y}_n))^T$  содержит значения исходной решетчатой функции.

Аналогичным образом можно выполнить интерполяцию решетчатых функций с любым числом аргументов.

### **2.2.2 Экспериментальные исследования точности разных видов интерполяции при внесении деформирующих искажений**

Выбор длины отрезка  $n$  для одномерного случая, размеров блоков  $m \times n$  для двумерного случая, параметра влияния РБФ-функций  $\sigma$  и параметра регуляризации  $\lambda$  (в случае ее использования) осуществлялся путем оценки невязки интерполированной функции  $\tilde{f}(x)$  и исходной функции  $f(x)$  —  $|f(x) - \tilde{f}(x)|$  для разных значений перечисленных параметров. С этой целью использовались непрерывные функции  $f(x)$ , для которых проводилась дискретизация и по полученным результатам осуществлялось построение интерполяции  $\tilde{f}(x)$ .

На рисунке 2.2 изображена зависимость усредненной невязки (в данном случае средней модульной ошибки интерполирования) от  $\lambda$  при фиксированных  $n=2$  и  $\sigma=2$  для функции  $f(x) = \cos(x^2 / w)$  в области от 1 до 100 с шагом интерполирования 1 и шагом проверки величиной в 0,1 при постепенном изменении  $w$  от 1 до 200.

В ходе тестирования было выявлено, что для функций одного аргумента хорошего приближения можно достичь, используя размер блока  $n=25$  и параметр влияния РБФ-функций  $\sigma=1$ . Эксперименты показали, что псевдообращение и регуляризация (при  $\lambda = 10^{-16}$ ) обеспечивают примерно одинаковую точность интерполяции, но при этом использование регуляризации требует меньших временных затрат.

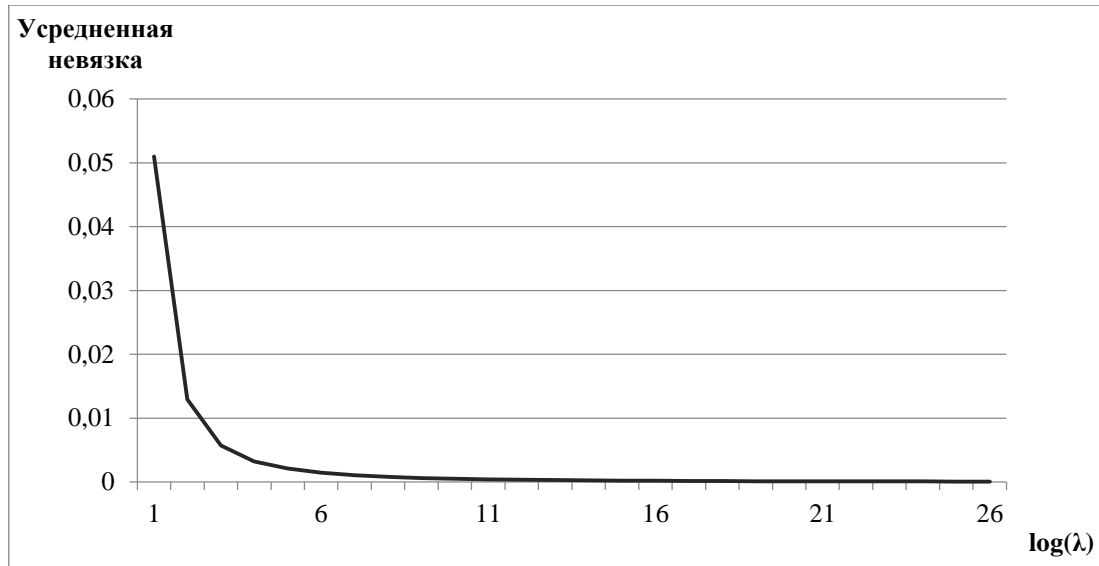


Рисунок 2.2 – Зависимость усредненной невязки от  $\lambda$  при фиксированных  $n = 2$  и  $\sigma = 2$  для функции  $f(x) = \cos(x^2 / w)$  в области от 1 до 100 с шагом интерполирования 1 и шагом проверки величиной в 0.1 при постепенном изменении  $w$  от 1 до 200

Тестирование точности работы разных способов интерполяции показало, что в зависимости от типа и частоты интерполируемой функции лидируют разные способы интерполяции. Так, сравнение точности интерполяции для функции  $f(x) = \cos(x^2 / w)$  в области от 1 до 100 с шагом интерполирования 1 и шагом проверки величиной в 0,1 при постепенном изменении  $w$  от 1 до 200 показывает зависимости, представленные на рисунке 2.3.

Также были произведены замеры скорости работы алгоритмов деформации, использующих разные способы интерполяции, при внесении ДИ в функцию  $f(x) = \cos(x^2 / w)$  (таблица 2.1). Тестирование производилось для разных границ области определения. При проведении экспериментов в качестве испытательного стенда использовался персональный компьютер со следующими характеристиками: процессор Intel Core i5-3210M 2.5 ГГц; оперативная память 6 ГБ DDR3; ОС — 64-х битная версия Windows 8.1; среда разработки Matlab R2013b.

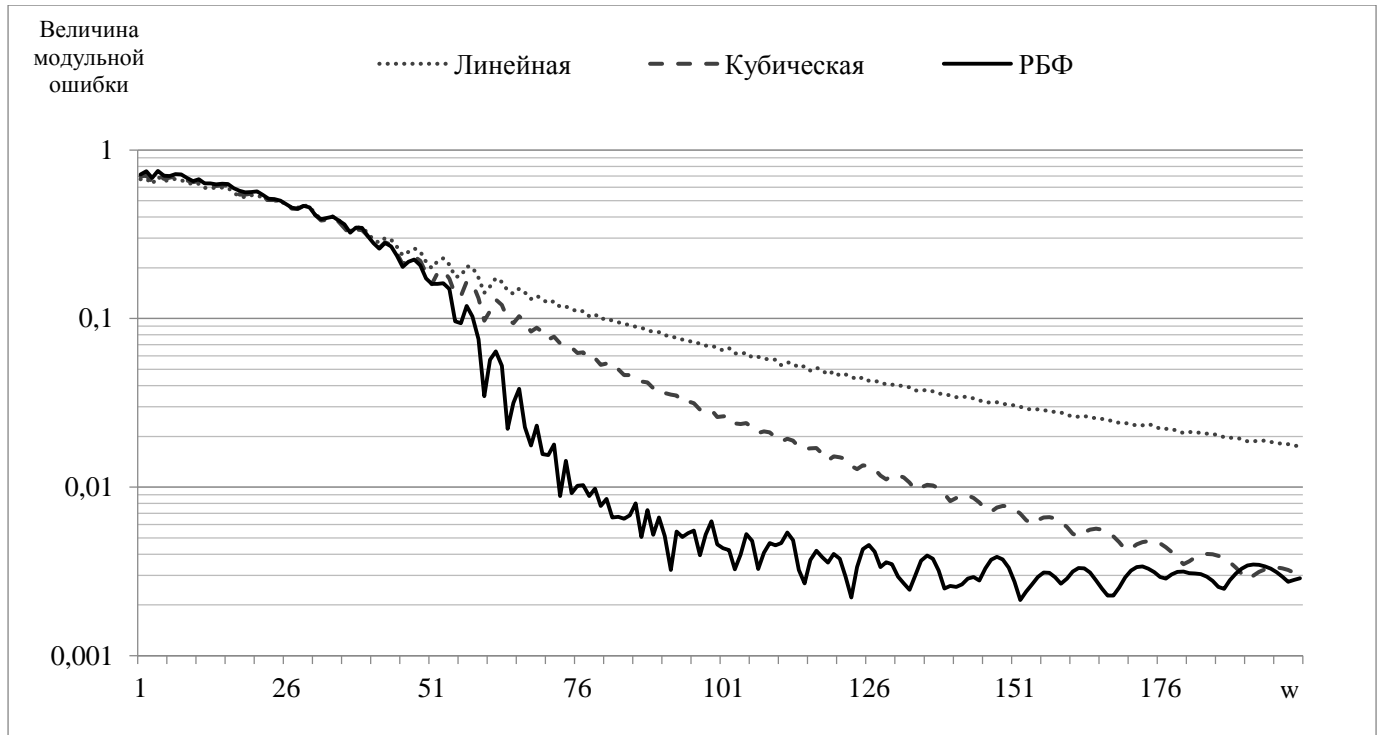


Рисунок 2.3 – Сравнение точности интерполяции для функции  $f(x) = \cos(x^2 / w)$  в области от 1 до 100 с шагом интерполирования 1 и шагом проверки величиной в 0,1 при постепенном изменении  $w$  от 1 до 200

Таблица 2.1 – Сравнение быстродействия различных способов интерполяции (в секундах)

Границы области определения	Билинейная	Бикубическая	РБФ	
			Расчет псевдообратной матрицы	Регуляризация по Тихонову
От 1 до 50	0.0018	0.0024	0.0039	0.0033
От 1 до 100	0.0033	0.0044	0.0070	0.0059
От 1 до 500	0.0154	0.0211	0.0328	0.0278

Для двумерного случая в ходе тестирования было выявлено, что достаточно хорошего приближения можно достичь, если использовать параметр влияния

РБФ-функций  $\sigma = 2$ , размер блока  $4 \times 4$  и использовать псевдообратную матрицу при решении получаемых систем линейных уравнений. Эксперименты также показали, что при «сшивке» на границах соседних блоков применение усреднения значений полученной для них интерполяции не улучшает результат.

На рисунке 2.4 показаны примеры результатов внесения деформирующих искажений в цифровое изображение с использованием интерполяции на основе RBF-функций и периодической функции ДИ.

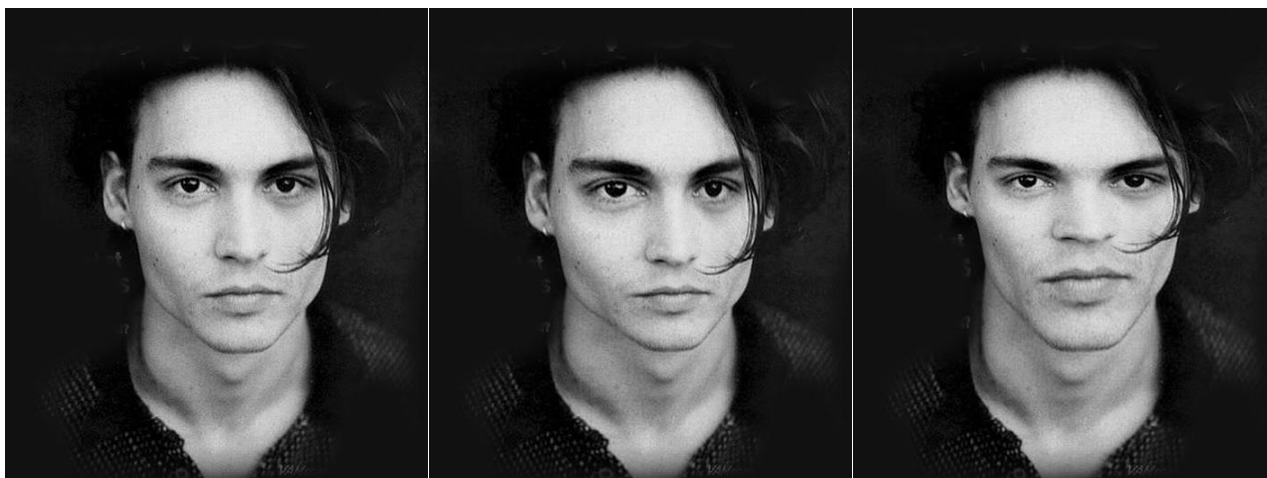


Рисунок 2.4 – Примеры результатов внесения деформирующих искажений с применением основанной на РБФ-функциях интерполяции и периодической функции искажения (изображение слева является исходным)

Было произведено сравнение быстродействия основанных на разных способах интерполяции вариантов внесения ДИ в изображения (таблица 2.2), и была осуществлена оценка точности разных способов интерполяции (таблица 2.3). Тестирование точности производилось в приложении к задаче изменения размера изображения. При этом осуществлялось повторное увеличение его предварительно уменьшенной копии ( $32 \times 24$  пикселя) и сравнение полученного результата с оригиналом ( $320 \times 240$  пикселей) для каждого из способов интерполяции.

Таблица 2.2 – Сравнение быстродействия различных способов интерполяции при деформации реальных изображений (в секундах)

Размер черно-белого изображения (в пикселях)	Билинейная	Бикубическая	РБФ
160x120	0,85	1,60	1,90
320x240	9,22	17,67	20,92
640x480	27,44	54,00	60,20

Таблица 2.3 – Сравнение точности разных способов интерполяции в приложении к задаче изменения размера изображения

Тип интерполяции	Билинейная	Бикубическая	РБФ
Величина среднемодульной ошибки для 8-ми битного кодирования черно-белого изображения	14,64	13,90	14,08

Также было осуществлено сравнение точности работы разных способов интерполяции для функций двух аргументов. Результаты показывают, что аналогично одномерному случаю здесь также в зависимости от типа и частоты интерполируемой функции лидируют разные способы интерполяции. Так, для  $f(x, y) = \cos((x^2 + y^2)/w)$  в области значений от 1 до 50 по обеим координатам с шагом интерполирования 1 и шагом проверки величиной в 0,1 при постепенном изменении  $w$  от 1 до 200 получаются зависимости, представленные на рисунке 2.5.

Исходя из представленных результатов, можно сделать вывод, что использование предложенного подхода внесения деформации в изображения на основе радиально-базисных функций показывает сравнимые результаты с другими методами интерполяции, а в ряде случаев превосходит их по точности и, таким образом, является перспективным для использования в процессе внесения ДИ при

размножении обучающих выборок в ходе решения различных задач распознавания изображений.

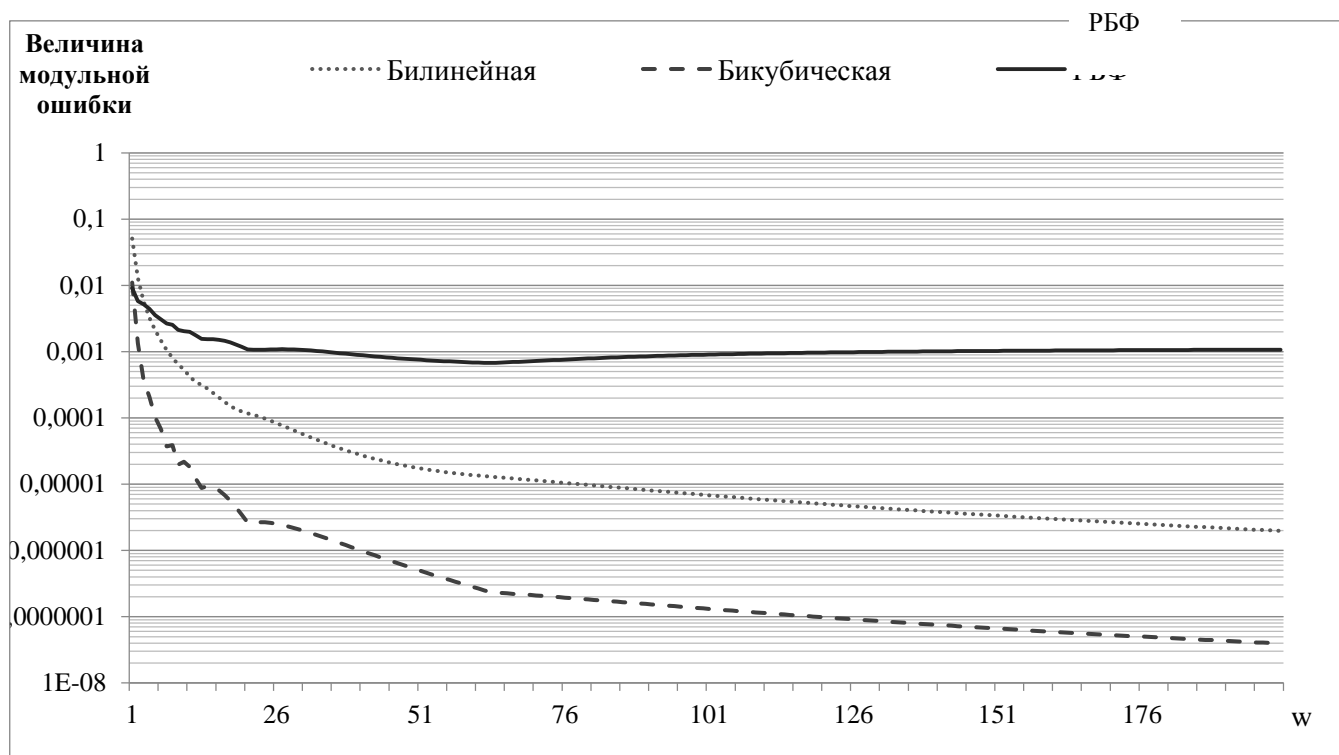


Рисунок 2.5 – Сравнение точности интерполяции для функции

$f(x, y) = \cos((x^2 + y^2) / w)$  в области от 1 до 50 по обеим координатам с шагом интерполирования 1 и шагом проверки величиной в 0,1 при постепенном изменении  $w$  от 1 до 200

### 2.3 Эквивалентная статистическая модель влияния деформирующих искажений, основанная на приближенном представлении нелинейного оператора деформации как оператора перестановки

С учетом потребностей решений задач синтеза и анализа алгоритмов распознавания цифровых сигналов и изображений автором в работе [65] была предложена эквивалентная статистическая модель влияния ДИ, основанная на приближенном представлении нелинейного оператора деформации как оператора размещения с повторениями элементов исходного дискретного сигнала с

добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации.

Рассмотрим без существенного ограничения общности следующую постановку задачи. Пусть на интервале времени  $\Xi_t = [0, T]$  задан сигнал  $S(t)$ ,  $t \in \Xi_t$ . Наблюдению доступна реализация случайной функции, формируемая на основе этого сигнала при воздействии ДИ

$$\tilde{S}(t) = S[t + A(t)], \quad t \in \Xi_t,$$

где  $\tilde{S}(t)$  — реализация случайной функции, описывающей деформированный полезный сигнал, являющаяся дифференцируемой детерминированной функцией;  $A(t)$  — однозначная дифференцируемая функция, являющаяся в общем случае реализацией случайной функции, описывающей ДИ с известными статистическими характеристиками, такая, что  $t + A(t) \in \Xi_t$ .

При цифровой обработке в результате дискретизации по времени с интервалом  $\Delta t$  формируется цифровой эквивалент наблюдаемого в непрерывном времени сигнала путем фиксации его значений в моменты времени  $t_k = (k-1)\Delta t$ ,  $k = \overline{1, n}$ ,  $\Delta t = T / (n-1)$ ,  $t_1 = 0$ ,  $t_n = T$ :

$$\tilde{S}(t_k) = S[t_k + A(t_k)]. \quad (2.5)$$

Введем также следующие векторные представления данных

$$\mathbf{s} = (s_1, \dots, s_n)^T = [S(t_1), \dots, S(t_n)]^T, \quad s_k = s[k] = S[(k-1)\Delta t], \quad k = \overline{1, n},$$

$$\tilde{\mathbf{s}} = (\tilde{s}_1, \dots, \tilde{s}_n)^T = [\tilde{S}(t_1), \dots, \tilde{S}(t_n)]^T, \quad \tilde{s}_k = \tilde{s}[k] = \tilde{S}[(k-1)\Delta t], \quad k = \overline{1, n},$$

$$\mathbf{a} = (a_1, \dots, a_n)^T = [A(t_1), \dots, A(t_n)]^T, \quad a_k = a[k] = A[(k-1)\Delta t], \quad k = \overline{1, n}.$$

Для проведения дальнейших преобразований рассмотрим исходное представление процесса внесения ДИ. Введем целочисленный индекс  $r_k = \text{round}(a_k / \Delta t)$ ,  $k = \overline{1, n}$ , где  $\text{round}()$  обозначает операцию округления. Тогда компоненты вектор-функции ДИ можно представить как



$$a_k = A(t_k) = \Delta t r_k + \varepsilon_k, \quad \varepsilon_k = A(t_k) - \Delta t r_k, \quad -\Delta t / 2 < \varepsilon_k < \Delta t / 2. \quad (2.6)$$

Учитывая в общем случае случайный характер функции  $A(t)$ , можно считать, что величина  $\varepsilon_k$  добавки к  $\Delta t r_k$  в (2.6) также является случайной величиной, характеризующей погрешность квантования по уровню заданной в непрерывном времени функции деформации  $A(t_k) = \Delta t r_k + \varepsilon_k$ , сопровождающую процесс дискретизации сигналов по времени в (2.5):  $\tilde{S}(t_k) = S[t_k + A(t_k)] = S(t_k + \Delta t r_k + \varepsilon_k)$ .

На основе представления ДИ в (2.6) используем для элементов цифрового эквивалента деформированного сигнала  $\tilde{\mathbf{s}} = (\tilde{s}_1, \dots, \tilde{s}_n)^T = [\tilde{S}(t_1), \dots, \tilde{S}(t_n)]^T$ ,  $k = \overline{1, n}$  следующее приближение:

$$\tilde{s}_k = S(t_k + r_k \Delta t + \varepsilon_k) \cong S(t_k + r_k \Delta t) + S'(t_k + r_k \Delta t) \varepsilon_k, \quad (2.7)$$

где  $S'$  — производная функции  $S$ . Если учесть, что  $s_k = S(t_k) = S[(k-1)\Delta t]$ , то

$$s_{k+r_k} = S(t_k + r_k \Delta t) = S[(k + r_k - 1)\Delta t], \quad (2.8)$$

и, следовательно, (2.7) можно переписать в виде  $\tilde{s}_k \cong s_{k+r_k} + s'_{k+r_k} \varepsilon_k$  или в виде

$$\tilde{\mathbf{s}} = (\tilde{s}_1, \dots, \tilde{s}_n)^T \cong \hat{\mathbf{s}} + \text{diag}(\hat{\mathbf{s}}') \boldsymbol{\varepsilon} = F(\mathbf{s}, \mathbf{a}) + \text{diag}[F(\mathbf{s}', \mathbf{a})] \boldsymbol{\varepsilon}, \quad (2.9)$$

где  $F(\mathbf{s}, \mathbf{a})$  — оператор перестановки, такой, что  $\hat{s}_k = s_{k+r_k}$  и, синхронно —  $\hat{s}'_k = s'_{k+r_k}$ , а  $\text{diag}()$  — обозначение диагональной матрицы.

### 2.3.1 Анализ свойств оператора перестановки

Исходя из (2.9) мы видим, что воздействие ДИ в рамках введенного приближения сводится к действию нелинейного оператора перестановки и добавлению дополнительной аддитивной помеховой составляющей, обусловленной погрешностью квантования непрерывной деформирующей функции.

**Утверждение 2.2.** Для любого индекса  $k$  элемента  $\hat{s}_k = s_{k+r_k}$  вектора  $\hat{\mathbf{s}}$ , полученного после применения оператора  $F(\mathbf{s}, \mathbf{a})$ , при фиксированном значении  $\mathbf{a}$

существует только один используемый элемент  $s_{k+r_k}$  исходной последовательности  $\mathbf{s}$ , при этом для любого индекса  $j$  элемента  $s_j$  исходной последовательности  $\mathbf{s}$  при фиксированном  $\mathbf{a}$  таком, что существует  $k$  такое, что  $r_k \neq 0$ , возможно более одной позиции, в которую данный элемент может быть переставлен.

Первая часть утверждения вытекает из того, что функция  $t + A(t)$  является однозначной функцией времени, и, следовательно, значение  $k + r_k$  в (2.8), где  $r_k = \text{round}[A(t_k) / \Delta t]$ , тоже является единственным.

Вторая часть утверждения вытекает из того, что могут существовать два индекса  $k'$ ,  $k''$  ( $k' \neq k''$ ) и связанные с ними индексы  $r_{k'}$ ,  $r_{k''}$  такие, что  $j = k' + r_{k'} = k'' + r_{k''}$ . Действительно, достаточно взять, например,  $k' = j$ ,  $k'' = j + 1$ ,  $r_{k'} = 0$  и  $r_{k''} = -1$ . Тогда  $k' + r_{k'} = (j) + 0 = j = (j + 1) - 1 = k'' + r_{k''}$ . В противном случае не выполняется условие существования такого  $k$ , что  $r_k \neq 0$ , т.е. фактически все  $r_k = 0$ . Последняя ситуация означает, что все значения  $r_k = \text{round}[A(t_k) / \Delta t]$  равны нулю, и при внесении ДИ перемещений элементов вектора  $\mathbf{s}$  не происходит.

Из доказательства утверждения непосредственно следует, что оператор  $F(\mathbf{s}, \mathbf{a})$  является случайным оператором, определяющим возможные размещения с повторениями элементов векторов  $\mathbf{s}$  и  $\mathbf{s}'$  в элементы векторов  $\hat{\mathbf{s}}$  и  $\hat{\mathbf{s}}'$ , соответственно:  $\hat{\mathbf{s}} = F(\mathbf{s}, \mathbf{a})$ ,  $\hat{\mathbf{s}}' = F(\mathbf{s}', \mathbf{a})$ . Для случайной функции деформации он задает вероятностные переходы при преобразовании любой последовательности  $\mathbf{s}$  в конечное множество последовательностей  $\hat{\mathbf{s}}$ . При этом выполняются размещения с повторениями в получаемых последовательностях одних элементов и выпадениями других. Оператор для конкретной реализации ДИ не является взаимно однозначным.

Рассмотрим теперь возможное число получающихся размещений с повторениями при различных ограничениях, накладываемых на функцию деформации.

**Утверждение 2.3.** Пусть деформирующее искажение формируется так, что  $|r_k| \leq k$ , при  $0 \leq k \leq n/2$  и  $|r_k| \leq 1-k$ , при  $n/2 \leq k \leq n$ , что означает условие невыхода за пределы области определения дискретной функции. Тогда оператор  $F(\mathbf{s}, \mathbf{a})$  определяет вероятностные переходы при преобразовании любой последовательности  $\mathbf{s}$  длины  $n$  в конечное множество из  $R \leq n!$  последовательностей  $\hat{\mathbf{S}} = \{\hat{\mathbf{s}}\}$ , каждая из которых соответствует возможной комбинации элементов исходной последовательности.

Пусть  $n = 2m + 1$  — нечетное число. Тогда, учитывая ранее введенное ограничение на амплитуду ДИ, количество допустимых перестановок определяется соотношением:

$$R = 1 \times 3 \times 5 \times \dots \times (2m-1) \times (2m+1) \times (2m-1) \times \dots \times 5 \times 3 \times 1 = (2m+1)!!(2m-1)!! \leq n!.$$

Пусть теперь  $n = 2m$  — четное число. Тогда количество допустимых перестановок составляет

$$R = 1 \times 3 \times 5 \times \dots \times (2m-1) \times (2m-1) \times \dots \times 5 \times 3 \times 1 = (2m-1)!!(2m-1)!! \leq n!.$$

**Утверждение 2.4.** Если не учитывать границы области определения  $\Xi_t$  и принять  $r_k \in [-\max(|r_k|), \max(|r_k|)]$  в отсутствие ограничения  $t + A(t) \in \Xi_t$ , то число используемых элементов исходной последовательности  $\mathbf{s}$  для размещения одного элемента в последовательности  $\hat{\mathbf{s}}$  составляет величину  $2\max(|r_k|) + 1$ , и число возможных размещений с повторениями для всех  $n$  элементов  $\hat{\mathbf{s}}$ , таким образом, равно

$$[2\max(|r_k|) + 1]^n. \quad (2.10)$$

Если при этом ограничить набор допустимых значений  $A(t)$  при приближении к границам  $\Xi_t$  при помощи трапецевидной оконной функции:

$$O(t) = \begin{cases} \frac{t}{|\max[A(t)]|}, & 0 \leq t < |\max[A(t)]|, \\ 1, & |\max[A(t)]| \leq t \leq T - |\max[A(t)]|, \\ \frac{T-t}{|\max[A(t)]|}, & T - |\max[A(t)]| < t \leq T, \end{cases} \quad (2.11)$$

заменяя  $A(t)$  на  $O(t)A(t)$ , то в границах промежутка  $[|\max[A(t)]|, T - |\max[A(t)]|]$  производится размещение с повторениями по  $n - 2 \max(|r_k|)$  элементов последовательности  $\hat{s}$  из  $2 \max(|r_k|) + 1$  элементов последовательности  $s_i$  на основе (2.10). При этом на краю области определения  $\Xi_t$  для каждого значения  $t_k$ , число возможных размещений с повторениями 1 элемента последовательности  $\hat{s}$  из

$$\begin{aligned} 2 \max(|\text{round}[O(t_k)A(t_k) / \Delta t]|) + 1 &= 2 \max\left(|\text{round}\left[\frac{t_k}{\max[A(t_k)]} A(t_k) / \Delta t\right]|\right) + 1 = \\ &= 2 \left| \text{round}\left[\frac{t_k}{\max[A(t_k)]} \max\{A(t_k)\} / \Delta t\right] \right| + 1 = \\ &= 2 \text{round}((k-1)\Delta t / \Delta t) + 1 = 2(k-1) + 1 = 2k - 1 \end{aligned} \quad (2.12)$$

элементов последовательности  $s$  составляет  $2k - 1$  варианта размещения и, в итоге, с учетом симметричности оконной функции  $O(t)$  (2.12) на другом конце области определения  $\Xi_t$ , общее число размещений с повторениями составляет

$$[2 \max(|r_k|) + 1]^{n-2 \max(|r_k|)} \times [2 \max(|r_k|) - 1]^2 \times \dots \times 3^2 \times 1^2.$$

**Утверждение 2.5.** Если на деформированный сигнал  $\tilde{S}(t) = S[t + A(t)]$ ,  $t \in \Xi_t$  наложить ограничение

$$t_1 + A(t_1) \leq t_2 + A(t_2), \quad \forall t_1 < t_2,$$

т.е. ввести запрет перестановки его значений, то производная функции  $A(t)$  —  $A'(t)$  должна удовлетворять следующему условию:

$$A'(t) \geq -1.$$

В этом случае для цифрового сигнала

$$k + a_k \geq (k - 1) + a_{k-1}, \quad k = \overline{2, n},$$

и общее число возможных вариантов размещений с повторениями в отсутствие ограничения  $t + A(t) \in \Xi_t$  становится меньшим (2.10).

Вывод аналитической формулы расчета точного количества размещений в данном случае не представляется возможным из-за необходимости рассмотрения каждого из возможных вариантов значений вектора  $\mathbf{a}$ . Тем не менее в частных случаях определение этого количества может быть произведено путем проверки возможности появления каждого конкретного значения вектора  $\mathbf{a}$ .

### 2.3.2 Анализ свойств дополнительной помеховой составляющей

Как уже было отмечено ранее, при использовании приближенного представления (2.7) и (2.9) дополнительная помеховая составляющая, определяемая вектором  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$ , носит аддитивный характер и формируется в результате квантования или дискретизации по уровню заданной в непрерывном времени функции  $t + A(t)$ ,  $t \in [0, T]$ . Каждая величина  $\varepsilon_k$  может рассматриваться как точечное значение  $\varepsilon_k = \varepsilon(t_k)$  сигнала  $\varepsilon(t) = U_A(t) - \tilde{U}_A(t)$ , где  $\tilde{U}_A(t) = t + \text{round}[A(t) / \Delta t] \Delta t$  — сигнал, полученный при квантовании  $U_A(t) = t + A(t)$  методом ступенчатой интерполяции.

В известных работах, посвященных квантованию сигналов [75, 76], достаточно подробно исследованы вопросы анализа статистических характеристик возникающих при этом ошибок. В соответствии с их результатами погрешность квантования случайной функции при относительно малом интервале квантования практически распределена по равномерному закону и имеет малый уровень корреляции во времени. Таким образом, можно предположить, что при достаточно малом значении  $\Delta t$ , выбираемым в соответствии со второй теоремой квантования

Видроу [76], каждая составляющая  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$  имеет близкий к равномерному распределению вид:

$$\varepsilon_k = \varepsilon(t_k) \sim p(\varepsilon_k) = \begin{cases} 1 / \Delta t, & |\varepsilon_k| \leq \Delta t / 2, \\ 0, & |\varepsilon_k| > \Delta t / 2. \end{cases} \quad (2.13)$$

При этом взаимная корреляция элементов последовательности  $\varepsilon_1, \dots, \varepsilon_n$  весьма мала, поскольку автокорреляционная функция сигнала  $\varepsilon(t)$  в непрерывном времени имеет узкий пик с шириной, сопоставимой с величиной  $\Delta t$ .

Для подтверждения предположений вида закона распределения и корреляционных свойств последовательности  $\varepsilon_1, \dots, \varepsilon_n$  с учетом специфики решаемой задачи было проведено моделирование указанных статистических характеристик погрешностей квантования функции  $U_A(t) = t + A(t)$ . Модель была разработана в среде MATLAB. При проведении исследования рассматривались две ситуации, связанные с конкретным способом задания ДИ.

Первый вариант предусматривал задание функции  $A(t)$  как детерминированной функции колокольного вида

$$A(t) = A_m \exp \left[ -\frac{(t - T / 2)^2}{2\sigma^2} \right],$$

где  $A_m$  — амплитуда деформирующей функции;  $\sigma$  — параметр, определяющий ширину области локализации колокольной функции на заданном интервале времени. Применение подобной функции деформации к исходному сигналу приводит к смещению и перекосу формы деформируемого сигнала влево.

На рисунке 2.6 представлены получаемые при квантовании по уровню зависимости для сигналов  $U_A(t)$  и  $\tilde{U}_A(t)$  (рисунок 2.6а), а также для ошибки  $\varepsilon(t)$  (рисунок 2.6б) при задании  $T = 1$ ,  $\Delta t = 0,1$  (общее число уровней квантования  $M = 10$ ),  $A_m = 0,5$ ,  $\sigma = 0,1$ .

На рисунке 2.7а,б представлены гистограмма ошибки  $\varepsilon(t)$  и оценка ее автокорреляционной функции для интервала квантования  $\Delta t = 0,1$ . На рисунке 2.7в,г представлены для сравнения аналогичная гистограмма и автокорреляционная функция, полученные при прочих одинаковых параметрах для  $\Delta t = 0,05$ .

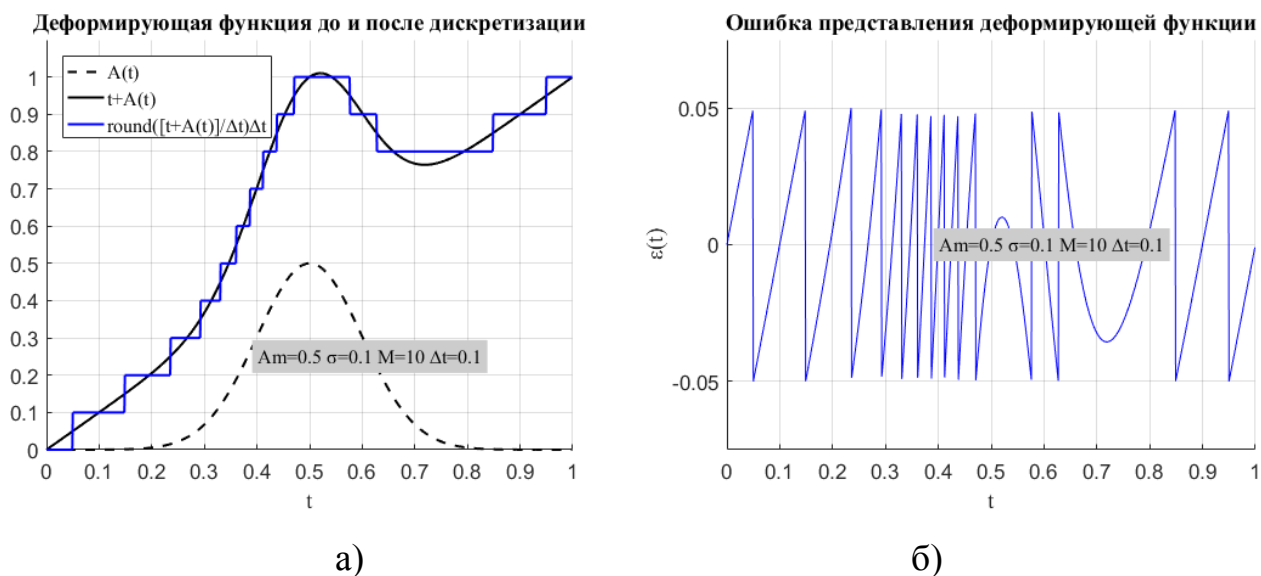


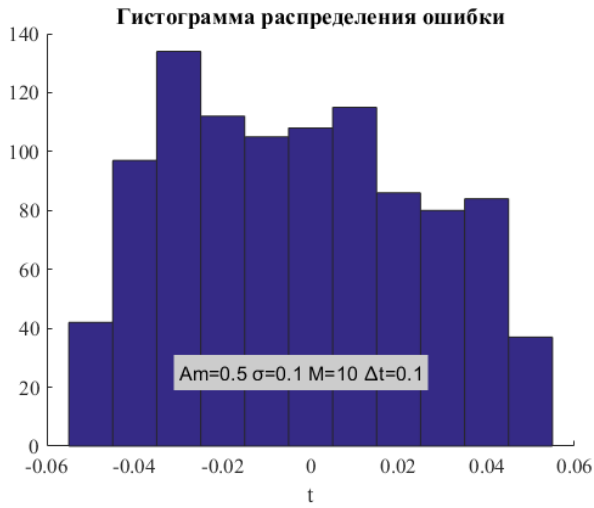
Рисунок 2.6 – Зависимости для  $U_A(t) = t + A(t)$  и  $\tilde{U}_A(t) = t + \text{round}[A(t) / \Delta t] \Delta t$  (а), а также  $\varepsilon(t)$  (б) при задании  $T = 1$ ,  $\Delta t = 0,1$  (общее число уровней квантования  $M = 10$ ),  $A_m = 0,5$ ,  $\sigma = 0,1$  для детерминированной функции деформации

Второй вариант предусматривал задание функции  $A(t)$  как произведения реализации стационарного случайного процесса  $A_0(t)$  с заданной функцией корреляции  $B_0(\tau)$  на оконную функцию  $O(t)$ , обеспечивающую выполнение условия  $|A(t)| < \min\{t, T - t\}$ :

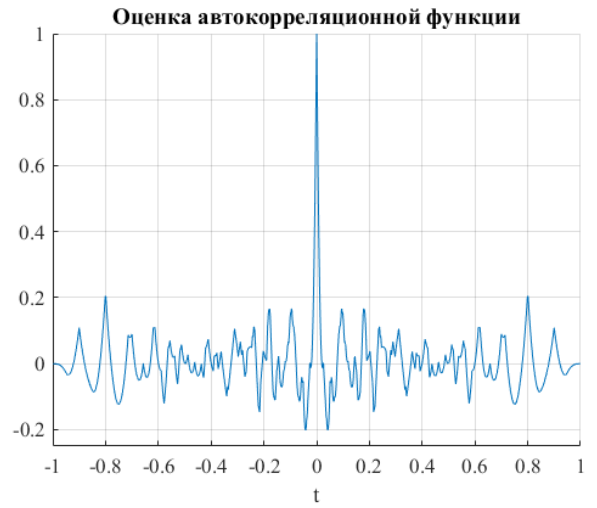
$$A(t) = O(t)A_0(t),$$

$$A_0(t) \sim B_0(\tau) = B_m^2 \exp[-\alpha\tau^2], \quad O(t) = \sin(\pi t / T),$$

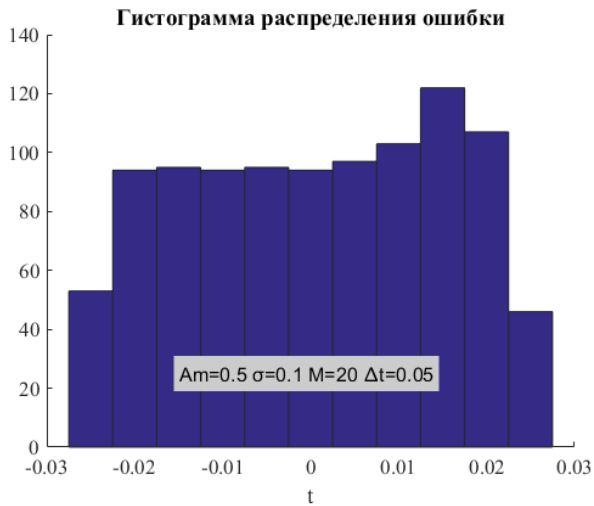
где  $B_m^2$  — дисперсия деформирующей функции  $A_0(t)$ ;  $\alpha$  — параметр, определяющий величину корреляции на заданном интервале времени.



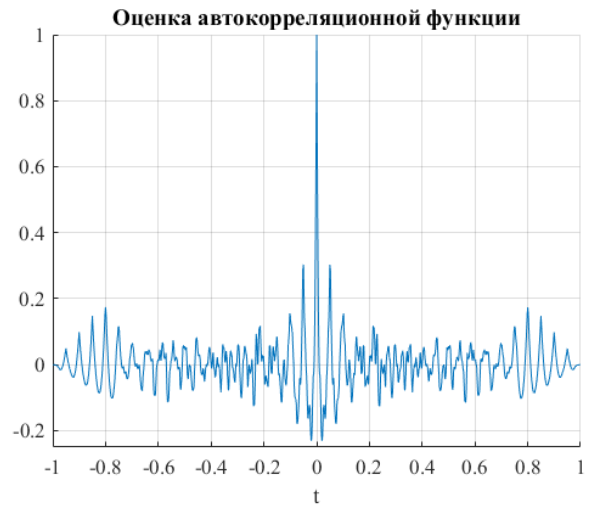
а)



б)



в)



г)

Рисунок 2.7 – Гистограмма распределения ошибки  $\varepsilon(t)$  и оценка ее автокорреляционной функции для интервалов квантования  $\Delta t = 0,1$  (а, б) и  $\Delta t = 0,05$  (в, г) для детерминированной функции деформации

На рисунке 2.8 представлены получаемые при квантовании зависимости для сигналов  $U_A(t)$  и  $\tilde{U}_A(t)$  (рисунок 2.8а), а также ошибки  $\varepsilon(t)$  (рисунок 2.8б) при



задании  $T = 1$ ,  $\Delta t = 0,1$  (общее число уровней квантования  $M = 10$ ),  $B_m = 0,1$ ,  $\alpha = 0,0001$  для функции деформации на основе случайного процесса  $A_0(t)$ .

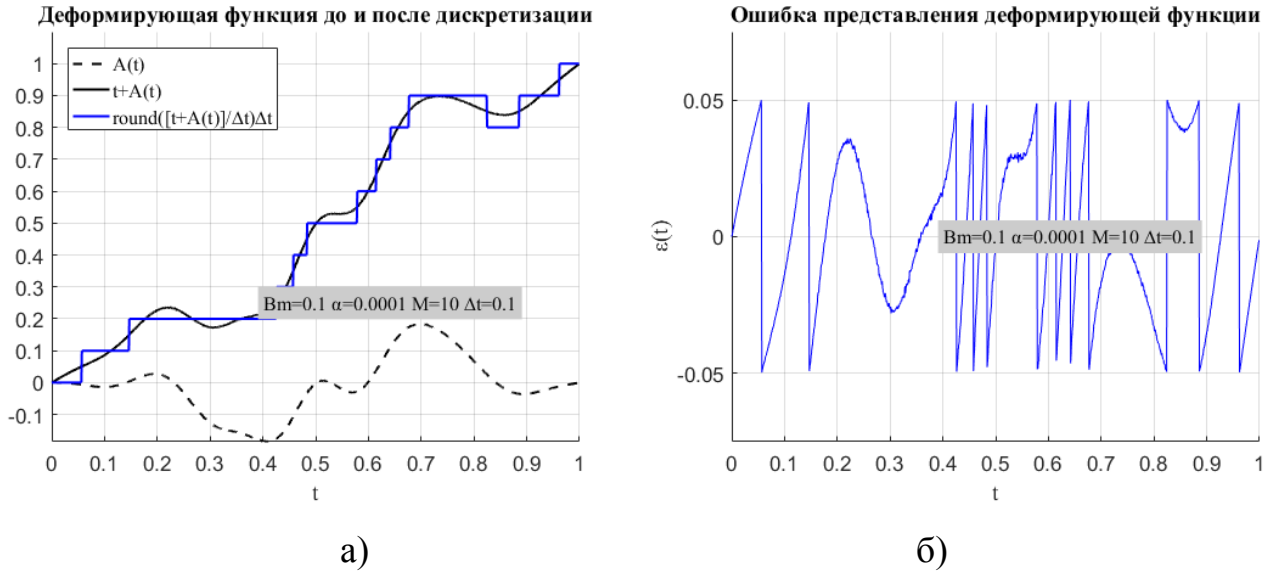
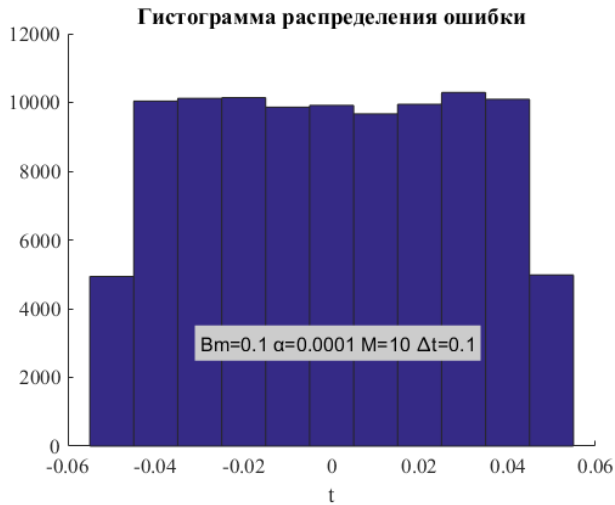
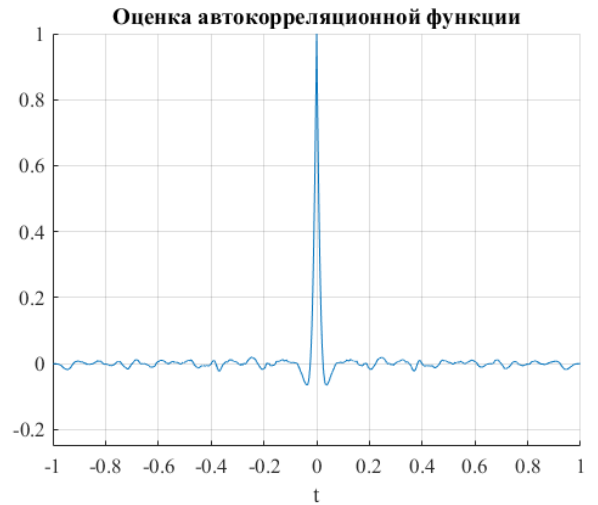


Рисунок 2.8 – Зависимости для  $U_A(t) = t + A(t)$  и  $\tilde{U}_A(t) = t + \text{round}[A(t) / \Delta t]\Delta t$  (а), а также  $\varepsilon(t)$  (б) при задании  $T = 1$ ,  $\Delta t = 0,1$  (общее число уровней квантования  $M = 10$ ),  $B_m = 0,1$ ,  $\alpha = 0,0001$  для функции деформации на основе случайного процесса  $A_0(t)$

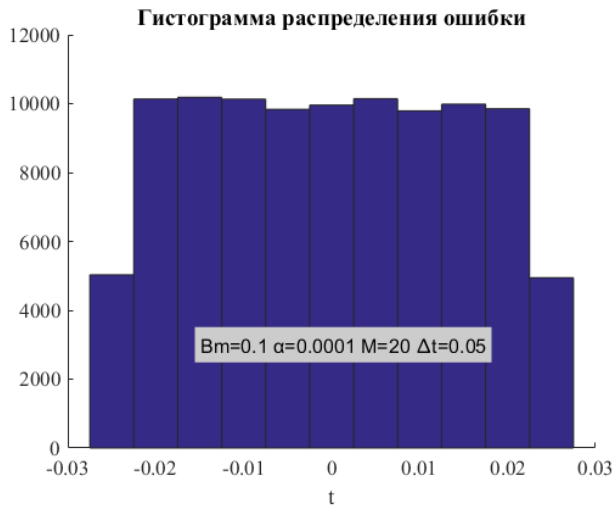
На рисунке 2.9а,б представлены гистограмма ошибки  $\varepsilon(t)$  и ее усредненная автокорреляционная функция, полученные на основе 100 реализаций данного случайного процесса для интервала квантования  $\Delta t = 0,1$ . На рисунке 2.9в,г для сравнения представлены аналогичная гистограмма и автокорреляционная функция, полученные при прочих одинаковых параметрах для  $\Delta t = 0,05$ .



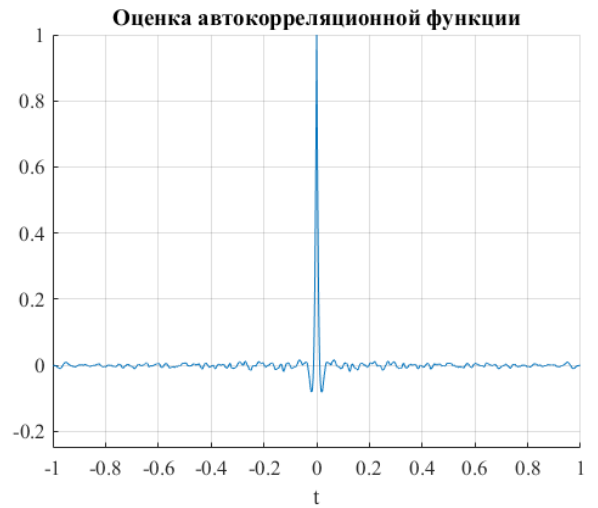
а)



б)



в)



г)

Рисунок 2.9 – Гистограмма распределения ошибки и оценка ее автокорреляционной функции для интервалов квантования  $\Delta t = 0,1$  (а, б) и  $\Delta t = 0,05$  (в, г) для функции деформации на основе случайного процесса  $A_0(t)$

Анализ представленных зависимостей позволяет подтвердить сделанное ранее предположение относительно вида закона распределения компонентов вектора  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$  как элементов последовательности погрешностей квантования, оказывающих дополнительное помеховое воздействие на обрабатываемый цифровой сигнал, а также относительно степени коррелированности элементов этой

последовательности. Как показывают приведенные графики, с весьма большой степенью достоверности можно считать закон распределения погрешности равномерным в интервале  $[-\Delta t / 2, \Delta t / 2]$ . При этом время корреляции процесса  $\varepsilon(t)$  можно оценить величиной порядка  $\Delta t / 2$ , что позволяет считать его значения, полученные при дискретизации с указанным интервалом, практически некоррелированными.

Таким образом, вносимая за счет ДИ дополнительная аддитивная помеховая составляющая может рассматриваться как центрированная последовательность распределенных по равномерному закону случайных чисел, которые практически не коррелированы между собой. При проведении синтеза и анализа алгоритмов распознавания далее будем считать их независимыми.

### **Выводы по главе**

1. Представлена математическая модель внесения деформирующих искажений (ДИ) для любой исходной функции от нескольких переменных. Описаны особенности и алгоритмы внесения ДИ в решетчатые функции дискретных аргументов с использованием различных видов интерполяции, в том числе на основе применения радиально-базисных функций. Сформулировано и доказано утверждение о верхней границе ошибки интерполяции деформируемой функции, возникающей при внесении ДИ.

Приведены результаты экспериментов по сравнению скорости и точности работы алгоритмов внесения ДИ, использующих разные типы интерполяции, а также по определению оптимальных параметров интерполяции, основанной на применении радиально-базисных функций. На основе проведенных исследований, сделан вывод, что предложенный способ внесения деформации в сигналы и изображения на основе радиально-базисных функций показывает сравнимые результаты с другими методами интерполяции, а в ряде случаев превосходит их по

точности и является перспективным для использования при размножении обучающих выборок путем внесения в них ДИ в ходе решения различных задач распознавания сигналов и изображений.

2. Для задач синтеза алгоритмов распознавания цифровых сигналов в условиях деформирующих искажений предложена эквивалентная статистическая модель влияния ДИ, основанная на приближенном представлении нелинейного оператора деформации как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. Доказан ряд утверждений относительно свойств преобразований, получаемых на основе такого представления оператора деформации, и даны оценки получающегося количества размещений в зависимости от ограничений, накладываемых на функцию внесения ДИ.

В ходе исследований теоретически и экспериментально показано, что при квантовании деформирующей функции с относительно малым интервалом квантования для наблюдаемого сигнала возникает дополнительная аддитивная помеховая составляющая, имеющая приближенно равномерный закон распределения и малый уровень корреляции во времени.

### 3 Статистический синтез и анализ алгоритмов распознавания в условиях деформирующих искажений и аддитивного шума

#### 3.1 Синтез и анализ оптимальных и квазиоптимальных алгоритмов распознавания цифровых сигналов в условиях деформирующих искажений

##### 3.1.1 Постановка задачи распознавания детерминированных сигналов в присутствии шумов и деформирующих искажений

Рассмотрим без существенного ограничения общности следующую постановку задачи. Пусть на интервале времени  $\Xi_t = [0, T]$  заданы сигналы известной формы, определяющие эталонные описания  $M$  классов образов  $\omega_i : S_i(t), i = \overline{1, M}, t \in \Xi_t$ . Наблюдению доступны реализации случайных функций, формируемые на основе этих сигналов при воздействии ДИ и аддитивного гауссовского шума и имеющие вид

$$U_i(t) = \tilde{S}_i(t) + V(t) = S_i[t + A(t)] + V(t), \quad t \in \Xi_t, \quad i = \overline{1, M}, \quad (3.1)$$

где  $U_i(t)$  — реализация наблюдаемой случайной функции;  $\tilde{S}_i(t)$  — реализация случайной функции, описывающей деформированный полезный сигнал;  $A(t)$  — однозначная дифференцируемая функция, являющаяся в общем случае реализацией случайной функции, описывающей ДИ с известными статистическими характеристиками, такая, что  $t + A(t) \in \Xi_t$ ;  $V(t)$  — реализация случайной функции, описывающей действие аддитивного гауссовского шума с параметрами  $M[V(t)] = 0$ ,  $M[V(t)V(t + \tau)] = \sigma^2 R(\tau)$ .

При цифровой обработке в результате дискретизации по времени с интервалом  $\Delta t$  формируются цифровые эквиваленты наблюдаемых в непрерывном времени сигналов путем фиксации их значений в моменты времени  $t_k = (k - 1)\Delta t, k = \overline{1, n}$ ,  $\Delta t = T / (n - 1), t_1 = 0, t_n = T$ :

$$U_i(t_k) = \tilde{S}_i(t_k) + V(t_k) = S_i[t_k + A(t_k)] + V(t_k), \quad i = \overline{1, M}. \quad (3.2)$$

Введем также векторное представление наблюдаемой функции как  $\mathbf{u}_i = [U_i(t_1), \dots, U_i(t_n)]^T$ . Пусть сигналы различных классов появляются на входе распознающей системы с априорными вероятностями  $p(\omega_i)$ ,  $i = \overline{1, M}$  в виде вектора наблюдений  $\mathbf{x} = (x_1, \dots, x_n)^T$ , принадлежащего к одному из классов  $\omega_i$ :  $\mathbf{x} = \mathbf{u}_i$ . Требуется сформировать решение, относящее каждое наблюдение к тому или иному классу. Как известно [3, 77], в этом случае решение задачи при использовании простой функции потерь определяется оптимальным решающим правилом вида

$$\omega_i : l_{ij}(x) = \frac{p(x/\omega_i)}{p(x/\omega_j)} \underset{\omega_j}{\overset{\omega_i}{>}} l_{0ij} = \frac{p(\omega_j)}{p(\omega_i)}, \quad j = \overline{1, M}, \quad i \neq j.$$

Далее, с целью упрощения анализа (без ограничений общности по структуре самого правила) будем рассматривать двухальтернативное решающее правило

$$l(x) = \frac{p(\mathbf{x}/\omega_1)}{p(\mathbf{x}/\omega_2)} \underset{\omega_2}{\overset{\omega_1}{>}} l_0 = \frac{p(\omega_2)}{p(\omega_1)}, \quad (3.3)$$

где  $l(x)$  — величина отношения правдоподобия, сравниваемая при принятии решения с порогом  $l_0$ . Далее также будем считать априорные вероятности одинаковыми ( $l_0 = 1$ ).

Для синтеза алгоритма распознавания в виде (3.3) выполним анализ исходного представления (3.1) и (3.2). Запишем функцию правдоподобия (ФП) любого из классов  $\omega_i$ ,  $i = \overline{1, 2}$ , в виде

$$p(\mathbf{x}/\omega_i) = \int_{\Gamma} p(\mathbf{x}, \mathbf{a}/\omega_i) d\mathbf{a} = \int_{\Gamma} p(\mathbf{x}/\mathbf{a}, \omega_i) p(\mathbf{a}/\omega_i) d\mathbf{a}, \quad (3.4)$$

где  $p(\mathbf{a} / \omega_i) = p(\mathbf{a})$  — плотность распределения вероятностей вектор-функции деформации в (3.2):  $\mathbf{a} = (a_1, \dots, a_n)^T = [A(t_1), \dots, A(t_n)]^T$ , вид которой, будем считать, не зависит от класса сигнала.

На основе эквивалентного представления ДИ (раздел 2.3) как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации, используем для элементов цифрового эквивалента деформированного сигнала следующее приближение:

$$\tilde{s}_{i,k} = S_i(t_k + r_k \Delta t + \varepsilon_k) \cong S_i(t_k + r_k \Delta t) + S'_i(t_k + r_k \Delta t) \varepsilon_k, \quad (3.5)$$

где  $S'_i$  — производная функции  $S_i$ ;  $r_k = \overline{\text{round}(a_k / \Delta t)}$ ,  $k = \overline{1, n}$ ;  $\varepsilon_k$  — случайная величина, характеризующая погрешность квантования по уровню заданной в непрерывном времени функции деформации  $A(t_k) = \Delta t r_k + \varepsilon_k$ .

Если учесть, что  $s_{i,k} = S_i(t_k) = S_i[(k-1)\Delta t]$ , то

$$s_{i,k+r_k} = S_i(t_k + r_k \Delta t) = S_i[(k+r_k-1)\Delta t],$$

и, следовательно, (3.5) можно переписать в виде  $\tilde{s}_{i,k} \cong s_{i,k+r_k} + s'_{i,k+r_k} \varepsilon_k$  или в виде

$$\tilde{\mathbf{s}}_i = (\tilde{s}_{i,1}, \dots, \tilde{s}_{i,n})^T \cong \hat{\mathbf{s}}_i + \text{diag}(\hat{\mathbf{s}}'_i) \boldsymbol{\varepsilon} = F(\mathbf{s}_i, \mathbf{a}) + \text{diag}[F(\mathbf{s}'_i, \mathbf{a})] \boldsymbol{\varepsilon},$$

где  $F(\mathbf{s}_i, \mathbf{a})$  — оператор размещения с повторениями элементов исходного дискретного сигнала, такой, что  $\hat{s}_{i,k} = s_{i,k+r_k}$  и, синхронно, —  $\hat{s}'_{i,k} = s'_{i,k+r_k}$ ;  $\text{diag}()$  — обозначение диагональной матрицы.

Таким образом, вектор наблюдений  $\mathbf{u}_i$ , формируемый сигналом каждого класса, может быть представлен в виде

$$\mathbf{u}_i \cong \hat{\mathbf{s}}_i + \text{diag}(\hat{\mathbf{s}}'_i) \boldsymbol{\varepsilon} + \mathbf{v} = F(\mathbf{s}_i, \mathbf{a}) + \text{diag}[F(\mathbf{s}'_i, \mathbf{a})] \boldsymbol{\varepsilon} + \mathbf{v}, \quad u_{i,k} \cong \hat{s}_{i,k} + \hat{s}'_{i,k} \varepsilon_k + v_k, \quad (3.6)$$

где  $\hat{\mathbf{s}}_i = (\hat{s}_{i,1}, \dots, \hat{s}_{i,n})^T = (s_{i,1+r_1}, \dots, s_{i,n+r_n})^T$  — вектор, получаемый после размещения с повторениями элементов вектора  $\mathbf{s}_i$  по закону, определяемому оператором  $F(\mathbf{s}_i, \mathbf{a})$ ;

$\hat{\mathbf{s}}'_i = (\hat{s}'_{i,1}, \dots, \hat{s}'_{i,n})^T = (s'_{i,1+r_1}, \dots, s'_{i,n+r_n})^T$  — вектор, получаемый после синхронного с  $\mathbf{s}_i$  размещения с повторением элементов вектора производных  $\hat{\mathbf{s}}'_i$ ;  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$  — вектор дополнительных возмущений, вызванных погрешностью дискретизации, и распределенный по закону (2.13);  $\mathbf{v} = (v_1, \dots, v_n) = [V(t_1), \dots, V(t_n)]^T$  — аддитивный гауссовский шум.

### 3.1.2 Основные соотношения для проведения синтеза и анализа алгоритмов распознавания

Выполненный анализ модели наблюдений (3.6) позволяет перейти к синтезу алгоритмов распознавания. Запишем общие выражения для ФП каждого из классов в (3.4) в виде

$$p(\mathbf{x} / \omega_i) = \int_{\Gamma} p(\mathbf{x}, \mathbf{a} / \omega_i) d\mathbf{a} = \sum_{\{\psi\}} P_{\mathbf{a}}(\psi) p(\mathbf{x} / \psi, \omega_i), \quad (3.7)$$

где  $P_{\mathbf{a}}(\psi)$  — вероятности возможных вариантов размещений с повторениями  $\psi = \{r_1, \dots, r_n\}$ , которые могут быть получены на множестве случайных реализаций вектора  $\mathbf{a}$ . Каждая такая вероятность определяется на основе многомерной плотности распределения  $p(\mathbf{a}) = p(a_1, \dots, a_n)$  в точном и приближенном (при малых  $\Delta t$ ) виде следующим образом:

$$P_{\mathbf{a}}(\psi) = \int_{r_1\Delta t - \Delta t/2}^{r_1\Delta t + \Delta t/2} \dots \int_{r_n\Delta t - \Delta t/2}^{r_n\Delta t + \Delta t/2} p(a_1, \dots, a_n) da_1, \dots, da_n, \quad P_{\mathbf{a}}(\psi) \cong \frac{p(r_1\Delta t, \dots, r_n\Delta t)}{\sum_{\{\psi\}} p(r_1(\psi)\Delta t, \dots, r_n(\psi)\Delta t)}. \quad (3.8)$$

Условная плотность  $p(\mathbf{x} / \psi, \omega_i)$  в (3.7) представляется как композиция законов распределения аддитивного гауссовского шума  $\mathbf{v} = \mathbf{x} - \mathbf{z}$  и дополнительной помеховой составляющей  $\boldsymbol{\varepsilon} = \mathbf{z}$  из (2.13):



$$p(\mathbf{x} / \psi, \omega_i) = \int p_1(\mathbf{x} - \mathbf{z} / \psi, \omega_i) p_2(\mathbf{z} / \psi, \omega_i) d\mathbf{z}, \quad p_1(\mathbf{x} - \mathbf{z} / \psi, \omega_i) = N(\mathbf{x} - \mathbf{z}, \hat{\mathbf{s}}_i, \mathbf{C}),$$

$$p_2(\mathbf{z} / \psi, \omega_i) = \prod_{k=1}^n p_2(z_k / \psi, \omega_i), \quad p_2(z_k / \psi, \omega_i) = \begin{cases} \frac{1}{\hat{s}'_{i,k} \Delta t}, & |z_k| \leq \hat{s}'_{i,k} \frac{\Delta t}{2}, \\ 0, & |z_k| > \hat{s}'_{i,k} \frac{\Delta t}{2}. \end{cases} \quad (3.9)$$

где  $N(\mathbf{x} - \mathbf{z}, \hat{\mathbf{s}}_i, \mathbf{C})$  — обозначение гауссовской плотности распределения вектора аддитивного шума  $\mathbf{v} = \mathbf{x} - \mathbf{z}$  с математическим ожиданием  $\hat{\mathbf{s}}_i$  и матрицей ковариаций —  $\mathbf{C}$ .

При малом или сопоставимом с единицей соотношении  $\sigma_{\varepsilon}^2 / \sigma^2$  дисперсий шума дискретизации  $\sigma_{\varepsilon}^2$  и гауссовского шума  $\sigma^2$  плотность распределения композиции (3.9) близка к гауссовской, и результирующее условное относительно  $\psi = \{r_1, \dots, r_n\}$  распределение вектора признаков в этом случае будет иметь вид

$$p(\mathbf{x} / \psi, \omega_i) = N\left(\mathbf{x}, \hat{\mathbf{s}}_i, \mathbf{C} + \frac{1}{12} \text{diag}\{(\hat{s}'_{i,1} \Delta t)^2, \dots, (\hat{s}'_{i,n} \Delta t)^2\}\right). \quad (3.10)$$

На основе полученных выше представлений ФП может быть синтезировано оптимальное (в рамках введенных допущений) решающее правило в форме (3.3).

Следует отметить, что непосредственное применение оптимального алгоритма на основе (3.7) очевидно вызывает определенные трудности, поскольку при большом числе возможных вариантов размещений с повторениями процесс вычисления  $p(\mathbf{x} / \omega_i)$  становится весьма затратным по времени и используемому объему памяти. Поэтому для реализации экономного в вычислительном отношении алгоритма принятия решений предлагается также рассмотреть гауссовское приближение для распределений  $p(\mathbf{x} / \omega_i)$ . С этой целью получим выражения для условных векторов математических ожиданий и матриц ковариаций. Опуская громоздкие, но не сложные преобразования, можно увидеть, что они представляются в конечном виде следующим образом:

$$\begin{aligned}
\mathbf{m}_i &= M[\mathbf{x} / \omega_i] = \sum_{\{\psi\}} \mathbf{m}_i(\psi) P_a(\psi), \\
\mathbf{m}_i(\psi) &= M[\mathbf{x} / \psi, \omega_i] = \int \mathbf{x} p(\mathbf{x} / \psi, \omega_i) d\mathbf{x}, \\
\mathbf{C}_i &= M[(\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T / \omega_i] = \\
&= \sum_{\{\psi\}} \mathbf{C}_i(\psi) P_a(\psi) + \sum_{\{\psi\}} (\mathbf{m}_i(\psi) - \mathbf{m}_i)(\mathbf{m}_i(\psi) - \mathbf{m}_i)^T P_a(\psi), \\
\mathbf{C}_i(\psi) &= M[\{\mathbf{x} - \mathbf{m}_i(\psi)\} \{\mathbf{x} - \mathbf{m}_i(\psi)\}^T / \psi, \omega_i] = \\
&= \int \{\mathbf{x} - \mathbf{m}_i(\psi)\} \{\mathbf{x} - \mathbf{m}_i(\psi)\}^T p(\mathbf{x} / \psi, \omega_i) d\mathbf{x} = \\
&= \mathbf{C} + \frac{1}{12} \text{diag}\{(\hat{s}'_{i,1} \Delta t)^2, \dots, (\hat{s}'_{i,n} \Delta t)^2\}.
\end{aligned} \tag{3.11}$$

В данном приближении решающее правило имеет структуру, которая может быть представлена в явном виде на основе следующих соотношений:

$$\mathbf{g}(\mathbf{x}) = \ln \frac{N(\mathbf{x}, \mathbf{m}_1, \mathbf{C}_1) \stackrel{\omega_1}{>}}{N(\mathbf{x}, \mathbf{m}_2, \mathbf{C}_2) \stackrel{\omega_2}{<}} l'_0. \tag{3.12}$$

т.е. фактически реализует распознавание гауссовского вектора для распределений классов, имеющих различные математические ожидания и матрицы ковариаций.

Для проведения анализа ошибок в представленных выше решающих правилах (3.3) и (3.12) помимо универсального подхода, основанного на прямом имитационном моделировании, могут быть использованы различные оценки и границы. Так, для разделяющей функции вида (3.12) могут быть вычислены значения математических ожиданий  $m_{gi}$  и дисперсий  $D_{gi}$ ,  $i = \overline{1, 2}$  для каждой из гипотез на основе общих соотношений, представленных в [3, 77]:

$$\begin{aligned}
m_{g_1} &= M[g / \omega_1] = \frac{1}{2} \text{tr}(\mathbf{C}_2^{-1} \mathbf{C}_1 - \mathbf{I}) + \frac{1}{2} (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2)^\top \mathbf{C}_2^{-1} (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2) - \frac{1}{2} \ln \frac{|\mathbf{C}_1|}{|\mathbf{C}_2|}, \\
D_{g_1} &= M[(g - m_{g_1})^2 / \omega_1] = \frac{1}{2} \text{tr}(\mathbf{C}_2^{-1} \mathbf{C}_1 - \mathbf{I})^2 + (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2)^\top \mathbf{C}_2^{-1} \mathbf{C}_1 \mathbf{C}_2^{-1} (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2), \\
m_{g_2} &= M[g / \omega_2] = \frac{1}{2} \text{tr}(\mathbf{I} - \mathbf{C}_1^{-1} \mathbf{C}_2) - \frac{1}{2} (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2)^\top \mathbf{C}_1^{-1} (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2) + \frac{1}{2} \ln \frac{|\mathbf{C}_2|}{|\mathbf{C}_1|}, \\
D_{g_2} &= M[(g - m_{g_2})^2 / \omega_2] = \frac{1}{2} \text{tr}(\mathbf{C}_1^{-1} \mathbf{C}_2 - \mathbf{I})^2 + (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2)^\top \mathbf{C}_1^{-1} \mathbf{C}_2 \mathbf{C}_1^{-1} (\hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2), \\
p(g / \omega_1) &\approx N(g, m_{g_1}, D_{g_1}), \quad p(g / \omega_2) \approx N(g, m_{g_2}, D_{g_2}).
\end{aligned} \tag{3.13}$$

На этой основе в гауссовском приближении можно определить вероятности ошибок первого и второго рода, а также суммарной ошибки в виде

$$\alpha \cong \int_{-\infty}^{l'_0} N(g, m_{g_1}, D_{g_1}) dg, \quad \beta \cong \int_{l'_0}^{\infty} N(g, m_{g_2}, D_{g_2}) dg, \quad E_r \cong p(\omega_1) \alpha + p(\omega_2) \beta. \tag{3.14}$$

### 3.1.3 Результаты экспериментальных исследований алгоритмов распознавания цифровых сигналов в условиях деформирующих искажений

В ходе экспериментальных исследований, описание и результаты которых были опубликованы автором в [65], проводился анализ алгоритмов распознавания сигналов сложной формы, имеющих вид гармонических рядов или полиномов с варьируемыми коэффициентами. С точки зрения возможностей реализации оптимального и квазиоптимального алгоритмов распознавания были рассмотрены две принципиально различающиеся ситуации: воздействие ДИ квазидетерминированного типа, при которой они случайным образом могут принимать форму одной из конечного числа детерминированных функций, и воздействие ДИ, являющихся реализациями случайной функции (нестационарного случайного процесса). Для моделирования различных вариантов построения алгоритмов распознавания использованы программные средства, разработанные в среде MATLAB.

**Воздействие квазидетерминированных деформирующих искажений.** В данном случае предполагается, что форма ДИ в каждой реализации определяется набором функций  $A_r(t)$  детерминированного вида и случайным параметром  $r_x$ , принимающим конечный ряд значений с заданными вероятностями. При этом случайная функция  $A(t)$  представлена как

$$A(t) = \sum_{r=1}^R A_r(t) \tilde{\delta}(r - r_x), \quad \tilde{\delta}(r - r_x) = \begin{cases} 1, & r = r_x, \\ 0, & r \neq r_x, \end{cases} \quad (3.15)$$

где  $\tilde{\delta}(r - r_x)$  дискретная дельта-функция Кронекера. Таким образом, при дискретном представлении получается  $R$  вариантов ДИ:  $\mathbf{a}_r = (a_{r,1}, \dots, a_{r,n})^T$ ,  $r = \overline{1, R}$ , которые порождают  $R$  вариантов размещений с повторениями  $\{\psi_r, r = \overline{1, R}\}$ . Для них в интересах реализации оптимального (3.7) – (3.9) и квазиоптимального (3.10) – (3.12) алгоритмов могут быть определены вероятности  $P_a(\psi)$  как  $P_a(\psi_r) = \Pr(r_x = r)$ ,  $r = \overline{1, R}$ . Для анализа ошибок может быть применен метод прямого статистического имитационного моделирования, а также аналитические соотношения (3.13) и (3.14), основанные на гауссовском приближении ФП.

Рассмотрим пример, в котором используется 4 варианта задания ДИ, возникающие с равной вероятностью и имеющие вид, представленный на рисунке 3.1. В качестве исходных сигналов рассматриваются сигналы  $S_i(t)$ , форма которых определяется гармоническим рядом с варьируемыми коэффициентами. На рисунке 3.2а представлены исходные сигналы и пример их деформации для значительно отличающихся друг от друга сигналов, а на рисунке 3.2б — для сигналов с достаточно близкой формой. Для гарантированного невыхода из области определения в соответствии с условием утверждения 2.4 на функцию ДИ дополнительно накладывалась оконная функция трапецевидной формы (2.11).

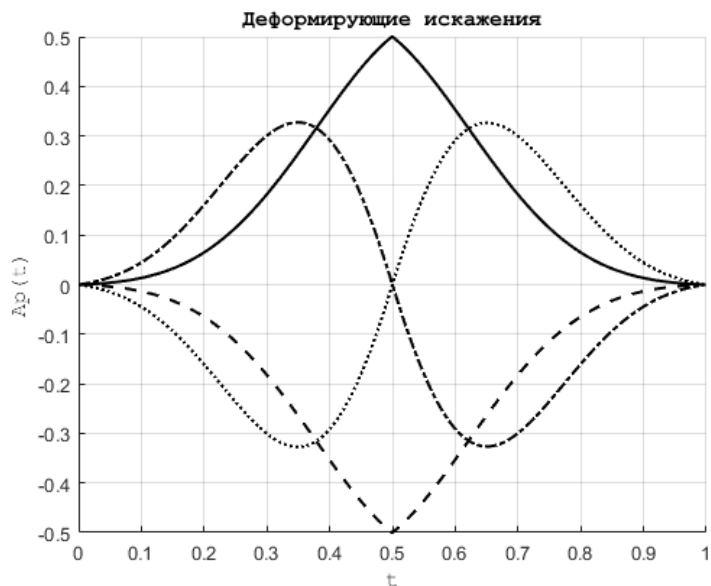


Рисунок 3.1 – Форма возможных вариантов функции деформации

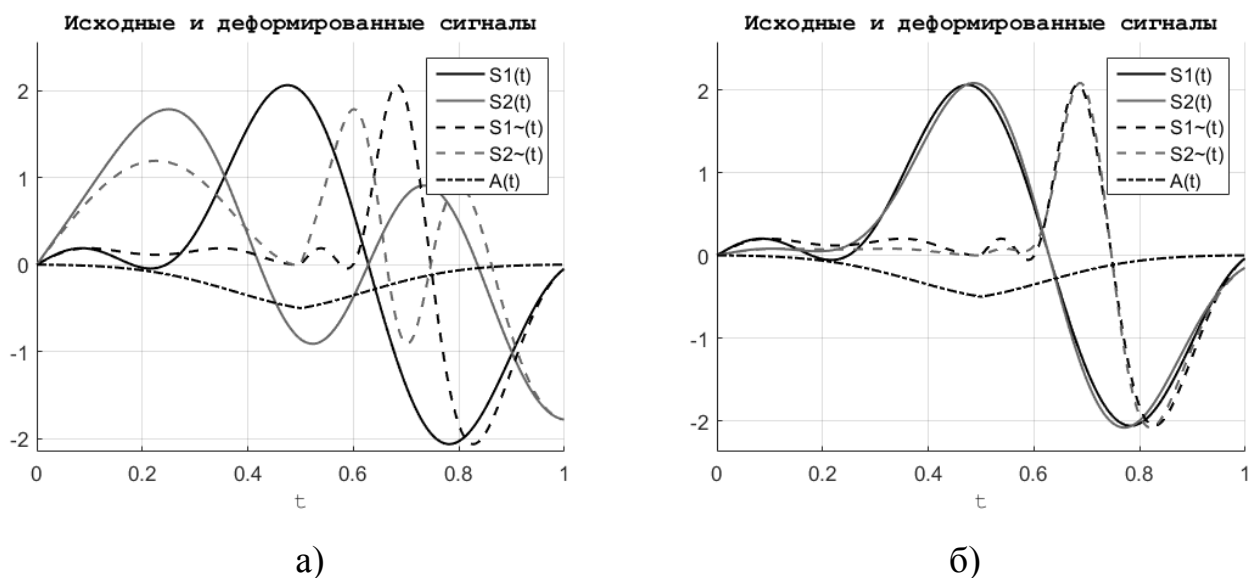


Рисунок 3.2 – Исходные сигналы и примеры их деформации в случае квазидетерминированных ДИ для значительно отличающихся по форме сигналов (а) и для сигналов с достаточно близкой формой (б)

Результаты экспериментальной и теоретической оценки суммарной вероятности ошибки распознавания  $E_r$  в условиях воздействия ДИ вида (3.15) и аддитивного гауссовского шума для данных примеров в виде ее зависимости от величины отношения сигнал-шум представлены на рисунках 3.3 и 3.4. Здесь и далее

в тексте главы в случае наличия малых ошибок графики приводятся в двух видах друг под другом: в линейной шкале, чтобы наглядно отобразить общий ход зависимостей, и в логарифмической шкале, когда требуется отобразить отличие алгоритмов при малых ошибках.

Зависимости получены при числе дискретных отсчетов  $n=16$  и фиксированном  $T=1$  и  $\Delta t = T / (n-1)$ . Обозначения на графиках, вынесенные в текстовую дорожку, соответствуют ранее введенным параметрам, в том числе, обозначение  $A_m = \max A_r(t)$ ,  $r = \overline{1, R}$  определяет максимальное значение амплитуды функции деформации, а обозначение  $r_0$  определяет коэффициент корреляции для аддитивного гауссовского шума, генерируемого при задании матрицы ковариаций вида  $\mathbf{C} = \sigma^2 \| r_0^{|k-l|} \|$ ,  $k = \overline{1, n}$ ,  $l = \overline{1, n}$ . Отношение сигнал-шум для полученных зависимостей вводилось как корень из отношения усредненной «энергии» разностного сигнала к дисперсии аддитивного шума в виде

$$\rho_r = \frac{1}{\sigma} \left[ \frac{1}{n} \sum_{k=1}^n (s_{1,k} - s_{2,k})^2 \right]^{1/2}.$$

На рисунке 3.3 представлены зависимости для сигналов со значительно отличающейся формой (рисунок 3.2а) в условиях наличия ( $A_m = 0,5$ ) и отсутствия ДИ ( $A_m = 0$ ). На рисунке 3.4 показаны результаты моделирования алгоритмов для двух значений коэффициента корреляции аддитивного гауссовского шума для исходных сигналов достаточно близкой формы (рисунок 3.2б). Сплошной линией на графиках представлены зависимости для теоретической оценки вероятности ошибки (3.14). Пунктирными — зависимости, полученные в результате имитационного моделирования с усреднением ошибки по 10000 реализациям.

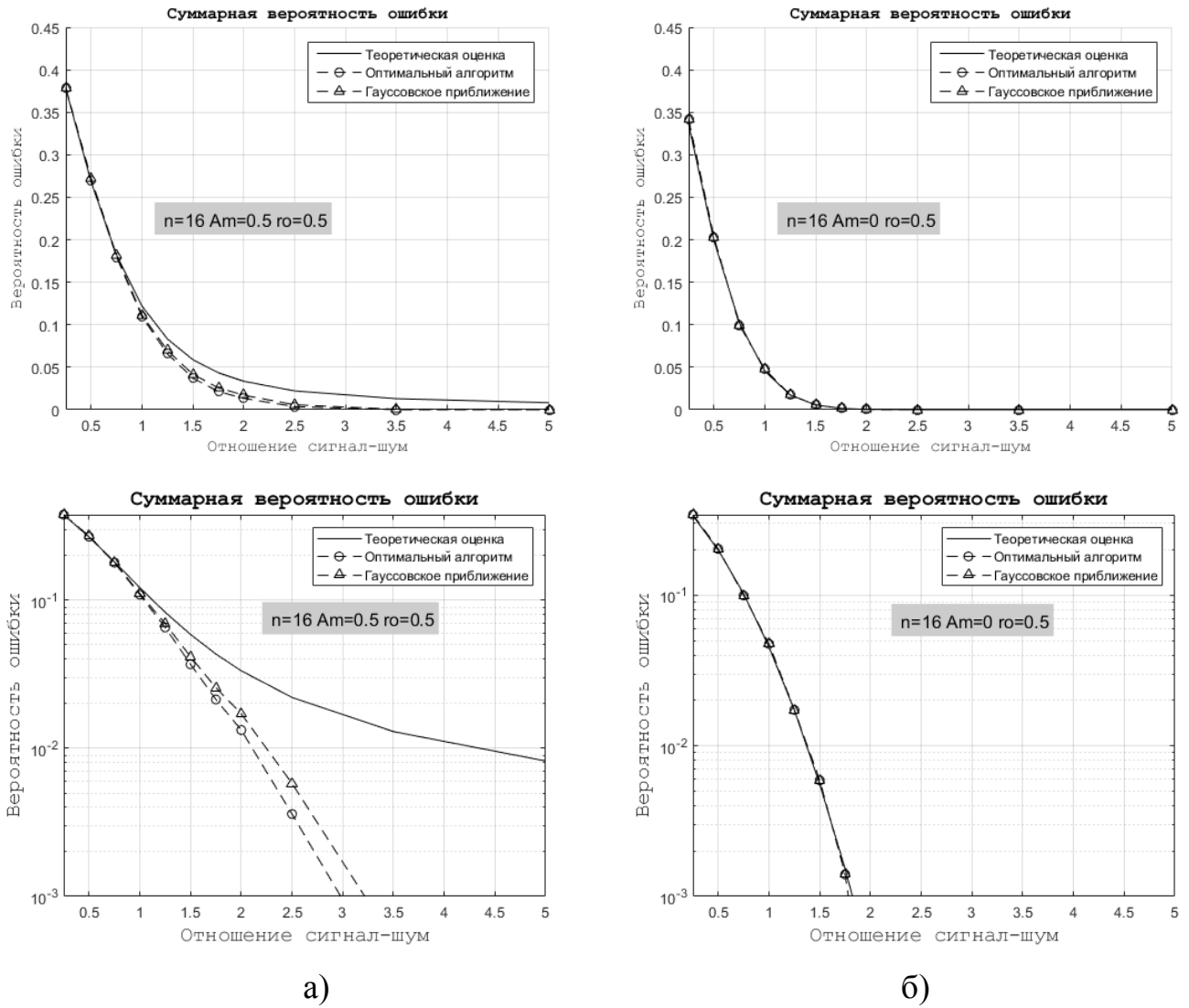
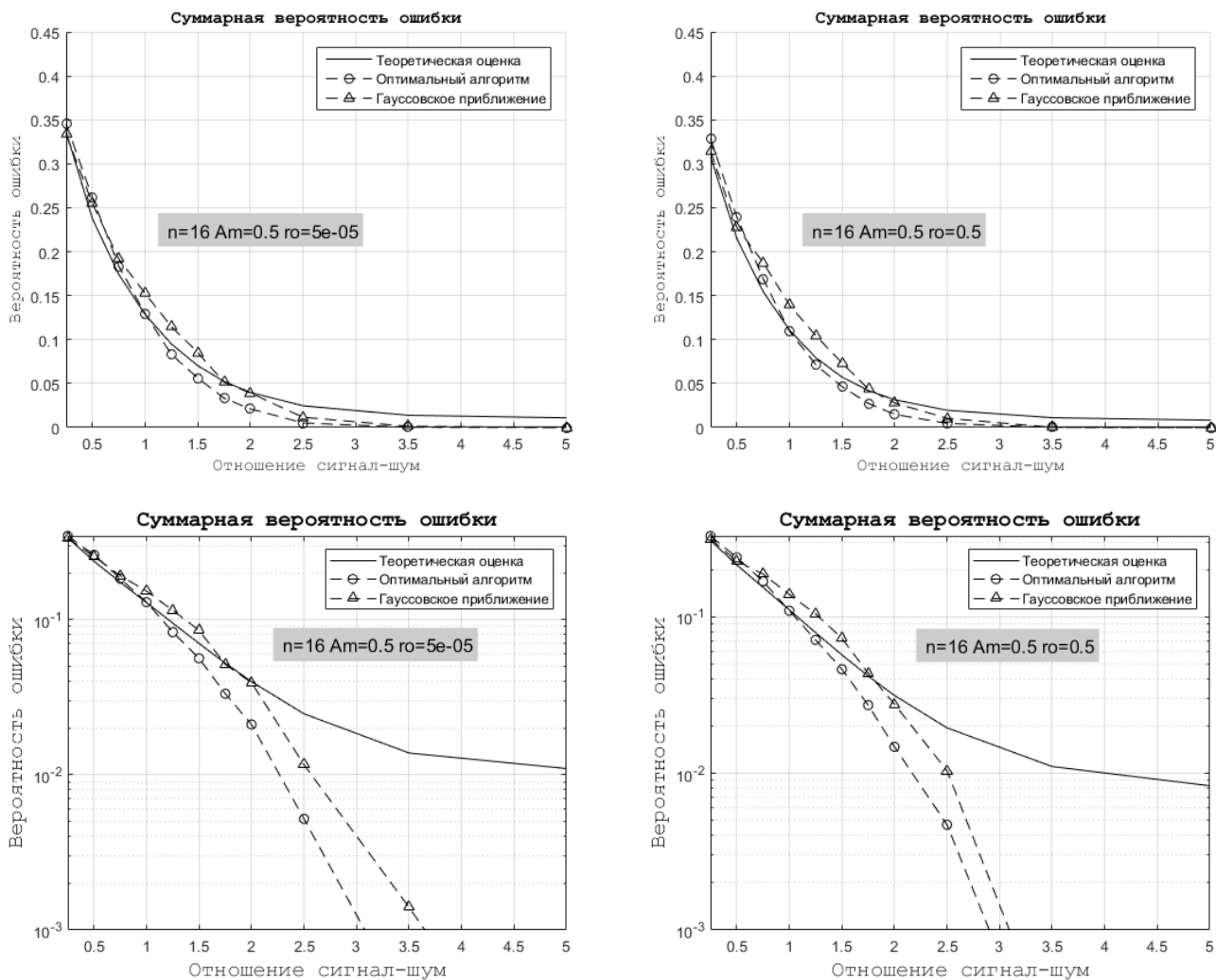


Рисунок 3.3 – Зависимости для оценок вероятностей ошибки от отношения сигнал-шум  $\rho$ , для сигналов со значительно отличающейся формой



а)

б)

Рисунок 3.4 – Зависимости для оценок вероятностей ошибки от отношения сигнал-шум  $\rho$ , для исходных сигналов достаточно близкой формы

**Воздействие случайных деформирующих искажений.** В данном случае предполагается, что форма функции ДИ является реализацией случайной функции (случайного процесса) и имеет бесконечное число вариантов. При этом, в результате дискретизации по времени (3.2) для введенной выше вектор-функции ДИ получается конечное, ограниченное сверху число комбинаций ее значений (утверждение 2.4 из раздела 2.3.1), возникающих при действии оператора размещения с повторениями.



Сама случайная функция  $A(t)$  в общем случае может быть представлена своей плотностью распределения. Рассмотрен вариант, когда это распределение является гауссовским. Тогда для описания случайной функции деформации достаточно задать математическое ожидание  $m(t) = M[A(t)]$  и функцию ковариации  $B(t, \tau) = M([A(t) - m(t)][A(t + \tau) - m(t + \tau)])$ ,  $0 \leq t \leq T$ . Пусть, для определенности,  $B(\tau) = A_m^2 \exp[-\alpha\tau^2]$ , где  $A_m^2$  — дисперсия случайной функции  $A(t)$ , а  $\alpha$  — параметр, определяющий величину ковариации.

Главной особенностью данной ситуации является наличие весьма значительного числа вариантов размещений с повторениями, для каждого из которых в случае непосредственного расчета ФП на основе (3.7) – (3.9) должны быть определены значения вероятностей  $P_a(\psi)$  и плотностей  $p(\mathbf{x}/\psi, \omega_i)$ ,  $i = 1, 2$ . Так, уже при количестве отсчетов  $n = 10$  это число имеет порядок  $10^5$ , и время, затрачиваемое на вычисление ФП на основе (3.7) – (3.9), занимает порядка суток уже для 2000 реализаций алгоритма. Таким образом, в данном случае актуальным становится применение алгоритма (3.10) – (3.12), основанного на использовании гауссовского приближения с заранее рассчитанными векторами математических ожиданий и матриц ковариаций классов.

На рисунке 3.5 представлены исходные сигналы и пример их деформации при внесении ДИ на основе реализаций указанного вида случайных функций.

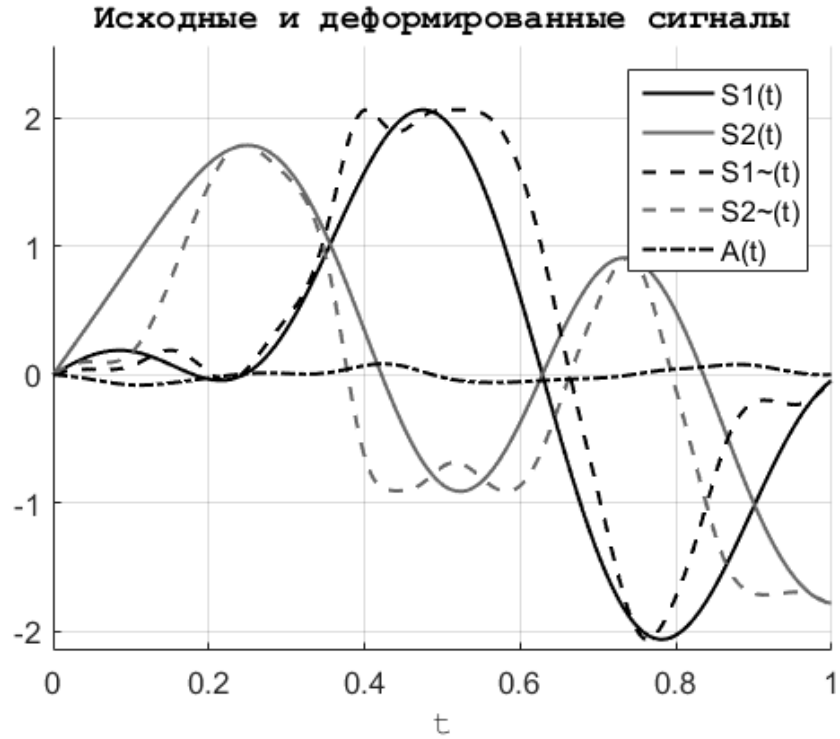


Рисунок 3.5 – Исходные сигналы и пример их деформации в условиях случайных деформирующих искажений

На рисунке 3.6а приведены зависимости вероятности суммарной ошибки распознавания  $E_r$  от величины отношения сигнал-шум  $\rho_r$  в условиях случайных ДИ для числа дискретных отсчетов  $n=10$  при фиксированном  $T=1$  и  $\Delta t = T / (n-1)$ . Как и в предыдущем случае, сплошной линией на графиках отмечены зависимости для теоретической оценки вероятности ошибки. Пунктирными — зависимости, полученные в результате имитационного моделирования с усреднением по 2000 реализациям. Обозначения на графиках, вынесенные в текстовую дорожку, соответствуют ранее введенным параметрам ковариационной функции  $B(\tau)$  ( $A_m$  соответствует  $A_m$  — корню из дисперсии  $A(t)$ , и  $alf$  —  $\alpha$ ). Величина  $ro$ , как и в прошлый раз, обозначает коэффициент корреляции для аддитивного гауссовского шума. На рисунке 3.6б приведены аналогичные зависимости, полученные в условиях отсутствия ДИ ( $A_m = 0$ ).

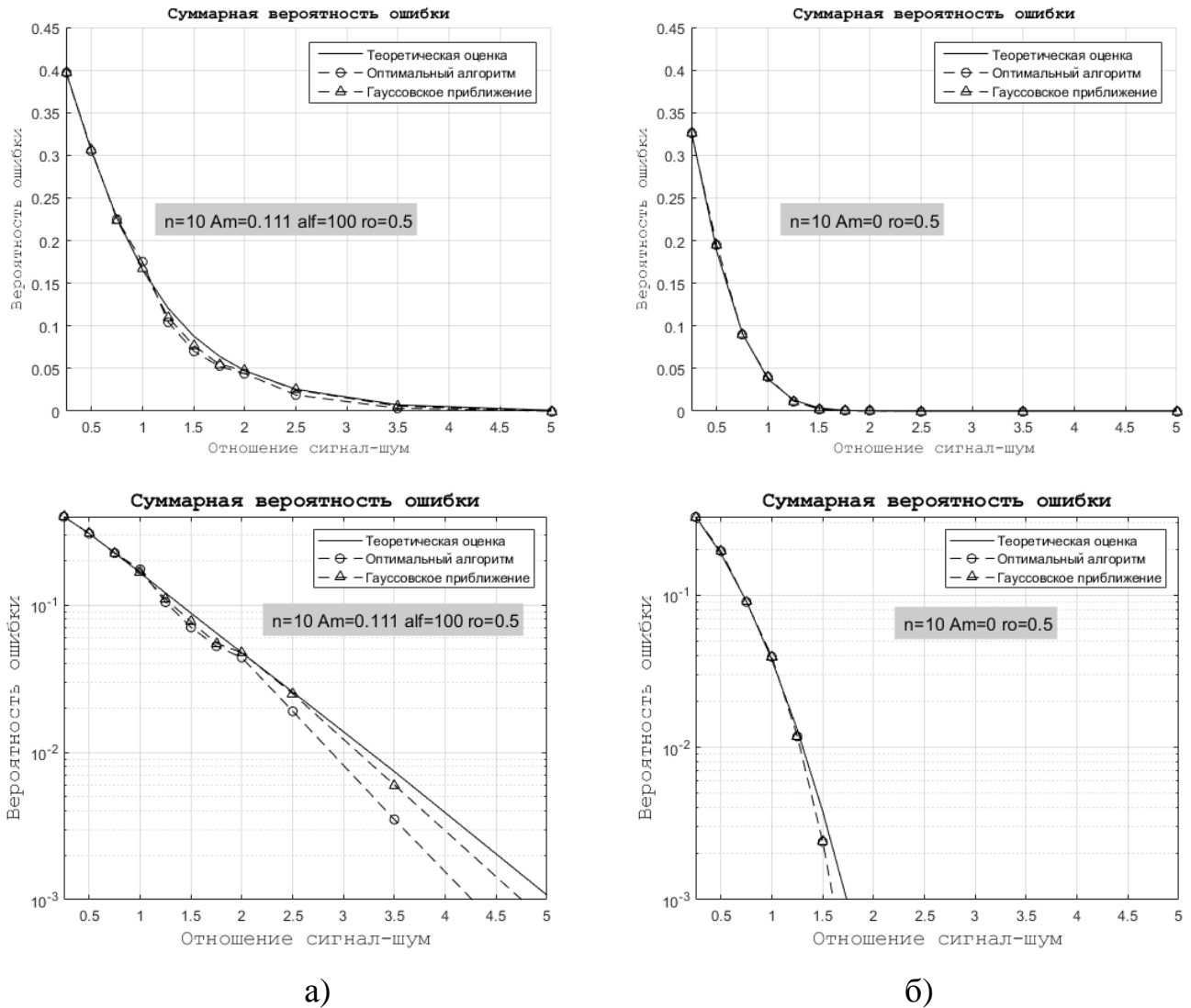


Рисунок 3.6 – Зависимости для оценок вероятностей ошибки от отношения сигнал-шум в условиях случайных деформирующих искажений (а) и в условиях отсутствия деформирующих искажений (б)

**Обсуждение результатов.** Анализ полученных результатов позволил сделать ряд выводов и выявить определенные закономерности.

Прежде всего, следует отметить, что во всех приведенных примерах наблюдается определенное преимущество оптимального алгоритма по отношению к квазиоптимальному, основанному на использовании гауссовского приближения, алгоритму. Это преимущество может быть как незначительным (не более 5-10%)

(рисунок 3.3а и рисунок 3.6а), если распознаваемые сигналы имеют большие различия формы, так и достаточно существенным в противном случае (рисунок 3.4) и может составлять при этом до 2-2,5 раз. Такое расхождение, как показывают многочисленные эксперименты, возникает в случае близости формы исходных распознаваемых сигналов, когда функция их разности становится сопоставимой с произведениями их производных на значения функции деформации (рисунок 3.2б).

Также эксперименты показали, что значение коэффициента корреляции аддитивного гауссовского шума не оказывает существенного влияния на степень превосходства оптимального алгоритма по отношению к квазиоптимальному алгоритму. При этом уменьшение значения коэффициента корреляции увеличивает вероятность ошибок распознавания для обоих алгоритмов.

Принципиальное значение имеют выводы относительно влияния ДИ на общее качество обработки информации, основанные на сравнительном анализе представленных графиков (рисунок 3.3 и рисунок 3.6). Они заключаются в том, что влияние ДИ проявляется при всех отношениях сигнал-шум в виде увеличения суммарной вероятности ошибки, но имеет при этом нелинейный, «кумулятивный» характер, обусловленный природой воздействия ДИ на вид плотностей  $p(\mathbf{x} / \omega_i)$ : их «растяжение» влево-вправо. Установлено, что влияние ДИ в соотношении вероятностей ошибок будет проявляться в большей степени при увеличении отношения сигнал-шум. Так, например, как видно на рисунке 3.6, при отношении сигнал-шум  $\rho_r = 0,5$  в условиях ДИ вероятность суммарной ошибки составляет величину порядка 0,3, а при отсутствии ДИ для тех же условий — менее 0,2. При этом, при более высоких значениях  $\rho_r = 1,5 \dots 2$  в условиях ДИ вероятность ошибки составляет от 0,08 до 0,05, тогда как в отсутствии ДИ вероятность ошибки не превышает 0,01.

Полученные результаты позволяют провести также анализ адекватности предложенных моделей сигналов и помех на основе сравнения зависимостей для

экспериментальных и теоретических оценок (3.13) и (3.14) вероятностей ошибки для квазиоптимального алгоритма с использованием гауссовского приближения ФП. При малых  $n = 6$  и  $n = 8$  здесь наблюдалось заметное расхождение этих оценок. Это, на наш взгляд, можно объяснить тем, что в данных случаях при сложной форме исходных сигналов и еще большего ее усложнения в результате применения ДИ в (3.2) не выполнялись условия теоремы Котельникова. Однако, уже при больших значениях  $n$  зависимости для теоретических и экспериментальных оценок становятся достаточно близкими к друг к другу. При этом теоретические оценки при увеличении отношения сигнал-шум практически всегда превышают экспериментальные, что обеспечивают получение гарантированного результата. Очевидно, также, что точное совпадение этих оценок невозможно, так как соотношения (3.13) и (3.14) сами по себе являются приближенными [3, 77].

На основе полученных результатов можно сделать ряд итоговых рекомендаций относительно практической применимости предложенных моделей и синтезированных алгоритмов распознавания цифровых сигналов в условиях ДИ.

Установлено, что применение оптимального алгоритма, основанного на вычислении ФП классов как взвешенной суммы условных плотностей распределения для возможных комбинаций перестановок элементов сигналов, в большинстве случаев дает относительно небольшое преимущество по отношению к алгоритму, основанному на использовании гауссовского приближения ФП с вычислением математических ожиданий и матриц ковариаций в соответствии с полученными соотношениями. В дополнение к этому, при воздействии ДИ, являющихся реализациями случайных функций, использование оптимального алгоритма связано с необходимостью проведения значительных по объему вычислений непосредственно в процессе принятия решений, что делает его практическое применение затруднительным, тогда как применение квазиоптимального алгоритма позволяет выполнить все необходимые вычисления предварительно и не вызывает затруднений с точки зрения практической реализации.

В то же время, в определенных ситуациях, для которых характерны малые отличия формы распознаваемых сигналов, как показывают приведенные примеры, оптимальный алгоритм может давать заметное преимущество по сравнению с алгоритмом, основанным на использовании гауссовского приближения. В подобных случаях целесообразно переходить к применению непараметрических оценок для восстановления функций правдоподобия классов или к применению других методов машинного обучения на основе использования искусственно сгенерированной выборки обучающих данных путем внесения деформирующих искажений в исходные сигналы в соответствии с используемой статистической моделью.

### **3.2 Синтез и анализ алгоритмов распознавания цифровых изображений в условиях деформирующих и аддитивных искажений**

#### **3.2.1 Исходная статистическая модель и оптимальный алгоритм распознавания**

Пусть без существенного ограничения общности на множестве значений  $\Omega_{xy} = \Omega_x \times \Omega_y$ ,  $\Omega_x = [0, x_{\max}]$ ,  $\Omega_y = [0, y_{\max}]$  заданы изображения известной формы, определяющие эталонные описания двух классов образов  $\omega_1 : S_1(x, y)$ ,  $\omega_2 : S_2(x, y)$ ,  $(x, y) \in \Omega_{xy}$ . (Следует отметить, что здесь, также как и ранее для сигналов в начале главы, весь последующий анализ не ограничивает возможность рассмотрения случая многих классов по аналогичной схеме.)

Наблюдению доступны реализации случайных функций, формируемые на основе этих изображений при воздействии ДИ и аддитивного гауссовского шума и имеющие вид

$$U_i(x, y) = \tilde{S}_i(x, y) + V(x, y) = S_i[x + A_x(x, y), y + A_y(x, y)] + V(x, y), \quad i = \overline{1, 2},$$

где  $U_i(x, y)$  — реализация случайной функции, описывающей наблюдаемое изображение;  $\tilde{S}_i(x, y)$  — реализация случайной функции, описывающей деформированное изображение;  $A(x, y) = (A_x(x, y), A_y(x, y))^T$  — однозначная дифференцируемая вектор-функция координат, являющаяся реализацией векторного случайного поля, описывающего ДИ вдоль осей изображений  $x$  и  $y$  соответственно, такая, что  $(x + A_x(x, y), y + A_y(x, y))^T \in \Omega_{xy}$ ;  $V(x, y)$  — реализация скалярного случайного поля, описывающего действие аддитивного гауссовского шума с параметрами  $M[V(x, y)] = 0$ ,  $M[V(x, y)V(x', y')^T] = \sigma^2 R(x - x', y - y')$ .

Векторное случайное поле  $A(x, y) = (A_x(x, y), A_y(x, y))^T$  в общем случае может быть представлено своей плотностью распределения, а в простейшем случае математическим ожиданием  $m(x, y) = M[A(x, y)]$ ,  $(x, y) \in \Omega_{xy}$  и функцией корреляции,  $B(x, x', y, y') = M[(A(x, y) - m(x, y))(A(x', y') - m(x', y'))^T]$ .

Без существенного ограничения общности предположим, что форма функции ДИ по каждой координате определяется реализациями случайных полей  $A_x(x, y)$ ,  $A_y(x, y)$ . Здесь можно рассмотреть две ситуации: случайные поля  $A_x(x, y)$ ,  $A_y(x, y)$  независимы и, наоборот, полностью зависимы, т.е. фактическое воздействие осуществляется общим случайным полем  $A_{xy}(x, y) = A_x(x, y) = A_y(x, y)$  по каждой координате. Далее будем рассматривать первый вариант — независимая деформация по каждой координате. Реализации  $A_{x(y)}(x, y)$  будем задавать как произведения реализаций однородного случайного поля  $A_{0x(y)}(x, y)$  с заданной функцией корреляции  $B_0(x - x', y - y')$  на оконную функцию  $O_{x(y)}(x(y))$ , используемую для ограничения значений функции деформации,

$$A(x, y) = (\mathbf{O}_x(x)A_x(x, y), \mathbf{O}_y(y)A_y(x, y))^T,$$

$$\mathbf{O}_{x(y)}(x(y)) = \begin{cases} \frac{x(y)}{A_m}, & 0 \leq x(y) < A_m, \\ 1, & A_m \leq x(y) \leq x(y)_{\max} - A_m, \\ \frac{x(y)_{\max} - x(y)}{A_m}, & x(y)_{\max} - A_m < x(y) \leq x(y)_{\max}. \end{cases}$$

где  $A_m^2$  – дисперсия случайного поля  $A_0(x, y)$ .

При цифровой обработке в результате дискретизации с интервалами  $\Delta x$  и  $\Delta y$  формируются цифровые эквиваленты наблюдаемых в непрерывном времени изображений путем фиксации их значений по координатам  $x_k = (k-1)\Delta x$ ,  $k = \overline{1, m_x}$ ,  $\Delta x = x_{\max} / (m_x - 1)$ ,  $x_1 = 0$ ,  $x_{m_x} = x_{\max}$  и  $y_p = (p-1)\Delta y$ ,  $p = \overline{1, n_y}$ ,  $\Delta y = y_{\max} / (n_y - 1)$ ,  $y_1 = 0$ ,  $y_{n_y} = y_{\max}$ :

$$u_i[k, p] = U_i(x_k, y_p) = \tilde{S}_i(x_k, y_p) + V(x_k, y_p) = \\ = S_i[x_k + A_x(x_k, y_p), y_p + A_y(x_k, y_p)] + V(x_k, y_p), \quad i = \overline{1, 2},$$

заданные на двумерной дискретной сетке  $\Omega_{kp} = \{k = \overline{1, m_x}, p = \overline{1, n_y}\}$ . В качестве эквивалентного описания этих изображений и случайных полей будем использовать их представление в виде векторов

$$\mathbf{u}_1 \in \mathbf{R}^n, \quad \mathbf{u}_2 \in \mathbf{R}^n, \quad \mathbf{s}_1 \in \mathbf{R}^n, \quad \mathbf{s}_2 \in \mathbf{R}^n, \quad \tilde{\mathbf{s}}_1 \in \mathbf{R}^n, \quad \tilde{\mathbf{s}}_2 \in \mathbf{R}^n,$$

$$\mathbf{a} \in \mathbf{R}^n, \quad \mathbf{v} \in \mathbf{R}^n, \quad n = m_x \times n_y,$$

полученных в ходе развертки исходных функций по столбцам [78]. Для функции внесения ДИ:  $A(x, y) = (A_x(x, y), A_y(x, y))^T$ , соответствующий развернутый вектор  $\mathbf{a}$  определяется соотношением:  $\tilde{s}_i[x_k] = s_i[x_k + a_k]$ .

Пусть исходные изображения появляются на входе распознающей системы с априорными вероятностями  $p(\omega_i)$ ,  $i = \overline{1, 2}$  в виде вектора наблюдений  $\mathbf{x} = (x_1, \dots, x_n)^T$ , принадлежащего к одному из двух классов  $\omega_1: \mathbf{x} = \mathbf{u}_1$ ,  $\omega_2: \mathbf{x} = \mathbf{u}_2$ . Требуется



сформировать решение в виде (3.3), относящее каждое наблюдение к тому или иному классу. Как и в случае с сигналами, далее будем считать априорные вероятности одинаковыми ( $I_0 = 1$ ).

Анализ результатов, представленных в разделе 3.1.3, показывает, что проведение строгого статистического синтеза алгоритмов распознавания сигналов наталкивается на определенные трудности, которые связаны с вычислительной сложностью при использовании многомерных плотностей распределения классов образов в виде взвешенных сумм условных плотностей распределения по числу комбинаций допустимых перестановок, вызванных внесением ДИ.

Еще в большей степени эта ситуация усугубляется в задачах обработки изображений, где в случае независимости случайных полей  $A_x(x, y)$ ,  $A_y(x, y)$  количество комбинаций перестановок элементов изображения составляет величину равную произведению числа возможных размещений по каждой его оси. В таких условиях работа оптимального алгоритма связана с необходимостью проведения значительных по объему вычислений, что делает его применение для обработки изображений практически невозможным. Кроме того, синтезируемые на этой основе алгоритмы оказываются чувствительными по отношению к адекватности используемых моделей данных, которые основаны на введении многочисленных допущений и приближений. Поэтому, в качестве альтернативного подхода к решению задачи целесообразно перейти к применению непараметрических ядерных оценок для восстановления ФП классов или к применению методов машинного обучения на основе использования искусственно размноженной выборки обучающих данных, полученной путем внесения ДИ в изображения исходной выборки в соответствии с используемой статистической моделью.

### 3.2.2 Использование непараметрических смешанных ядерных оценок

Для получения оценок ФП (3.3) без введения каких-либо ограничений, допущений и приближений можно воспользоваться методами непараметрического оценивания многомерных плотностей распределения вероятностей классов на основе генерации обучающего множества данных по каждому из них. Решение задачи подобным образом имеет смысл для случая воздействия ДИ как реализаций случайного поля с известными статистическими характеристиками.

Пусть для проведения оценок плотностей распределения вероятностей  $p(\mathbf{x} / \omega_i)$ ,  $i = \overline{1, 2}$  по каждому из классов заданы обучающие данные, полученные при внесении ДИ как реализаций случайного поля с известными характеристиками,

$$U^{H_i} = \{\mathbf{u}_i^{(1)}, \dots, \mathbf{u}_i^{(H_i)}\}, \mathbf{u}_i^{(h)} = (u_{i,1}^{(h)}, \dots, u_{i,n}^{(h)})^T \in \mathbf{R}^n, i = \overline{1, 2}, U^H = \bigcup_{i=1}^2 U^{H_i},$$

где  $U^H$  — общая обучающая выборка;  $H$  — число ее элементов (верхний индекс в скобках соответствует элементу обучающей выборки);  $n$  — размерность вектора наблюдений.

Для синтеза алгоритмов распознавания вместо плотностей  $p(\mathbf{u} / \omega_i)$  в этом случае могут быть использованы их непараметрические оценки  $\tilde{p}(\mathbf{u} / \omega_i)$ , обладающие, желательно, свойствами несмещенности и сходимости по вероятности. Наиболее приемлемые для решаемой задачи оценки, как показывает анализ [77], могут быть получены на основе метода Парзена (ядерные оценки).

Для их использования первоначально необходимо выполнить снижение размерности вектора признаков, по отношению к которому будут строиться оценки ФП классов. Это связано с тем, что большие значения размерности неминуемо ведут к увеличению необходимого объема обучающих данных. Кроме того, оценки Парзена, как и фактически все остальные типы оценок, имеют низкую скорость сходимости при большой размерности данных [77].

Снижение размерности целесообразно осуществить на основе метода главных компонент. В этом случае используется линейное преобразование

$$\mathbf{x} = (x_1, \dots, x_K)^T = \mathbf{Q}_K^T \mathbf{u}, \quad (3.16)$$

где  $\mathbf{Q}_K^T = (q_1, \dots, q_K)^T$  — прямоугольная матрица размера  $K \times n$  усеченного преобразования Карунена-Лоэва, строками которой являются собственные векторы выборочной матрицы ковариации исходной выборки  $q_1, \dots, q_K$ , соответствующие  $K$  ее максимальным собственным значениям [77]. Таким образом, при использовании оценок на основе метода Парзена далее будем рассматривать преобразованный вектор признаков  $\mathbf{x}$  из (3.16) и соответствующие ему обучающие выборки  $X^{H_i} = \{\mathbf{x}^{(i,1)}, \dots, \mathbf{x}^{(i,H_i)}\}$ ,  $i = \overline{1,2}$ .

Для каждого класса в отдельности производится восстановление плотности в виде

$$\tilde{p}(\mathbf{x} / \omega_i) = \frac{1}{H_i h^K} \sum_{k=1}^{H_i} \phi\left(\frac{\mathbf{x} - \mathbf{x}^{(i,k)}}{h}\right), \quad i = \overline{1,2},$$

где  $h^{-K} \phi(\{\mathbf{x} - \mathbf{x}^{(i,k)}\} / h)$  — функция ядра, центрированная относительно каждого вектора обучающей выборки класса;  $h$  — параметр оконной функции. Полученные таким образом оценки затем используются при построении алгоритмов принятия решения на основе (3.3) в виде

$$\frac{H_1}{H_1 + H_2} \tilde{p}(\mathbf{x} / \omega_1) \begin{matrix} > \\ < \end{matrix} \frac{H_2}{H_1 + H_2} \tilde{p}(\mathbf{x} / \omega_2) \begin{matrix} \omega_1 \\ \omega_2 \end{matrix}.$$

В качестве универсальной функции ядра, подходящей для многих случаев, целесообразно использовать функцию вида [77]:

$$\begin{aligned} \frac{1}{h^K} \phi\left(\frac{\mathbf{x} - \mathbf{x}^{(i,k)}}{h}\right) &= \frac{1}{(2\pi)^{K/2} h^K |\Xi_i|^{1/2}} \exp\left(-\frac{(\mathbf{x} - \mathbf{x}^{(i,k)})^T \Xi_i^{-1} (\mathbf{x} - \mathbf{x}^{(i,k)})}{2h^2}\right) = \\ &= N(\mathbf{x}, \mathbf{x}^{(i,k)}, h^2 \Xi_i), \end{aligned} \quad (3.17)$$

где  $N(\mathbf{x}, \mathbf{x}^{(i,k)}, h^2 \Xi_i)$  — обозначение гауссовской плотности распределения с математическим ожиданием  $\mathbf{x}^{(i,k)}$  и матрицей ковариаций  $h^2 \Xi_i$ , где  $\Xi_i$  задается исходя из выборочных оценок матриц ковариаций классов  $\Xi_i = \tilde{C}_i$ ,  $i = \overline{1, 2}$  или в виде диагональной матрицы  $\Xi_i = \mathbf{I}$ .

Для выбора  $h(H_i)$  в [3] рекомендуется использовать соотношение  $h(H_i) = H_i^{-s/K}$ ,  $0 < s < 0,5$ . При большой размерности вектора признаков используются более сложные методики [77, 79], среди которых наибольшее распространение получила методика максимизации функционала правдоподобия  $L(h)$  при реализации метода скользящего контроля (кросс-валидации):

$$L(h) = \ln \prod_{k=1}^{H_i} \tilde{p}(\mathbf{x}^{(i,k)} / \omega_i) = \sum_{k=1}^{H_i} \ln \left( \frac{1}{(H_i - 1)h^K} \sum_{p=1, p \neq k}^{H_i} \phi \left( \frac{\mathbf{x}^{(i,k)} - \mathbf{x}^{(i,p)}}{h} \right) \right) \rightarrow \max_h.$$

В [79] показано, что при использовании ядра с диагональной матрицей  $\Xi_i = \mathbf{I}$  область расположения максимума функции находится в пределах

$$h_i \in \left[ \sqrt{\frac{\bar{r}_{\min}}{K}}, \sqrt{\frac{\bar{r}_{\max}}{K}} \right],$$

где  $\bar{r}_{\min}$ ,  $\bar{r}_{\max}$  — усредненные по элементам обучающей выборки минимальное и максимальное расстояния между ними при попарном сравнении. При наличии существенной неравномерности распределений данных величина этого параметра может задаваться переменной на основе анализа расстояний до  $k$ -ближайших соседей [3, 4].

В интересах решения рассматриваемой задачи распознавания изображений в условиях наличия как деформирующих, так и аддитивных помех, в [67] предлагается модификация алгоритма построения ядерных оценок для случая воздействия аддитивного шума с известной плотностью распределения  $p(v)$ . Она заключается в формировании статистических описаний классов в два этапа.

1. На первом этапе путем генерации множества деформаций изображений в отсутствии аддитивного шума формируются выборки  $\tilde{X}^{H_i} = \{\tilde{\mathbf{x}}^{(i,1)}, \dots, \tilde{\mathbf{x}}^{(i,H_i)}\}$ ,  $i = \overline{1,2}$ , на основе которых строятся ядерные оценки плотностей  $\tilde{p}(\tilde{\mathbf{x}} / \omega_i)$ , где  $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{v}$ .

2. На втором этапе осуществляется свертка полученной оценки распределения вектора признаков  $\tilde{p}(\tilde{\mathbf{x}} / \omega_i)$  с плотностью распределения аддитивного шума в соответствии с соотношением:

$$\hat{p}(\mathbf{x} / \omega_i) = \int \tilde{p}(\mathbf{x} - \mathbf{v} / \omega_i) p(\mathbf{v}) d\mathbf{v}. \quad (3.18)$$

Рассмотрим пример использования ядра вида (3.17) для непараметрических оценок в присутствии аддитивного гауссовского шума, описываемого распределением  $p(\mathbf{v}) = N(\mathbf{v}, 0, \mathbf{C}_v)$ .

Пусть  $\mathbf{x}$  и  $\mathbf{z}$  два гауссовских вектора, причем

$$p_x(\mathbf{x}) = N(\mathbf{x}, \mathbf{m}_x, \mathbf{C}_x), \quad p_z(\mathbf{z} / \mathbf{x}) = N(\mathbf{z}, \mathbf{H}\mathbf{x}, \mathbf{C}_z),$$

где  $\mathbf{H}$  — матрица соответствующей размерности. Тогда плотность распределения вектора  $\mathbf{z}$  имеет вид свертки [80]:

$$p_z(\mathbf{z}) = \int N(\mathbf{z}, \mathbf{H}\mathbf{x}, \mathbf{C}_z) N(\mathbf{x}, \mathbf{m}_x, \mathbf{C}_x) d\mathbf{x} = N(\mathbf{z}, \mathbf{H}\mathbf{m}_x, \mathbf{H}\mathbf{C}_x\mathbf{H}^T + \mathbf{C}_z).$$

Отсюда при использовании гауссовского ядра (3.17) и гауссовской плотности аддитивного шума получается оценка в виде

$$\hat{p}(\mathbf{x} / \omega_i) = \frac{1}{H_i} \sum_{k=1}^{H_i} N(\mathbf{x}, \tilde{\mathbf{x}}^{(i,k)}, h^2 \mathbf{\Xi}_i + \mathbf{C}_v). \quad (3.19)$$

Полученные таким образом оценки плотностей распределения будем называть **смешанными**, акцентируя внимание на том, что одна из составляющих смеси (3.18) получена на основе непараметрической оценки по обучающей выборке, а другая изначально задана в аналитическом виде.

### 3.2.3 Анализ свойств смешанных оценок

Как известно [3, 4], стандартные оценки Парзена при определенных условиях являются несмещенными и состоятельными. Пусть для оценки  $\tilde{p}(\mathbf{x} / \omega_i)$  выполняются следующие свойства функции ядра:

$$\frac{1}{h^K} \phi\left(\frac{\mathbf{x} - \mathbf{x}^{(i,k)}}{h}\right) = \frac{1}{h^K} \phi(\mathbf{u} / h), \quad (3.20)$$

$$\int \frac{1}{h^K} \phi(\mathbf{u} / h) d\mathbf{u} = 1, \quad \int \frac{1}{h^K} |\phi(\mathbf{u} / h)| d\mathbf{u} < \infty, \quad \sup |\phi(\mathbf{u} / h)| < \infty,$$

что говорит о том, что ядро удовлетворяет свойствам плотности распределения [77].

Пусть также при задании параметра ядра  $h$  как функции наблюдений  $h(H_i)$ , где  $h^K(H_i) \xrightarrow{H_i \rightarrow \infty} 0$ ,  $H_i h^K(H_i) \xrightarrow{H_i \rightarrow \infty} \infty$  и  $\psi_{H_i}(\mathbf{u}) = h^{-K}(H_i) \phi(\mathbf{u} / h(H_i))$ , для каждой точки непрерывности исходной плотности  $p(\mathbf{x} / \omega_i)$  выполняется

$$\int \psi_{H_i}(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) d\mathbf{s} \xrightarrow{H_i \rightarrow \infty} p(\mathbf{x} / \omega_i), \quad \frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \xrightarrow{H_i \rightarrow \infty} 0.$$

Тогда оценка  $\tilde{p}(\mathbf{x} / \omega_i)$  является состоятельной в среднем квадратичном в каждой точке непрерывности  $p(\mathbf{x} / \omega_i)$  [81].

Отметим, что первое из этих условий означает, что  $\psi_{H_i}(\mathbf{x} - \mathbf{s})$  при  $H_i \rightarrow \infty$  сходится к дельта-функции с центром в точке  $\mathbf{s}$ . Для широкого класса функций, которые могут быть использованы в качестве ядра и обладают свойствами (3.20), такая сходимость имеет место, если  $h^K(H_i) \xrightarrow{H_i \rightarrow \infty} 0$ . Второе условие для таких функций выполняется, если  $H_i h^K(H_i) \xrightarrow{H_i \rightarrow \infty} \infty$  [81].

При этом асимптотическая несмещенность оценки  $\tilde{p}(\mathbf{x} / \omega_i)$ , полученной по совокупности обучающей выборки  $X^{H_i} = \{\mathbf{x}^{(i,1)}, \dots, \mathbf{x}^{(i,H_i)}\}$ , может быть показана на основе следующих преобразований [81]:

$$\begin{aligned}
M[\tilde{p}(\mathbf{x} / \omega_i)] &= \frac{1}{H_i h^K} \sum_{k=1}^{H_i} M \left[ \phi \left( \frac{\mathbf{x} - \mathbf{x}^{(i,k)}}{h} \right) \right] = \frac{1}{h^K} M \left[ \phi \left( \frac{\mathbf{x} - X^{H_i}}{h} \right) \right] = \\
&= \int \frac{1}{h^K} \phi \left( \frac{\mathbf{x} - \mathbf{s}}{h} \right) p(\mathbf{s} / \omega_i) d\mathbf{s} = \int \psi_{H_i}(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) d\mathbf{s} \xrightarrow{H_i \rightarrow \infty} p(\mathbf{x} / \omega_i).
\end{aligned} \tag{3.21}$$

**Утверждение 3.1.** Если плотность распределения  $\tilde{p}(\mathbf{x} / \omega_i)$  при  $h^K(H_i) \xrightarrow{H_i \rightarrow \infty} 0$  и  $H_i h^K(H_i) \xrightarrow{H_i \rightarrow \infty} \infty$  является асимптотически несмещенной оценкой плотности  $p(\mathbf{x} / \omega_i)$  для всех  $\mathbf{x}$ , то и смешанная оценка  $\hat{p}(\mathbf{x} / \omega_i)$  является асимптотически несмещенной оценкой  $p(\mathbf{x} / \omega_i)$ .

Пусть  $X^{H_i} = \{\mathbf{x}^{(i,1)}, \dots, \mathbf{x}^{(i,H_i)}\}$ , где каждый  $\mathbf{x}^{(k)} = \tilde{\mathbf{x}}^{(k)} + \mathbf{v}^{(k)}$ ,  $\tilde{X}^{H_i} = \{\tilde{\mathbf{x}}^{(i,1)}, \dots, \tilde{\mathbf{x}}^{(i,H_i)}\}$ ,  $V = \{\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(H_i)}\}$ . Как показано выше, для стандартной оценки  $\tilde{p}(\mathbf{x} / \omega_i)$  выполняется (3.21). С другой стороны

$$\begin{aligned}
M[\tilde{p}(\mathbf{x} / \omega_i)] &= \frac{1}{H_i h^K} \sum_{k=1}^{H_i} M \left[ \phi \left( \frac{\mathbf{x} - \tilde{\mathbf{x}}^{(i,k)} - \mathbf{v}^{(k)}}{h} \right) \right] = \frac{1}{h^K} M \left[ \phi \left( \frac{\mathbf{x} - \tilde{X}^{H_i} - V}{h} \right) \right] = \\
&= \int \frac{1}{h^K} \phi \left( \frac{\mathbf{x} - \tilde{\mathbf{s}} - \mathbf{v}}{h} \right) p(\tilde{\mathbf{s}} / \omega_i) p(\mathbf{v}) d\tilde{\mathbf{s}} d\mathbf{v} = \int \mathfrak{G}_{H_i}(\mathbf{x} - \tilde{\mathbf{s}}) p(\tilde{\mathbf{s}} / \omega_i) d\tilde{\mathbf{s}} = \\
&= M \left[ \mathfrak{G}_{H_i}(\mathbf{x} - \tilde{X}^{H_i}) \right] = M[\hat{p}(\mathbf{x} / \omega_i)].
\end{aligned}$$

Запись  $\int \mathfrak{G}_{H_i}(\mathbf{x} - \tilde{\mathbf{s}}) p(\tilde{\mathbf{s}} / \omega_i) d\tilde{\mathbf{s}}$  использует функцию, суммируемую в представлении смешанной оценки (3.18), и, следовательно, стандартная оценка имеет то же математическое ожидание, что и смешанная, т.е. для нее также выполняется свойство асимптотической несмещенности.

Как известно, если оценка сходится к истинному значению параметра «в среднем квадратичном» или если оценка является асимптотически несмещенной при стремлении ее дисперсии к нулю, то такая оценка будет состоятельной. Таким образом, состоятельность стандартной оценки  $\tilde{p}(\mathbf{x} / \omega_i)$  вытекает из ее сходимости в

среднеквадратичном, что может быть показано на основе следующих преобразований:

$$\begin{aligned}
D[\tilde{p}(\mathbf{x} / \omega_i)] &= M[\tilde{p}^2(\mathbf{x} / \omega_i)] - p^2(\mathbf{x} / \omega_i), \\
M[\tilde{p}^2(\mathbf{x} / \omega_i)] &= \frac{1}{H_i^2 h^{2K}} \sum_{r=1}^{H_i} \sum_{s=1}^{H_i} M \left[ \phi \left( \frac{\mathbf{x} - \mathbf{x}^{(i,r)}}{h} \right) \phi \left( \frac{\mathbf{x} - \mathbf{x}^{(i,s)}}{h} \right) \right] = \\
&= \frac{1}{H_i^2 h^{2K}} \sum_{r=1}^{H_i} M \left[ \phi^2 \left( \frac{\mathbf{x} - \mathbf{x}^{(i,r)}}{h} \right) \right] + \frac{1}{H_i^2 h^{2K}} \sum_{r=1}^{H_i} \sum_{s=1, s \neq r}^{H_i} M \left[ \phi \left( \frac{\mathbf{x} - \mathbf{x}^{(i,r)}}{h} \right) \phi \left( \frac{\mathbf{x} - \mathbf{x}^{(i,s)}}{h} \right) \right] = \quad (3.22) \\
&= \frac{1}{H_i} \int \frac{1}{h^{2K}} \phi^2 \left( \frac{\mathbf{x} - \mathbf{s}}{h} \right) p(\mathbf{s} / \omega_i) ds + \frac{H_i - 1}{H_i} \left( \int \frac{1}{h^K} \phi \left( \frac{\mathbf{x} - \mathbf{s}}{h} \right) p(\mathbf{s} / \omega_i) ds \right)^2 = \\
&= \frac{1}{H_i} \int \psi_{H_i}^2(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds + \frac{H_i - 1}{H_i} \left( \int \psi_{H_i}(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds \right)^2.
\end{aligned}$$

Теперь, учитывая, что  $\frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \xrightarrow{H_i \rightarrow \infty} 0$ , получим

$$\begin{aligned}
\frac{1}{H_i} \int \psi_{H_i}^2(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds &\leq \frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \int \psi_{H_i}(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds \xrightarrow{H_i \rightarrow \infty} 0, \\
\frac{H_i - 1}{H_i} \left( \int \psi_{H_i}(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds \right)^2 &\xrightarrow{H_i \rightarrow \infty} p^2(\mathbf{x} / \omega_i),
\end{aligned}$$

и, окончательно:

$$D[\tilde{p}(\mathbf{x} / \omega_i)] \xrightarrow{H_i \rightarrow \infty} 0.$$

**Утверждение 3.2.** Если плотность распределения  $\tilde{p}(\mathbf{x} / \omega_i)$  при  $h^K(H_i) \xrightarrow{H_i \rightarrow \infty} 0$  и  $H_i h^K(H_i) \xrightarrow{H_i \rightarrow \infty} \infty$  сходится в среднеквадратичном к плотности  $p(\mathbf{x} / \omega_i)$  для всех  $\mathbf{x}$ , то и смешанная оценка  $\hat{p}(\mathbf{x} / \omega_i)$  аналогичным образом сходится к  $p(\mathbf{x} / \omega_i)$ .

Рассмотрим сначала дисперсию стандартной оценки  $D[\tilde{p}(\mathbf{x} / \omega_i)]$  и на основе (3.22) представим в ней математическое ожидание квадрата оценки в виде

$$M[\tilde{p}^2(\mathbf{x} / \omega_i)] = \frac{1}{H_i} \int \psi_{H_i}^2(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds + \frac{H_i - 1}{H_i} \left( \int \psi_{H_i}(\mathbf{x} - \mathbf{s}) p(\mathbf{s} / \omega_i) ds \right)^2.$$



Для первого слагаемого при  $\frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \xrightarrow{H_i \rightarrow \infty} 0$  выполняется

$$\begin{aligned} & \frac{1}{H_i} \int \psi_{H_i}^2(\mathbf{x} - \check{\mathbf{s}} - \mathbf{v}) p(\check{\mathbf{s}} / \omega_i) p(\mathbf{v}) d\check{\mathbf{s}} d\mathbf{v} \leq \\ & \leq \frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \int \psi_{H_i}(\mathbf{x} - \check{\mathbf{s}} - \mathbf{v}) p(\check{\mathbf{s}} / \omega_i) p(\mathbf{v}) d\check{\mathbf{s}} d\mathbf{v} \xrightarrow{H_i \rightarrow \infty} 0. \end{aligned}$$

Для второго слагаемого в силу сходимости стандартных оценок выполняется

$$\begin{aligned} & \frac{H_i - 1}{H_i} \left( \int \psi_{H_i}(\mathbf{x} - \check{\mathbf{s}} - \mathbf{v}) p(\check{\mathbf{s}} / \omega_i) p(\mathbf{v}) d\check{\mathbf{s}} d\mathbf{v} \right)^2 = \\ & = \frac{H_i - 1}{H_i} \left( \int \vartheta_{H_i}(\mathbf{x} - \check{\mathbf{s}}) p(\check{\mathbf{s}} / \omega_i) d\check{\mathbf{s}} \right)^2 \xrightarrow{H_i \rightarrow \infty} p^2(\mathbf{x} / \omega_i). \end{aligned} \quad (3.23)$$

Теперь рассмотрим дисперсию смешанной оценки

$$D[\hat{p}(\mathbf{x} / \omega_i)] = M[\hat{p}^2(\mathbf{x} / \omega_i)] - p^2(\mathbf{x} / \omega_i).$$

Заметим, сначала, что

$$\begin{aligned} \frac{1}{H_i} \vartheta_{H_i}(\mathbf{x}) &= \frac{1}{H_i} \int \psi_{H_i}(\mathbf{x} - \mathbf{v}) p(\mathbf{v}) d\mathbf{v} \leq \frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \Rightarrow \\ &\Rightarrow \frac{1}{H_i} \sup \vartheta_{H_i}(\mathbf{x}) \leq \frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}). \end{aligned} \quad (3.24)$$

Тогда для смешанной оценки

$$\begin{aligned} M[\hat{p}^2(\mathbf{x} / \omega_i)] &= \frac{1}{H_i^2} \sum_{r=1}^{H_i} \sum_{s=1}^{H_i} M \left[ \vartheta_{H_i}(\mathbf{x} - \check{\mathbf{x}}^{(i,r)}) \vartheta_{H_i}(\mathbf{x} - \check{\mathbf{x}}^{(i,s)}) \right] = \\ &= \frac{1}{H_i^2} \sum_{r=1}^{H_i} M \left[ \vartheta_{H_i}^2(\mathbf{x} - \check{\mathbf{x}}^{(i,r)}) \right] + \frac{1}{H_i^2} \sum_{r=1}^{H_i} \sum_{s=1, s \neq r}^{H_i} M \left[ \vartheta_{H_i}(\mathbf{x} - \check{\mathbf{x}}^{(i,r)}) \vartheta_{H_i}(\mathbf{x} - \check{\mathbf{x}}^{(i,s)}) \right] = \\ &= \frac{1}{H_i} \int \vartheta_{H_i}^2(\mathbf{x} - \check{\mathbf{s}}) p(\check{\mathbf{s}} / \omega_i) d\check{\mathbf{s}} + \frac{H_i - 1}{H_i} \left( \int \vartheta_{H_i}(\mathbf{x} - \check{\mathbf{s}}) p(\check{\mathbf{s}} / \omega_i) d\check{\mathbf{s}} \right)^2. \end{aligned}$$

Первое слагаемое при  $\frac{1}{H_i} \sup \psi_{H_i}(\mathbf{x}) \xrightarrow{H_i \rightarrow \infty} 0$  на основе (3.24) стремится к нулю,

так как

$$\frac{1}{H_i} \int \mathfrak{Q}_{H_i}^2(\mathbf{x} - \tilde{\mathbf{s}}) p(\tilde{\mathbf{s}} / \omega_i) d\tilde{\mathbf{s}} \leq \frac{1}{H_i} \sup \mathfrak{Q}_{H_i}(\mathbf{x}) \int \mathfrak{Q}_{H_i}(\mathbf{x} - \tilde{\mathbf{s}}) p(\tilde{\mathbf{s}} / \omega_i) d\tilde{\mathbf{s}} \xrightarrow{H_i \rightarrow \infty} 0,$$

а для второго слагаемого выполняется (3.23), т.е. математические ожидания квадрата стандартной и смешанной оценок стремятся к одному пределу, определяемому величиной  $p^2(\mathbf{x} / \omega_i)$ . Таким образом,  $D[\hat{p}(\mathbf{x} / \omega_i)] \xrightarrow{H_i \rightarrow \infty} 0$ , что позволяет говорить о том, что смешанная оценка является состоятельной.

Кроме того, в [67] было доказано следующее утверждение.

**Утверждение 3.3.** При известной плотности распределения  $p(\mathbf{v})$ ,  $M[\mathbf{v}] = 0$  и фиксированной выборке  $\tilde{X}^{H_i} = \{\tilde{\mathbf{x}}^{(i,1)}, \dots, \tilde{\mathbf{x}}^{(i,H_i)}\}$  использование смешанной оценки эквивалентно реализации процедуры искусственного размножения обучающих данных, в ходе которой для каждого вектора  $\tilde{\mathbf{x}}^{(i,k)} \in \tilde{X}^{H_i}$ ,  $k = \overline{1, H_i}$  проводится генерация подвыборки, состоящей из  $P_{i,k} \rightarrow \infty$  образцов, на основе распределения  $p(\mathbf{x} / \tilde{\mathbf{x}}^{(i,k)})$ , где  $\mathbf{x} = \tilde{\mathbf{x}}^{(i,k)} + \mathbf{v} + \mathbf{g}$ , где  $\mathbf{g}$  — случайная величина с распределением, плотность которого определяется одним из возможных ядер, используемых при получении смешанной оценки,  $p_{\mathbf{g}}(\mathbf{g}) = h'^{-K} \varphi(\mathbf{g} / h') = \psi_{H_i}(\mathbf{g})$ ,  $M[\mathbf{g}] = 0$ . Здесь  $h' = h'(H_i)$  — значение, выбираемое исходя из заданного  $H_i$ .

Доказательство вытекает из сходимости ядерной оценки  $\tilde{p}(\mathbf{x} / \tilde{\mathbf{x}}^{(i,k)})$ , восстановленной по подвыборке  $X^{(P_{i,k})} = \{\mathbf{x}^{(i,k,1)}, \dots, \mathbf{x}^{(i,k,P_{i,k})}\}$ , к плотности  $\hat{p}(\mathbf{x} / \tilde{\mathbf{x}}^{(i,k)})$  при  $P_{i,k} \rightarrow \infty$ . Следствием данного утверждения является то, что при достаточно больших  $H_i \rightarrow \infty$  искусственное размножение данных может осуществляться как  $\mathbf{x} = \tilde{\mathbf{x}}^{(i,k)} + \mathbf{v}$ , так как  $h' = h'(H_i)$  при этом стремится к нулю, а  $\psi_{H_i}(\mathbf{g})$  — к дельта-функции. Для генерации подвыборки в этом случае может быть использован метод нелинейного функционального преобразования или метод исключений с

соответствующей коррекцией вида плотности  $p(v)$  для учета усеченного характера плотности генерируемых данных.

### **3.2.4 Результаты экспериментальных исследований алгоритмов распознавания цифровых изображений в условиях деформирующих искажений**

Указанные свойства смешанных ядерных оценок позволили предложить схему алгоритма обработки информации в интересах синтеза решающих правил при распознавании сигналов и изображений в случае, когда имеются ограничения вычислительной сложности или отсутствуют необходимые исходные данные, что не позволяет применить оптимальный (3.7) – (3.9) или квазиоптимальный алгоритмы (3.10) – (3.12), основанные на использовании эквивалентной статистической модели ДИ (2.9). Обобщенная структура алгоритма обработки информации представлена на рисунке 3.7. Исследование различных вариантов алгоритмов распознавания изображений при этом проводилось в среде MATLAB методом статистического моделирования. Его результаты были опубликованы автором в работе [65].

Первоначально в ходе экспериментальных исследований производился анализ алгоритмов распознавания искусственно синтезированных изображений, подверженных ДИ и аддитивным помехам. Использовались изображения двух видов: изображения колокольной (гауссовской) функции с различной степенью вытянутости и ориентации (рисунок 3.8 — на графиках далее обозначены как *gaussian*) и изображения, полученные на основе двумерного гармонического ряда (рисунок 3.9 — на графиках далее обозначены как *chess*). Исходные размеры изображений составляли  $21 \times 21$  пиксель.

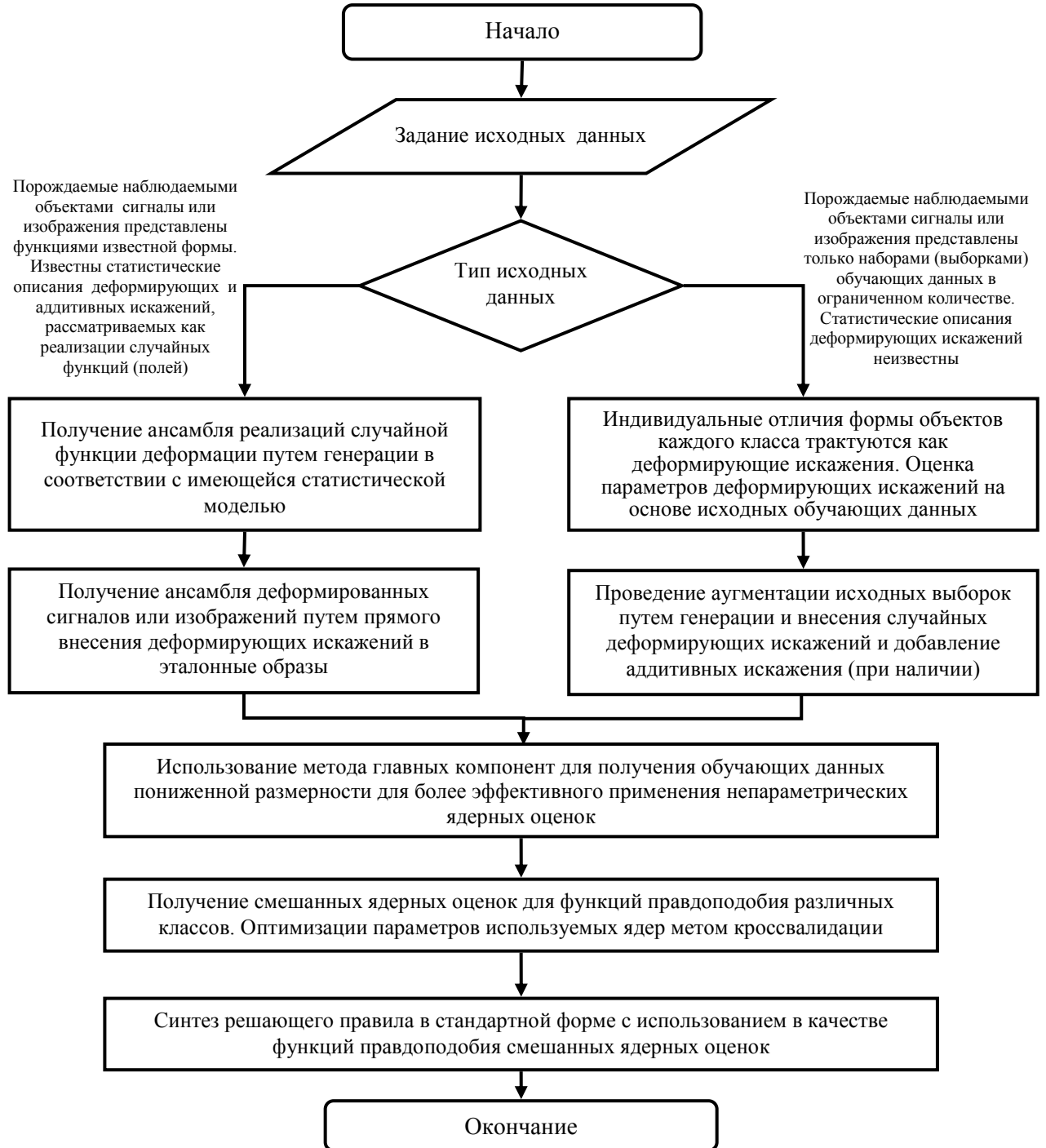


Рисунок 3.7 – Обобщенная структура алгоритма обработки информации

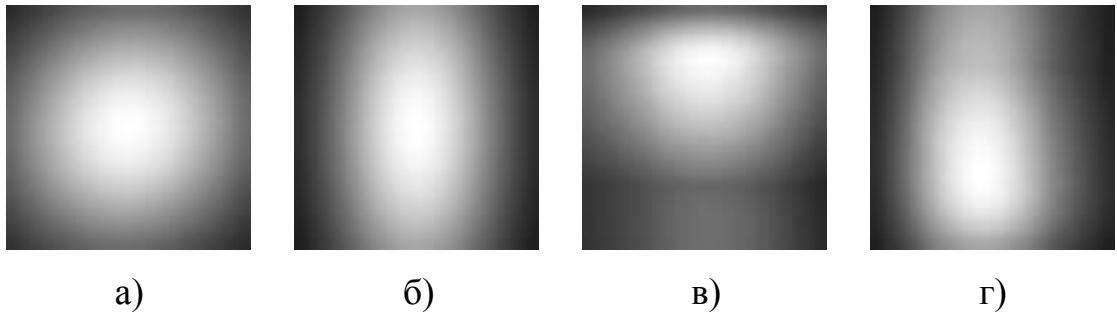


Рисунок 3.8 – Синтезированные изображения (а, б) в виде колокольной (гауссовской) функции и примеры их деформаций (в, г), (в) на основе (а) и (г) на основе (б)

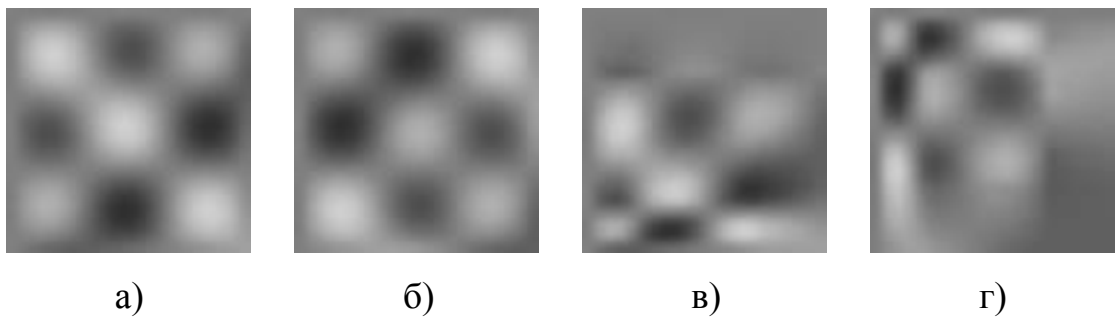


Рисунок 3.9 – Синтезированные изображения (а, б), полученные на основе двумерного гармонического ряда, и примеры их деформаций (в, г), (в) на основе (а) и (г) на основе (б)

В ходе сравнительного исследования алгоритмов распознавания изображений рассматривались следующие варианты их построения:

- алгоритм на основе смешанных непараметрических ядерных оценок  $\hat{p}(\mathbf{x} / \omega_i)$ ,  $i = \overline{1,2}$  (3.19);

- алгоритм на основе стандартных непараметрических ядерных оценок  $\tilde{p}(\mathbf{x} / \omega_i)$ ,  $i = \overline{1,2}$ ;

- алгоритм на основе гауссовского приближения функций правдоподобия  $\hat{p}(\mathbf{x} / \omega_i) = N(\mathbf{x}, \mathbf{m}_i, \mathbf{C}_i)$ ,  $i = \overline{1,2}$  с использованием оценок математических ожиданий и матриц ковариаций классов  $\mathbf{m}_i, \mathbf{C}_i$ ,  $i = \overline{1,2}$ ;

- алгоритм на основе искусственной двуслойной нейронной сети прямого распространения с  $2K$  нейронами в скрытом слое, обучаемой по выборкам  $X^{H_i} = \{\mathbf{x}^{(i,1)}, \dots, \mathbf{x}^{(i,H_i)}\}$ ,  $i = \overline{1,2}$ , где  $K$  соответствует числу главных компонент в (3.16).

В плане сравнительного анализа последний алгоритм представляет особый интерес, поскольку, в отличие от использования смешанной ядерной оценки или классической параметрической оценки, в стандартном варианте его реализации отсутствует возможность учета аддитивного шума (нейросети можно обучать только по исходным выборкам). Поэтому для учета влияния аддитивного шума в рассматриваемой задаче целесообразно дополнительно проводить искусственное размножение данных.

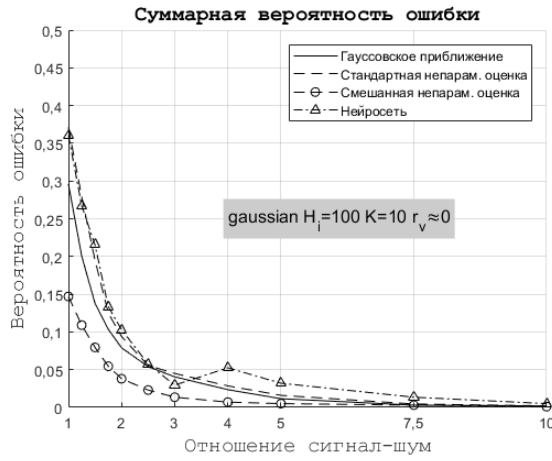
Корреляционные функции вносимых ДИ и аддитивного шума в ходе экспериментов задавались в виде

$$\begin{aligned} B_0(x-x', y-y') &= A_m^2 \exp(-\alpha_0(|x-x'|+|y-y'|)), \\ \sigma^2 R(x-x', y-y') &= \sigma^2 \exp(-\alpha_v(|x-x'|+|y-y'|)), \end{aligned}$$

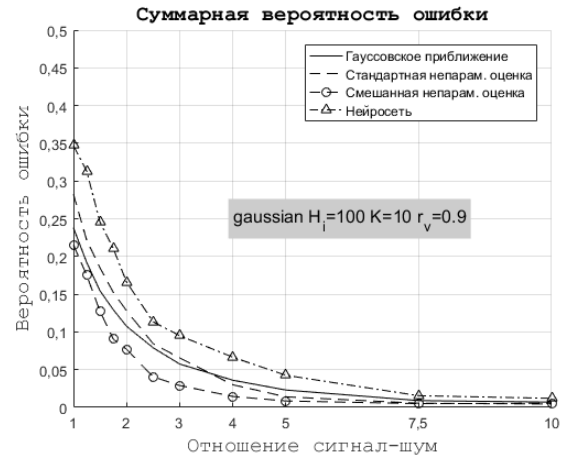
где параметры  $\alpha_0$ ,  $\alpha_v$  определялись исходя из задаваемых значений коэффициентов корреляции  $r_0$  и  $r_v$  соседних элементов случайных полей. Для всех представленных графиков использовались следующие значения параметров функции деформации:  $A_m = 0,25$ ;  $r_0 = 0,995$ .

Результаты, полученные при распознавании синтезированных изображений (рисунки 3.8, 3.9), представлены на рисунках 3.10, 3.11 в виде зависимостей вероятности суммарной ошибки распознавания  $E_r = p(\omega_1)\alpha + p(\omega_2)\beta$  от величины отношения сигнал-шум, которое задавалось как отношение пикового значения яркости изображений к  $\sigma$ :

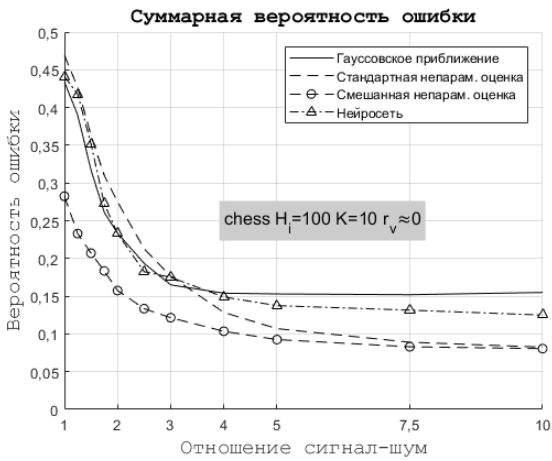
$$\rho_r = \frac{\max(\mathbf{s}_1, \mathbf{s}_2)}{\sigma}.$$



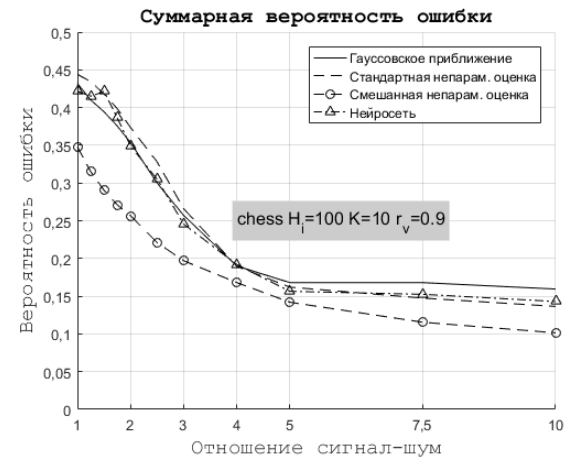
а)



б)



в)



г)

Рисунок 3.10 – Зависимости для оценок вероятности ошибки распознавания синтезированных изображений (рисунки 3.8, 3.9) от отношения сигнал-шум для  $H_i=100$  и  $K=10$

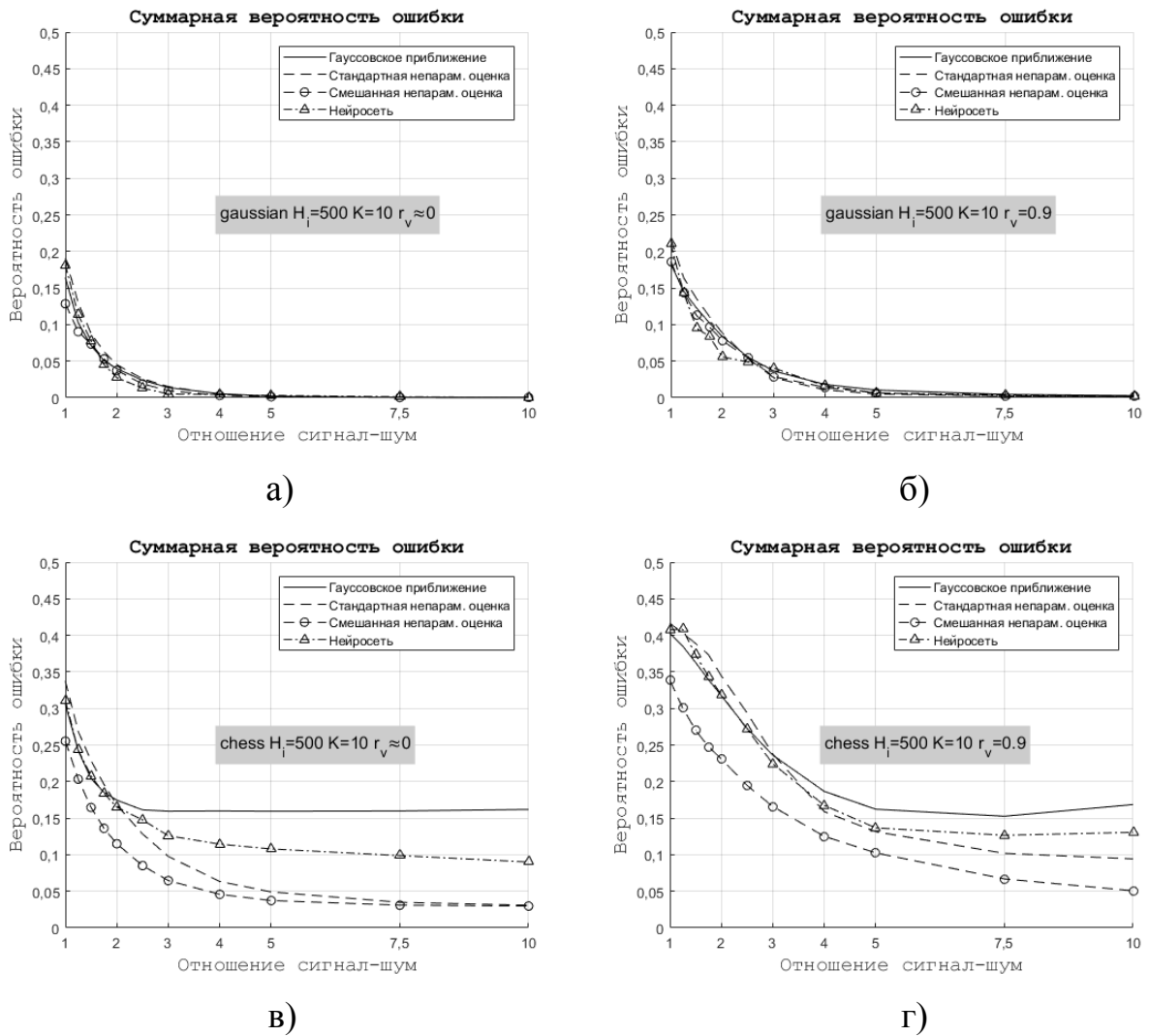


Рисунок 3.11 – Зависимости для оценок вероятности ошибки распознавания синтезированных изображений (рисунки 3.8, 3.9) от отношения сигнал-шум для  $H_i=500$  и  $K=10$

На рисунках 3.10, 3.11 представлены результаты, полученные на основе имитационного моделирования алгоритмов распознавания синтезированных изображений с усреднением по 10 реализациям их работы при  $K=10$ : на рисунке 3.10 — для  $H_i=100$ , и на рисунке 3.11 — для  $H_i=500$ . Здесь и далее обозначения на графиках, вынесенные в текстовую дорожку, соответствуют вариантам задания исходных данных: gaussian, chess — тип распознаваемых изображений;  $H_i$  — число элементов обучающей выборки;  $K$  соответствует количеству главных



компонент;  $r_v$  — коэффициент корреляции соседних элементов аддитивного шума. Представленные ниже зависимости получены для двух вариантов: воздействие аддитивного белого шума с низким уровнем корреляции  $r_v \approx 0$  (рисунки 3.10, 3.11, а, в) и воздействие значительно коррелированного аддитивного шума с коэффициентом  $r_v = 0,9$  (рисунки 3.10, 3.11, б, г).

Кроме того, на втором этапе наших исследований, как пример применимости развиваемого подхода, рассматривалась задача обработки реальных изображений, полученных при съемке движущегося неоднородного потока элементов зерновых смесей (ЭЗС) в системах фотосепарации реального времени [21] (рисунок 3.12). Нами были осуществлены эксперименты по распознаванию реальных изображений ЭЗС пшеницы (рисунок 3.13, а) и ячменя (рисунок 3.13, б). На графиках далее соответствующие результаты обозначены как *seeds*. Индивидуальные отличия формы элементов зерновых смесей каждого класса при этом трактовались как деформирующие искажения. Обработка осуществлялась в соответствии со схемой на рисунке 3.7.

Следует отметить, что при работе с изображениями необходимо учитывать изначально заданный целочисленный диапазон яркостей пикселей от 0 до 256, который переводится в процессе обработки изображений в диапазон чисел с плавающей точкой от 0 до 1. Если при проведении исследований необходимо внесение аддитивного гауссовского шума, то оно в этом случае будет не совсем корректным, так как результирующие значения яркостей пикселей могут выйти за границы исходного диапазона. В подобных случаях проводится точечное усечение уровня шума, которое проявляется только при малых соотношениях сигнал-шум.

Это оказывает следующее влияние на оценку эффективности алгоритмов распознавания — имеет место снижение вероятностей ошибок. Поэтому в рамках данной работы для получения гарантированного и неискаженного результата модификация уровня шума не проводилась.

На основе использования таких алгоритмов высокоуровневого анализа потока ЭЗС обеспечивается оценка состава исходного потока объектов и результатов их сортировки в относительно медленном масштабе времени. По результатам выполненной обработки производится динамическая настройка порогов принятия решений для одновременно выполняемой видеомодулями сепарации в режиме реального времени. Тем самым реализуется адаптивный режим управления фотосепаратором, так как, учитывая существенные ограничения по быстродействию, для принятия решений в реальном времени могут использоваться только простейшие правила порогового типа, основанные на использовании в качестве признаков таких элементарных характеристик ЭЗС как цвет, яркость и размеры.

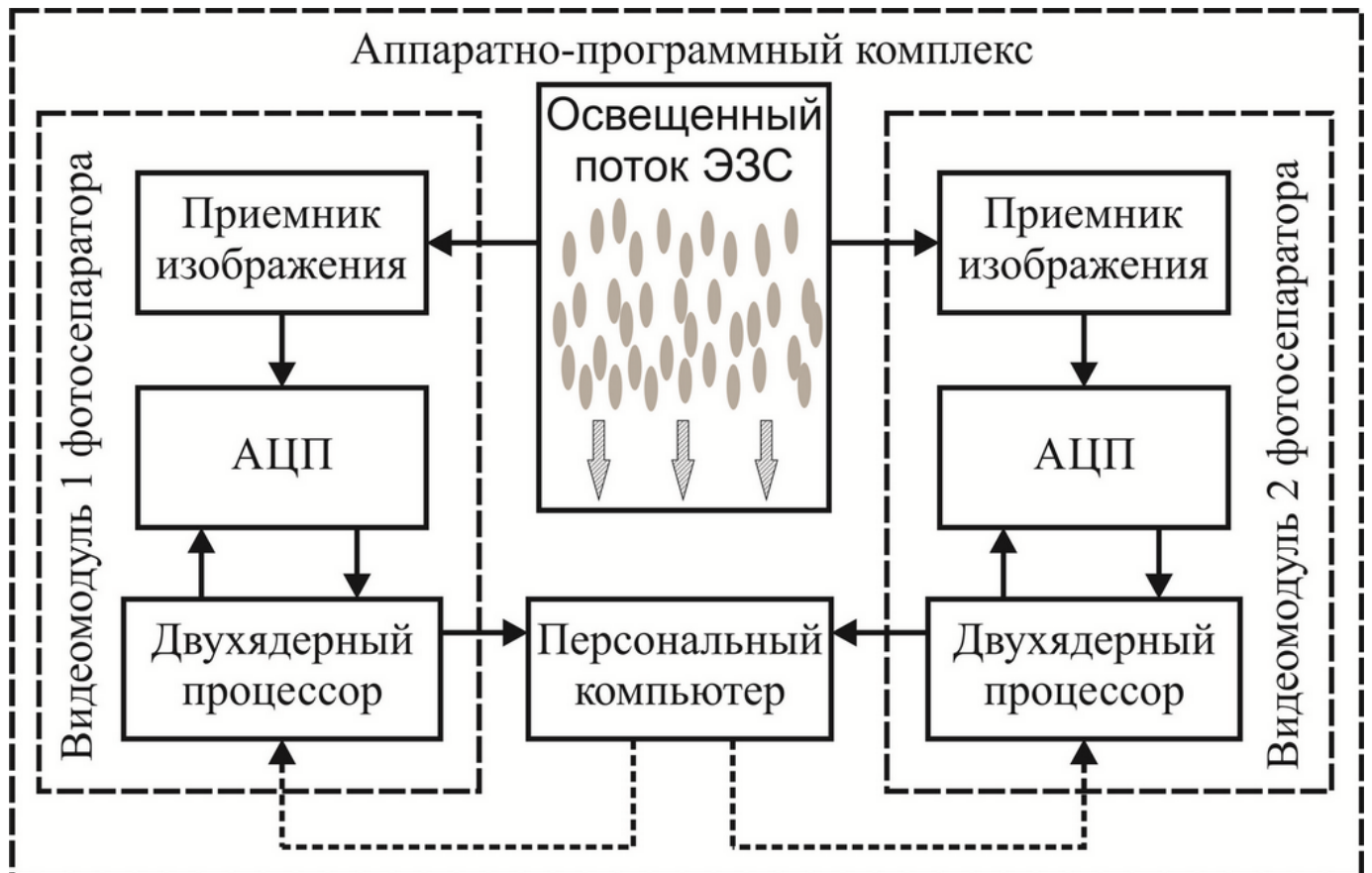


Рисунок 3.12 – Блок-схема аппаратно-программного комплекса, интегрированного в состав фотосепаратора

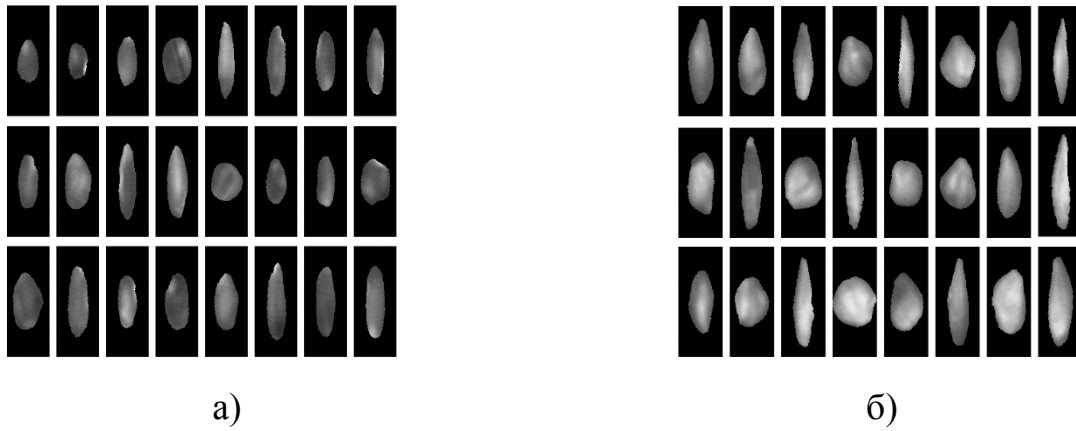


Рисунок 3.13 – Примеры подготовленных реальных изображений элементов зерновых смесей пшеницы (а) и ячменя (б), использованных при обучении алгоритмов распознавания

Зависимости для оценок вероятности ошибки распознавания этих изображений, полученные при значениях  $r_v \approx 0$  и  $r_v = 0,9$  для  $K = \{10, 20\}$  и  $H_i = 52$  с усреднением по 100 реализациям работы алгоритмов, представлены на рисунке 3.14.

**Обсуждение результатов.** При оценке вероятности ошибки распознавания изображений, имеющих вид колокольной (гауссовской) функции (рисунок 3.8), при малой величине обучающей выборки характерно преимущество алгоритмов, основанных на параметрических и непараметрических оценках, по отношению к алгоритму, основанному на использовании нейронной сети (рисунок 3.10, а, б). Это преимущество растет с увеличением коэффициента корреляции аддитивного шума, что, на наш взгляд, связано с однородностью структуры подвергаемых ДИ изображений и превалирующим в этом случае влиянием аддитивных помех. Это приводит к тому, что законы распределения значений главных компонент распознаваемых изображений достаточно хорошо описываются в гауссовском приближении, на котором и основаны использованные оценки. При малом объеме обучающей выборки выявлено значительное превосходство смешанной непараметрической оценки перед всеми другими алгоритмами, которое

нивелируется с увеличением объема обучающих данных (рисунок 3.11, а, б), что согласуется с представленными выше теоретическими выкладками.

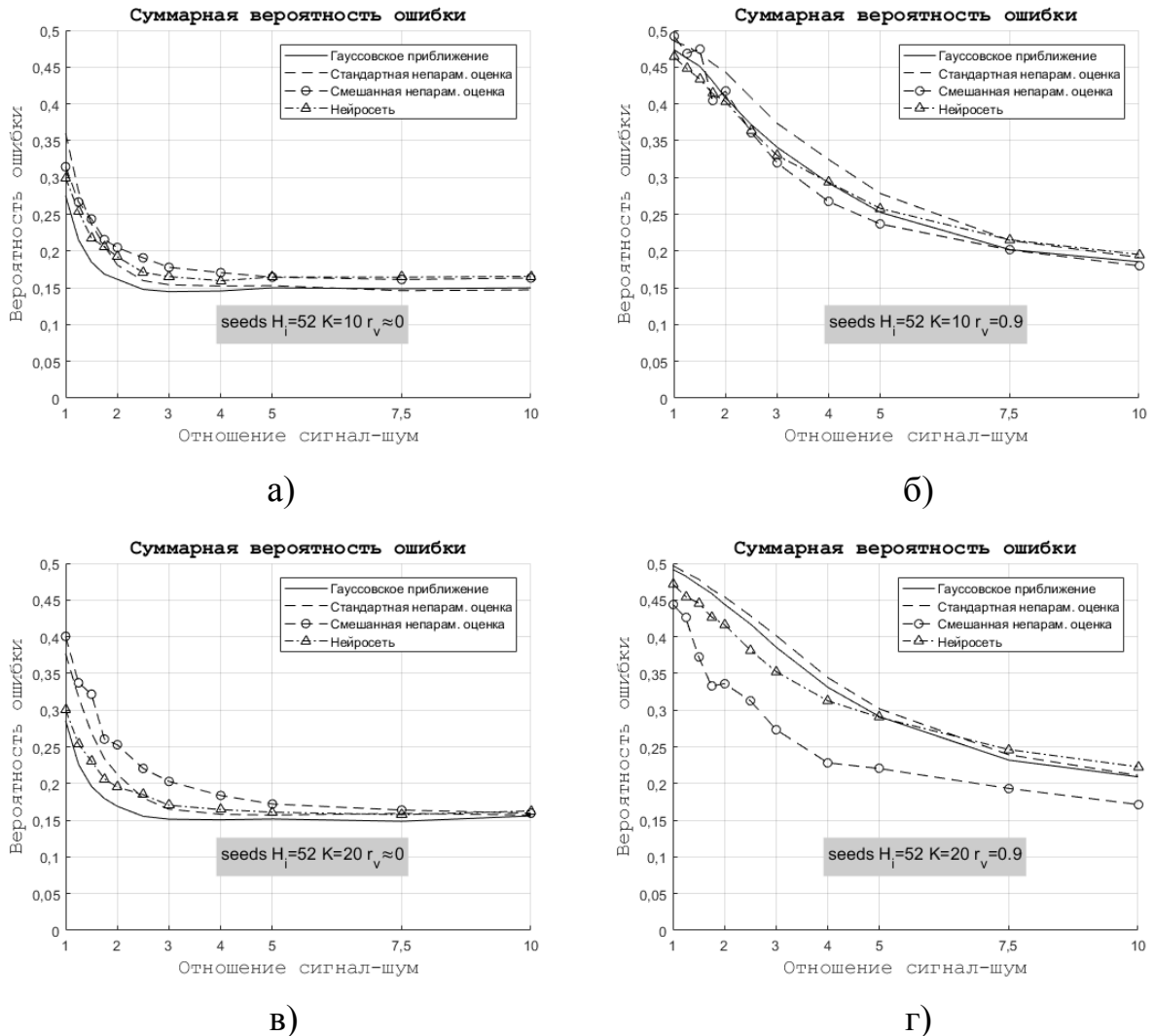


Рисунок 3.14 – Зависимости для оценок вероятности ошибки распознавания изображений элементов зерновых смесей (рисунок 3.13) от отношения сигнал-шум при  $H_i = 52$  и  $K = \{10, 20\}$

Для полученных на основе двумерного гармонического ряда исходных изображений квазипериодической структуры (рисунок 3.9), которая сильно подвержена влиянию со стороны ДИ, характерно еще более выраженное, чем в предыдущем случае, превосходство смешанной непараметрической оценки перед

всеми другими алгоритмами, которое незначительно уменьшается при увеличении объема обучающей выборки со 100 до 500 экземпляров (рисунок 3.10, в, г и рисунок 3.11, в, г). Нейросетевой алгоритм здесь в большинстве случаев показывает себя наравне с алгоритмами, основанными на параметрической и стандартной непараметрической оценках. Для последней, однако, вместе с увеличением отношения сигнал-шум характерны постепенное приближение к значениям, соответствующим смешанной непараметрической оценке, и отклонение от результатов работы остальных двух алгоритмов.

При распознавании изображений ЭЗС, подвергнутых дополнительному воздействию аддитивного шума, была получена несколько иная картина (рисунок 3.14). Для некоррелированного шума (левый столбец на рисунке 3.14), лучшие результаты показывает алгоритм, основанный на гауссовском приближении с использованием параметрических оценок. Это, по нашему мнению, связано с малым объемом доступных обучающих данных ( $N_i = 52$ ) и близостью закона распределения значений главных компонент, полученных при преобразовании относительно однородных изображений ЭЗС (рисунок 3.13), к гауссовскому закону распределения. Однако в композиции с коррелированным аддитивным шумом влияние последнего на закон распределения этих значений оказывается столь значительным, что учет параметров аддитивного шума в смешанной оценке позволяет получить наиболее точные результаты (рисунок 3.14, в, г). С увеличением числа главных компонент данный эффект становится все более выраженным, в то время как точность работы всех остальных алгоритмов снижается.

На основе полученных результатов можно сделать выводы о применимости предложенных моделей и синтезированных алгоритмов распознавания цифровых изображений в условиях деформирующих искажений и аддитивных помех. Показано, что использование алгоритмов, основанных на смешанных ядерных оценках, эквивалентно реализации процедуры искусственного размножения обучающих данных в соответствии со статистической моделью дополнительных аддитивных

искажений, оказывающих влияние на распознаваемые изображения, что представляет собой два возможных пути улучшения точности работы алгоритмов распознавания в условиях малых объемов обучающих выборок.

Искусственное размножение обучающих данных при этом становится единственной альтернативой при построении обучаемых алгоритмов распознавания, в которых невозможно напрямую учесть параметры используемых статистических моделей данных, например, алгоритмов, основанных на использовании нейронных сетей. В случае недостаточного объема обучающих данных выигрыш за счет использования смешанных оценок можно получить в условиях, когда искажения, эффекты которых требуется учесть, оказывают существенное влияние на законы распределения используемых признаков, например, в случае наличия коррелированного аддитивного шума.

### **Выводы по главе**

1. Представлена математическая постановка задачи распознавания детерминированных сигналов в присутствии шумов и деформирующих искажений и осуществлен синтез оптимальных алгоритмов распознавания цифровых сигналов на основе использования эквивалентной статистической модели представления ДИ как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. Синтезированы и исследованы оптимальный, основанный на точном вычислении функций правдоподобия, и квазиоптимальный, основанный на использовании гауссовского приближения функций правдоподобия, алгоритмы распознавания. Проведено моделирование алгоритмов при различных отношениях сигнал-шум и сравнение экспериментальной вероятности ошибки с ее теоретической оценкой при воздействии квазидетерминированных и случайных функций деформации.

Установлено, что применение оптимального алгоритма, основанного на вычислении функций правдоподобия классов как взвешенной суммы условных плотностей распределения для возможных комбинаций перестановок элементов сигналов, в большинстве случаев дает относительно небольшое преимущество по отношению к алгоритму, основанному на использовании гауссовского приближения функций правдоподобия с вычислением математических ожиданий и матриц ковариаций в соответствии с приведенными в главе соотношениями. В то же время выявлено, что в определенных ситуациях, для которых характерны малые отличия формы распознаваемых сигналов, оптимальный алгоритм может давать заметное преимущество по сравнению с алгоритмом, основанным на использовании гауссовского приближения.

2. Представлена эквивалентная статистическая модель деформирующих искажений в задачах распознавания изображений в присутствии шумов и осуществлен синтез алгоритмов их распознавания с использованием стандартных и смешанных непараметрических ядерных оценок функций правдоподобия классов. Доказаны свойства несмещенности и состоятельности смешанных ядерных оценок многомерных плотностей распределения вероятностей. Предложена обобщенная структура алгоритма обработки информации в интересах синтеза решающих правил при распознавании сигналов и изображений в случае, когда имеются ограничения вычислительной сложности или отсутствуют необходимые исходные данные. Показано, что использование алгоритмов, основанных на смешанных ядерных оценках, эквивалентно реализации процедуры искусственного размножения обучающих данных в соответствии со статистической моделью дополнительных аддитивных искажений, оказывающих влияние на распознаваемые изображения. Проведены моделирование и сравнение работы алгоритмов, основанных на использовании стандартных и смешанных непараметрических ядерных оценок, с алгоритмом, основанным на параметрических оценках и гауссовском приближении,

и алгоритмом, основанным на использовании нейронной сети прямого распространения.

3. Теоретически и экспериментально показано, что процедура искусственного размножения обучающих данных и использование смешанных ядерных оценок представляют собой два возможных пути улучшения точности работы алгоритмов распознавания в условиях малых объемов обучающих выборок. Искусственное размножение обучающих данных при этом становится единственной альтернативой при построении обучаемых алгоритмов распознавания, в которых невозможно напрямую учесть параметры используемых статистических моделей данных, например, алгоритмов, основанных на использовании нейронных сетей. В случае недостаточного объема обучающих данных существенный выигрыш за счет использования смешанных оценок можно получить в условиях, когда искажения, которые требуется учесть, оказывают значительное влияние на законы распределения используемых признаков, например, в случае наличия коррелированного аддитивного шума.



#### **4 Модели и алгоритмы искусственного размножения данных путем внесения деформирующих искажений для обучения алгоритмов распознавания лиц методом Виолы-Джонса**

Построение алгоритмов распознавания реальных изображений на основе методов машинного обучения при наличии необходимых обучающих данных обычно осуществляется по стандартным схемам, диктуемым парадигмой используемого подхода. Однако зачастую обучающих данных недостаточно, что может быть вызвано не только объективными ограничениями при их получении, но и большой трудоемкостью разметки обучающих выборок. В этой ситуации при построении алгоритмов распознавания реальных изображений одним из перспективных представляется подход, основанный на аугментации — искусственном размножении обучающего множества данных по каждому из классов.

Решение задачи таким способом имеет смысл, когда воздействие ДИ можно рассматривать как реализации некоторой случайной функции, которые описывают различия между элементами размножаемой обучающей выборки и которые можно сгенерировать при известной статистической модели деформирующих искажений. Можно ожидать, что подобный способ построения описаний классов хорошо встраивается в задачу машинного обучения алгоритмов на основе обучающих данных, полученных из набора реальных изображений.

Анализ возможностей применения подобного подхода был выполнен нами на примере алгоритма распознавания лиц по методу Виолы-Джонса. Данный метод был выбран в качестве инструмента исследования как один из самых надежных и популярных на данный момент детекторов фронтальных лиц [82].

## **4.1 Обоснование и исследование алгоритмов распознавания изображений на основе метода Виолы-Джонса с использованием технологии вычислений на графических процессорах CUDA**

Метод Виолы-Джонса подробно описан в работах [83-87] и основан на применении целого множества взаимосвязанных подходов и технологий. В центре данного метода лежит идея построения запускающихся друг за другом каскадов, реализующих точные или, как еще говорят, сильные классификаторы путем комбинирования их на основе так называемого «бустинга» из наборов других менее точных или слабых классификаторов — признаков Хаара. Применение обученного детектора производится путем перемещения и масштабирования сканирующего окна.

### **4.1.1 Реализация последовательной версии алгоритма на основе метода Виолы-Джонса**

Для создания последовательной версии метода была использована среда Matlab. Этот выбор обусловлен наличием в ней уже множества реализованных математических операторов и функций, зачастую отсутствующих в других языках программирования. Также среда Matlab содержит весь необходимый инструментарий, применимый для проведения анализа работы реализации метода.

Метод Виолы-Джонса можно разделить на два независимых алгоритма: обучение классификатора и анализ изображений готовым классификатором. Сам алгоритм обучения состоит из нескольких следующих друг за другом фаз (рисунок 4.1).

**Подготовка признаков.** Представляет собой фазу создания набора признаков Хаара. Производится перебор всех типов признаков с учетом их возможных размеров и положений внутри сканирующего окна размером 24x24.

**Подготовка и анализ положительных экземпляров выборки.** Представляет собой фазу подготовки обучающей выборки изображений лиц и получение первичных результатов ее анализа полным набором признаков. Оставшиеся три фазы выполняются в цикле до тех пор, пока не будут получены требуемые показатели достоверности работы алгоритма распознавания.

**Сбор и анализ отрицательных экземпляров выборки.** На этой фазе [83, 84, 87] производится сбор и первичный анализ полным набором признаков Хаара необходимого числа отрицательных экземпляров выборки. В качестве отрицательных экземпляров выборки используется содержимое сканирующих окон, вызвавших ложные срабатывания последнего слоя имеющегося на данный момент каскада классификаторов при обходе им в рабочем режиме заведомо не содержащих лица изображений.

**Сортировка результатов анализа тренировочной выборки.** На данном шаге алгоритма обучения для каждого отдельного признака Хаара производится объединение и сортировка по возрастанию первичных результатов обработки вместе положительных (изображений лиц) и отрицательных экземпляров обучающей выборки. Сортировка производится для ускорения процедуры определения порогового значения в процессе обучения слабого классификатора во время бустинга на следующей фазе работы алгоритма.

**Построение следующего слоя каскада.** Здесь осуществляется построение слоя каскада путем формирования его в ходе процедуры бустинга в виде линейной комбинации лучших из слабых классификаторов, обученных на текущем наборе положительных и отрицательных экземпляров обучающей выборки.

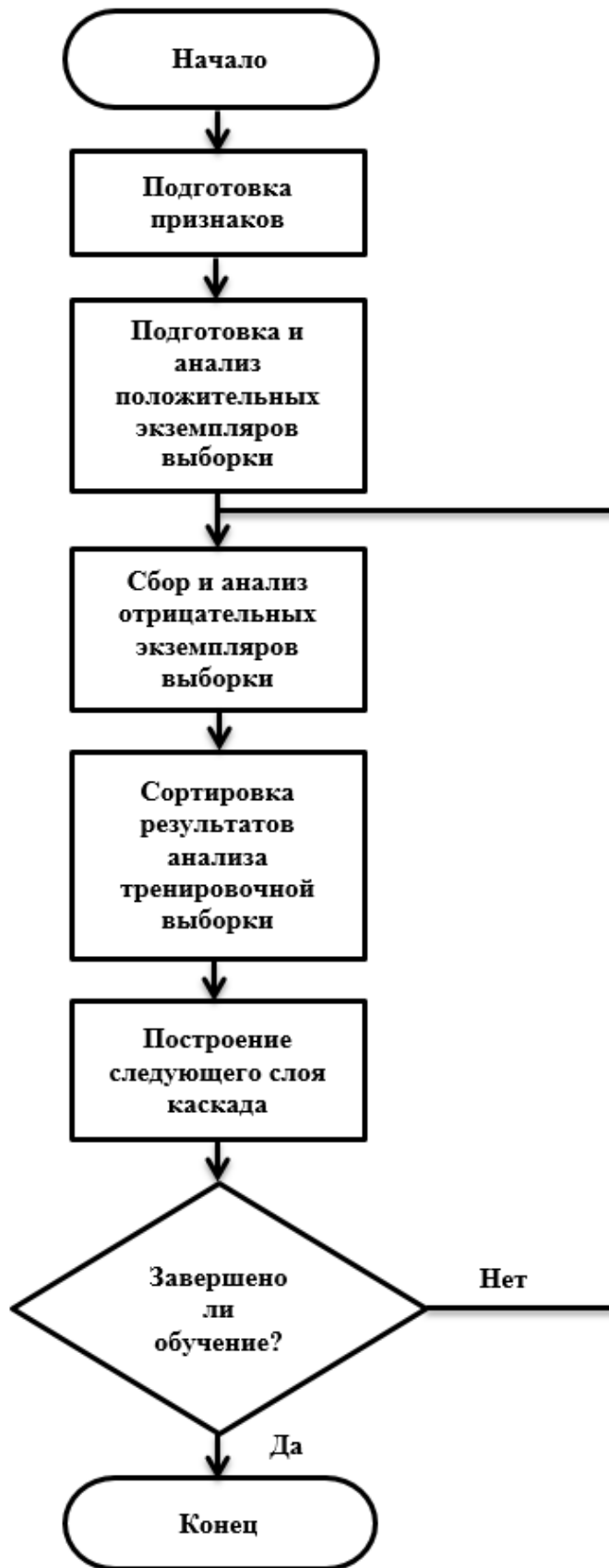


Рисунок 4.1 – Схема алгоритма обучения по методу Виолы-Джонса

#### **4.1.2 Реализация параллельных версий отдельных этапов работы алгоритма на основе метода Виолы-Джонса и экспериментальные исследования по сравнению скорости их работы с последовательными версиями**

Для создания параллельных версий отдельных этапов работы алгоритма в качестве механизма осуществления параллельных вычислений на графическом процессоре была выбрана архитектура CUDA, как единственно доступная из среды Matlab. При проведении всех экспериментов по оценке производительности алгоритмов в качестве испытательного стенда использовался компьютер со следующими характеристиками: процессор Intel Core i5-3210M 2.5 ГГц; 6 ГБ DDR3 оперативной памяти; видеокарта — NVIDIA GeForce GT 640M с 2 ГБ памяти и 384 ядрами CUDA (имеет возможность оперирования с числами с плавающей запятой двойной точности — double); ОС — 64-х битная версия Ubuntu 14.04.

На этапе обучения были распараллелены алгоритм сбора первичных результатов обработки признаками экземпляров обучающей выборки, как положительных, так и отрицательных, а также алгоритм обучения слабых классификаторов. Результаты опубликованы автором в работах [69, 70].

**Создание параллельной версии алгоритма сбора первичных результатов обработки.** Алгоритм расчета первичных результатов обработки признаками экземпляров обучающей выборки хорошо поддается распараллеливанию ввиду того, что анализ отдельным признаком отдельного экземпляра выборки производится независимо от остальных.

Из-за большого объема данных и ограниченных ресурсов оперативной памяти данный алгоритм решено было разбить на несколько этапов с постепенным сохранением результатов обработки на диск с целью последующего их использования во время обучения (рисунок 4.2). При этом в каждый файл для определенного числа признаков из общего набора (это число определяется доступным объемом оперативной памяти) записываются результаты первичного анализа ими

всех имеющихся экземпляров обучающей выборки. Замеры времени работы показали, что параллельная версия алгоритма первичного анализа обучающей выборки полным набором признаков Хаара по скорости работы превосходит последовательную его версию более чем в сто раз (таблица 4.1).

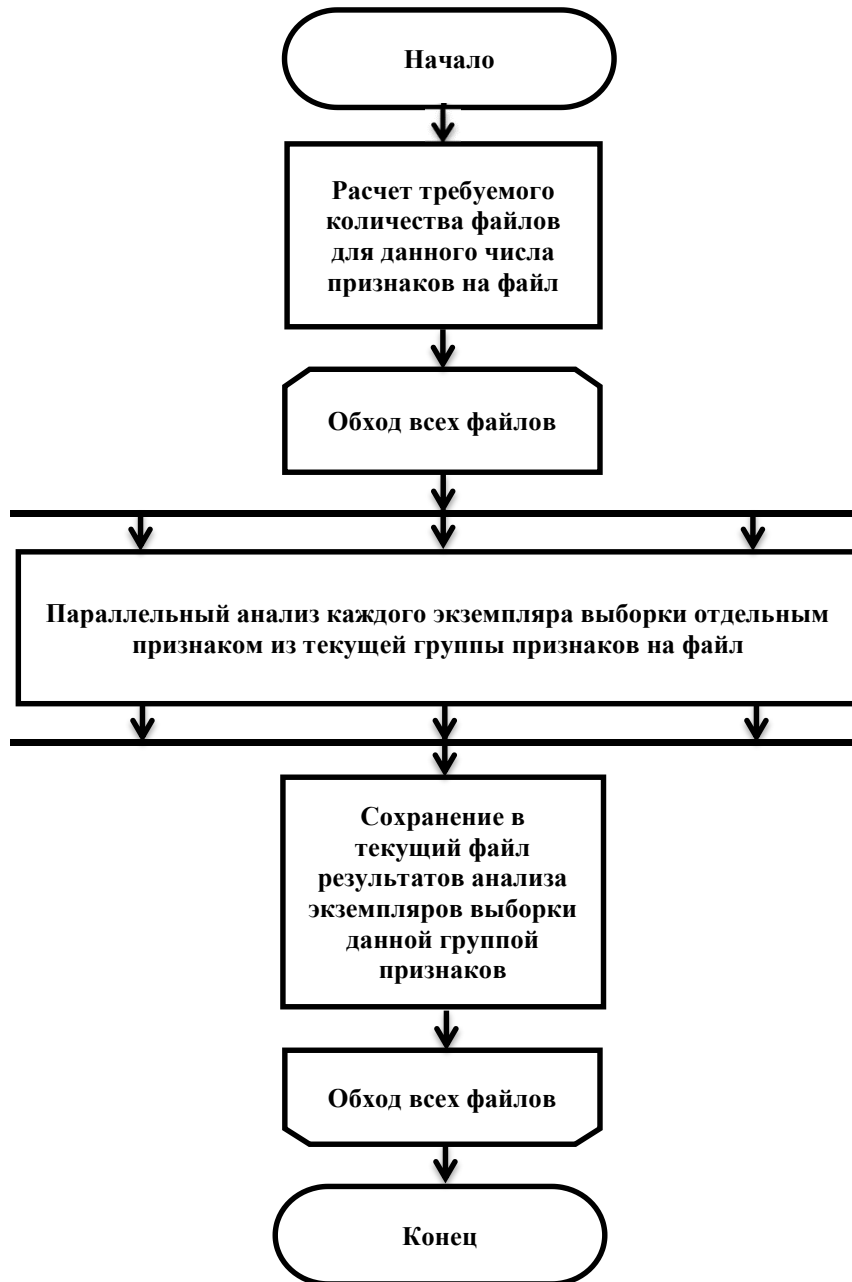


Рисунок 4.2 – Схема параллельной версии алгоритма сбора первичных результатов обработки экземпляров обучающей выборки полным набором классификаторов

Таблица 4.1 – Показатели времени работы последовательной и параллельной версий алгоритма сбора первичных результатов анализа экземпляров обучающей выборки полным набором признаков Хаара

Число обработанных экземпляров тренировочной выборки	Последовательная версия	Параллельная версия
2000	3 ч 23 мин	1 мин 33 с
4000	5 ч 13 мин	3 мин 30 с
6000	8 ч 38 мин	4 мин 26 с

**Создание параллельной версии алгоритма обучения слабого классификатора.** На данном этапе ускорение было достигнуто за счет параллельного обучения слабых классификаторов. Общая схема работы получившегося алгоритма приведена на рисунке 4.3. В этом случае из-за ограничений памяти количество одновременно работающих потоков невелико, и с увеличением объема обучающей выборки их число падает, что оказывает соответствующее влияние на производительность алгоритма (таблица 4.2).

Таблица 4.2 – Показатели времени выполнения последовательной и параллельных версий алгоритма обучения слабых классификаторов и выбора лучшего из них в ходе процедуры бустинга

Число экземпляров обучающей выборки	Последовательная версия	Параллельная версия
2000	0 мин 52 с	0 мин 24 с
4000	2 мин 04 с	1 мин 09 с
6000	2 мин 47 с	1 мин 53 с

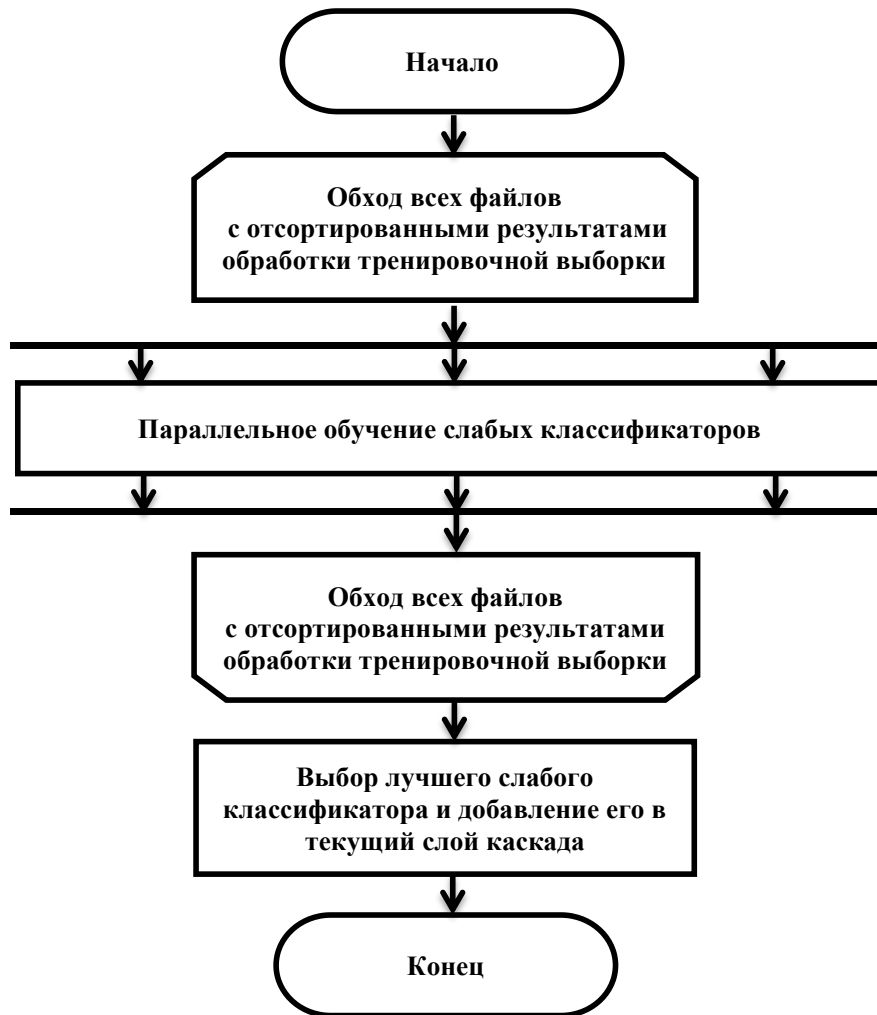


Рисунок 4.3 – Схема параллельной версии алгоритма обучения слабых классификаторов перед добавлением лучшего из них в текущий слой каскада

При этом следует отметить, что используемый механизм сохранения отсортированных первичных результатов обработки обучающей выборки малым числом признаков с последующим их считыванием всякий раз во время обучения каждого следующего слабого классификатора дает выигрыш только в случае, если время, затрачиваемое на чтение с диска, не превышает время расчета и сортировки первичных результатов обработки обучающей выборки данным числом признаков. В противном случае этот расчет выгоднее производить каждый раз «на лету» во время самого обучения.



**Применение ранее обученного классификатора для распознавания лиц.** На этапе использования ранее обученного классификатора производится перемещение сканирующего окна по всему изображению с постепенным увеличением масштаба анализа и для каждого из положений сканирующего окна рассчитывается возвращаемое полученным при обучении каскадом значение в виде результата классификации. Также производится объединение нескольких близких положительных результатов детектирования в один.

При создании параллельной версии алгоритма анализа изображения был реализован параллелизм сканирующих окон в ходе выполнения вычислительного процесса применения каскада. Из-за ограничений по объему памяти масштабирование сканирующих окон производится в последовательном режиме. Замеры производительности показали в среднем более чем тысячекратное превосходство параллельной версии алгоритма (таблица 4.3).

Таблица 4.3 – Показатели замеров времени работы последовательной и параллельной версий алгоритма анализа изображений готовым классификатором

Размер изображения в пикселях	Последовательная версия	Параллельная версия
160x120	0 ч 14 мин	1.16 с
320x240	1 ч 54 мин	4.15 с
640x480	9 ч 32 мин	21.87 с

## 4.2 Модели и алгоритмы внесения деформирующих искажений с целью аугментации обучающей выборки при распознавании лиц по методу Виолы-Джонса

При проведении исследования были рассмотрены три модели внесения ДИ с целью искусственного размножения обучающих данных. Результаты исследования опубликованы автором в работе [72].

### 4.2.1 Алгоритм внесения искусственной деформации на основе гармонических функций

Первая модель внесения ДИ основана на применении в качестве функции деформации  $\mathbf{r}(\mathbf{x}) = \mathbf{r}(x, y) = \{r_i(x, y), i = \overline{1, 2}\}$  гармонических функций (ГФ) вида

$$r_i(x, y) = A_i \sin(\omega_{xi}x + \omega_{yi}y + \psi_i),$$

где  $A_i$  — амплитуда смещения при внесении ДИ;  $\omega_{xi}$  и  $\omega_{yi}$  — частоты функции деформации по каждой из осей координат изображения;  $\psi_i$  — фаза.

Значение амплитуды  $A_i$  ограничивает максимальную величину смещения координат пикселей при внесении ДИ и определяется тем, насколько можно независимо друг от друга сместить черты лица относительно их исходных позиций так, чтобы не разрушить его общую структуру.

Частоты  $\omega_{xi}$  и  $\omega_{yi}$  определяют размеры деталей изображения, которые будут деформироваться. При малых значениях осуществляется смещение и плавная деформация изображения объекта в целом с сохранением его структурных свойств, а при их увеличении — деформируются все более мелкие его части. Исходя из этих соображений используемые в экспериментах значения  $\omega_{xi}$  и  $\omega_{yi}$  подбирались так,

чтобы период функции деформации  $\tau_i(x, y)$  примерно соответствовал размерам основных черт лица на изображении.

Фаза функции деформации  $\psi_i$  определяет ее смещение, и, для обеспечения возможности получения нескольких разных деформированных изображений на основе одного экземпляра исходной обучающей выборки, ее значение генерируется случайным образом по равномерному закону на всей области определения функции деформации.

В итоге реализуемый в соответствии с данной моделью алгоритм состоял в том, что для каждого из изображений исходной обучающей выборки заданное число раз с учетом описанных выше соображений генерировались параметры функции деформации, и на их основе производилось внесение ДИ с целью получить новые изображения.

На рисунке 4.4 показаны примеры результатов внесения ДИ такого типа в изображения лиц, которые могут использоваться в качестве обучающей выборки в методе Виолы-Джонса.



Рисунок 4.4 – Примеры внесения деформирующих искажений в изображения лиц с применением гармонической функции деформации (первый столбец слева является исходным)

## 4.2.2 Алгоритм внесения искусственной деформации на основе анализа смещений контрольных точек

Вторая модель внесения ДИ основана на часто применяемом при реализации морфинга [14, 88] подходе, заключающемся в ручной расстановке соответствий на двух совмещаемых изображениях в виде контрольных точек или контуров. Далее, на основе известных для отмеченных элементов попарных смещений или значений функции деформации в данных точках при помощи интерполяции или некоторого другого метода находят все остальные значения функции деформации. Одним из используемых при таком подходе способов интерполяции является применение радиально-базисных функций (РБФ) [89].

Реализуемый в соответствии с данной моделью алгоритм состоит в следующем. Для каждого из изображений подлежащей размножению обучающей выборки сначала производится расстановка контрольных точек  $(x_k, y_k)$ ,  $k = \overline{1, m}$ , где число контрольных точек  $m = 5$ : по одной на глаз, кончик носа и уголки рта (рисунок 4.5).



Рисунок 4.5 – Примеры расстановки контрольных точек на изображениях лиц

Далее рассчитываются средние значения  $(x_k, y_k)$  и их матрица ковариации, которые используются при генерации случайных положений контрольных точек для новых искусственно получаемых данных. После этого внесение ДИ в одно из изображений исходной обучающей выборки производится следующим образом.

Генерируется случайный набор координат контрольных точек  $(x_k, y_k)$  в соответствии с полученными ранее статистическими параметрами их распределения

(при генерации вместо средних значений также могут подаваться координаты самих контрольных точек размножаемого изображения). Для каждой из этих точек на основе (2.1) и (2.2) рассчитываются значения функции деформации  $r_i(x_k, y_k)$ ,  $i=\overline{1,2}$  в данной точке так, чтобы

$$f(x_k + r_1(x_k, y_k), y_k + r_2(x_k, y_k)) = f(\mathbf{u}(x_k, y_k)) = g(x_k, y_k),$$

где  $f(x, y)$  — деформируемое изображение;  $g(x, y)$  — результат деформации, соответствующий набору случайно сгенерированных положений контрольных точек. При этом  $(x_k + r_1(x_k, y_k), y_k + r_2(x_k, y_k))$  — координаты контрольных точек деформируемого изображения обучающей выборки.

В данной модели конечные функции деформации  $r_i(x_k, y_k)$ ,  $i=\overline{1,2}$  представляются в виде сумм RBF-функций:

$$r_i(x, y) = \sum_{k=1}^m a_{ki} \exp\left(-\frac{(x-x_k)^2 + (y-y_k)^2}{2\sigma^2}\right), \quad (4.1)$$

где  $\sigma$  — ширина окна RBF-функций;  $a_{ki}$  — неизвестные коэффициенты отображения. Элементом суммы соответствуют RBF-функции с центрами по координатам  $(x_k, y_k)$ .

Если для  $X$  и  $Y$  в координатах контрольных точек  $(x_k, y_k)$ ,  $k=\overline{1, m}$  вместо  $r_i(x, y)$  поочередно подставлять в (4.1) известные значения  $r_i(x_k, y_k)$ , то для каждой из функций деформации, соответствующей смещению по координатам  $X$  или  $Y$ , получится по системе из  $m$  линейных уравнений (число неизвестных также равно  $m$ ):

$$\sum_{l=1}^m \exp\left(-\frac{(x_k - x_l)^2 + (y_k - y_l)^2}{2\sigma^2}\right) a_{il} = r_i(x_k, y_k), \quad k = \overline{1, m}.$$

В матричном виде это эквивалентно следующему уравнению:

$$\mathbf{E} \mathbf{A}_i = \mathbf{R}_i,$$

где матрица  $\mathbf{E}$  состоит из элементов

$$\sum_{l=1}^m \exp\left(-\frac{(x_k - x_l)^2 + (y_k - y_l)^2}{2\sigma^2}\right)$$

со столбцами, пронумерованными значениями индексов  $l = \overline{1, m}$  и строками, пронумерованными значениями  $k = \overline{1, m}$ , вектор  $\mathbf{A}_i = (a_{i1}, \dots, a_{im})^T$  содержит неизвестные коэффициенты РБФ-функций, и вектор  $\mathbf{R}_i = (r_i(x_1, y_1), \dots, r_i(x_m, y_m))^T$  — известные значения функции деформации соответствующей компоненты изображения в контрольных точках. При этом матрица  $\mathbf{E}$  одинакова для всех  $r_i(x_k, y_k)$ ,  $i = \overline{1, 2}$  и не зависит от компоненты изображения.

В простейших вариантах значение  $\sigma$  определяется экспериментально путем перебора в заданном диапазоне значений или для каждой контрольной точки и соответствующей ей РБФ-функции выставляется равным расстоянию от нее до ближайшей к ней другой контрольной точки.

Получаемые таким образом системы уравнений не всегда могут иметь стандартное решение  $\mathbf{A}_i = \mathbf{E}^{-1}\mathbf{R}_i$  из-за плохой обусловленности матрицы уравнения или в ситуации, когда число контрольных точек меньше числа используемых для интерполяции функций. В этих случаях применяют методы регуляризации. В большинстве случаев, используют расчет псевдообратной матрицы (тогда  $\mathbf{A}_i = \mathbf{E}^+\mathbf{R}_i$ , где  $\mathbf{E}^+$  — псевдообратная матрица для  $\mathbf{E}$ ) или получают решение на основе регуляризации по А.Н. Тихонову [74]

$$\mathbf{A}_i = (\lambda \mathbf{I} + \mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T (\mathbf{R}_i - \mathbf{E} \mathbf{A}_{i0}),$$

где  $\mathbf{I}$  — единичная матрица размером  $m \times m$ ;  $\lambda$  — параметр регуляризации;  $\mathbf{A}_{i0}$  — априорная оценка решения.

На рисунке 4.6 показаны примеры результатов внесения ДИ на основе описанного подхода.



Рисунок 4.6 – Примеры внесения деформирующих искажений в изображения лиц на основе расстановки контрольных точек и интерполяции функции деформации при помощи РБФ (первый столбец слева является исходным)

#### **4.2.3 Алгоритм внесения искусственной деформации на основе анализа оптического потока с использованием фильтра энтропии**

Третья модель основана на идеях определения оптического потока (ОП) в процессе изменения положений и ориентации объектов сцены в разные моменты времени или при разных условиях съемки [37-44]. Вычисление оптического потока состоит в оценке смещений содержимого сцены по координатам  $x$  и  $y$ , которые и предлагается напрямую использовать в качестве значений функций деформации  $r_i(x, y)$ ,  $i=1,2$ .

Существует множество механизмов определения оптического потока [38-40]. Каждый из них ограничен своей областью применимости из-за заложенных в него допущений, на основе которых и производится его расчет. Для данной модели из-за своих особенностей был выбран метод Фарнебака [40], краткое описание которого представлено в разделе 1.3.

Из-за специфики предметной области — обработка изображений лиц, разницу между которыми нельзя увязать к смещению положений интенсивностей их

пикселей, — перед процедурой поиска оптического потока требуется провести дополнительное преобразование. Им является операция наложения фильтра энтропии [73], который для каждой координаты изображения и ее окрестности из  $n$  пикселей вычисляется следующим образом. Рассчитываются вероятности  $p_i, i = \overline{1, n}$  появления каждого значения интенсивности пикселя в данной области, и на их основе — энтропия ее центрального пикселя:

$$e(x, y) = -\sum_{i=1}^n p_i \log p_i,$$

где  $e(x, y)$  — результат наложения фильтра, который и используется для расчета оптического потока.

Таким образом, изображения лиц заменяются образами, отражающими их основную структуру. В них мелкие черты лица соответствуют высоким значениям энтропии, а фон и кожа лба и щек — более низким (рисунок 4.7).



Рисунок 4.7 – Примеры наложения фильтра энтропии на изображения лиц

В итоге размножение изображений обучающей выборки осуществлялось следующим образом. На каждое из них производилось наложение фильтра энтропии и вычисление  $e(x, y)$ . После чего рассчитывались выборочные параметры энтропии: среднее и матрица ковариации.

Далее, при размножении каждого изображения исходной обучающей выборки производилась генерация случайных значений энтропии  $e_g(x, y)$  и расчет функции



деформации  $r_i(x, y)$ ,  $i=\overline{1,2}$  путем вычисления оптического потока между сгенерированным значением энтропии и энтропией размножаемого изображения. Полученный результат  $r_i(x, y)$ ,  $i=\overline{1,2}$  и использовался при внесении ДИ в качестве функции деформации. На рисунке 4.8 показаны примеры работы алгоритма, основанного на данной модели.



Рисунок 4.8 – Примеры внесения деформирующих искажений в изображения лиц на основе наложения фильтра энтропии и расчета оптического потока (первый столбец слева является исходным)

#### **4.2.4 Экспериментальные исследования влияния искусственного размножения элементов обучающей выборки на качество распознавания лиц на основе метода Виолы-Джонса**

Для оценки влияния аугментации или искусственного размножения элементов обучающей выборки на качество детектирования были проведены эксперименты по обучению детекторов лиц по методу Виолы-Джонса с последующими замерами точности их работы [71, 72]. Обучение было осуществлено с использованием пяти вариантов формирования обучающих выборок:

- оригинальная, содержащая 4916 изображений лиц, собранных и подготовленных вручную авторами оригинального метода Виолы-Джонса [83, 84];

- усеченная, содержащая 492 изображений лиц из числа изображений оригинальной выборки;

- три искусственно размноженные выборки, содержащие по 4916 изображений лиц, из которых 492 изображения принадлежали усеченному набору изображений, а остальные (по 9 на каждое из исходных) сгенерированы искусственно путем внесения ДИ на основе каждой из описанных выше моделей.

Оценка качества детектирования производилось следующим образом. Для каждого детектора модифицировался порог его реагирования, и в соответствии с полученными откликами были построены зависимости относительной частоты правильно найденных лиц от числа ложных срабатываний (рисунок 4.9). В качестве изображений для тестирования был использован специальный набор изображений из выборки CMU+ MIT [90], часто применяемый для сравнения работы детекторов фронтальных лиц и содержащий 42 изображения со 170 лицами.

В таблице 4.4 представлены типичные результаты, полученные для каждого из детекторов при схожем числе ложных срабатываний. Анализ данных таблицы 4.4 и полученных зависимостей (рисунок 4.9) показывает, что при схожем числе ложных срабатываний детектор, обученный с использованием усеченной выборки, стабильно выдает результаты хуже оригинального примерно на 20%.

В то же время детекторы, основанные на трех рассмотренных выше моделях аугментации обучающих данных путем внесения в них ДИ, дают сравнимый с оригиналом результат с отклонением от него не более 5%. При этом наилучшие результаты (практически такие же, как и полученные на основе оригинальной версии) показал алгоритм, основанный на использовании расчета оптического потока с применением фильтра энтропии.

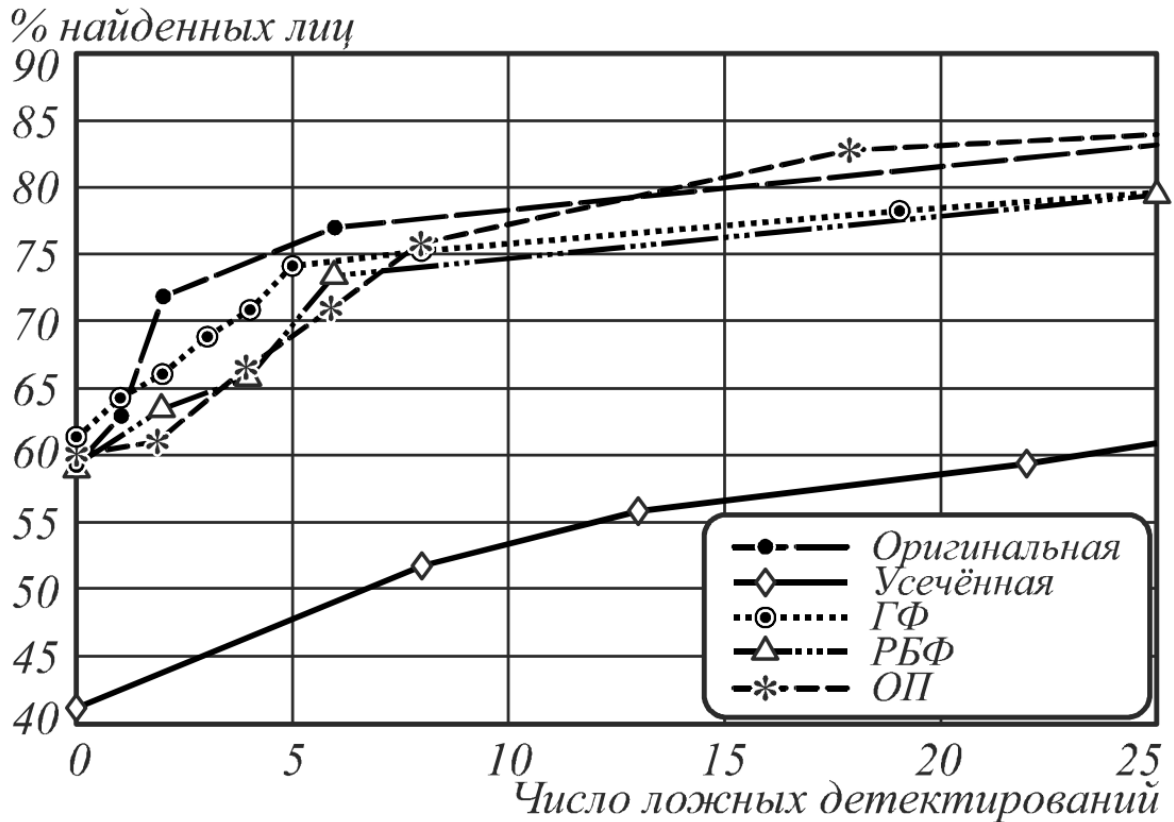


Рисунок 4.9 – Зависимости для относительной частоты правильно найденных лиц от числа ложных срабатываний, полученные при обучении по оригинальной и усеченной выборкам, а также с использованием трех рассмотренных моделей аугментации обучающих данных

Таблица 4.4 – Относительные частоты правильно найденных лиц при схожем числе ложных срабатываний, полученные при обучении по оригинальной и усеченной выборкам, а также с использованием трех рассмотренных моделей аугментации обучающих данных

	Оригинальная	Усеченная	ГФ	РБФ	ОП
Процент найденных лиц	83,53	59,41	78,23	79,41	82,94
Число ложных срабатываний	26	22	19	25	18

Таким образом, показано значительное увеличение точности работы детекторов лиц при сравнении их с классификатором, основанным на усеченной обучающей выборке (которая и была использована для размножения). Полученные результаты позволяют говорить о применимости данного подхода к задачам построения детекторов обнаружения объектов, имеющих ярко выраженные структурные характеристики, для сокращения затрат на поиск и подготовку обучающих данных, а также в условиях невозможности сбора достаточного для начала процесса обучения количества элементов обучающей выборки.

### **Выводы по главе**

1. Описана реализация метода Виолы-Джонса с использованием технологии параллельных вычислений на графических процессорах CUDA. Осуществлены эксперименты по сравнению производительности параллельной и последовательной версий различных этапов выполнения алгоритма. Результаты проведенных экспериментов выявили заметные преимущества в скорости параллельной версии реализации алгоритма Виолы-Джонса над последовательной и позволили рекомендовать применение параллельных вычислений на графических процессорах как перспективного способа увеличения производительности метода Виолы-Джонса с учетом относительной простоты и низкой стоимости их использования.

2. В интересах решения задач машинного обучения, связанных с обработкой изображений, на примере задачи фронтального распознавания лиц методом Виолы-Джонса рассмотрены три модели внесения деформирующих искажений для аугментации обучающей выборки: на основе применения при деформации гармонических функций, на основе смещения контрольных точек и на основе поиска оптического потока по методу Фарнебака при наложении фильтра энтропии. Для исходных и искусственно размноженных в соответствии с описанными моделями обучающих выборок проведены эксперименты по обучению алгоритмов

распознавания изображений лиц методом Виолы-Джонса и даны результаты сравнения точности работы полученных детекторов.

3. Анализ полученных результатов показывает, что при схожем числе ложных срабатываний детектор, обученный с использованием усеченной выборки, стабильно выдает результаты хуже оригинального примерно на 20%. В то же время алгоритмы, основанные на использовании трех рассмотренных моделей аугментации обучающих данных путем внесения в них деформирующих искажений, дают сравнимый с оригиналом результат с отклонением от него не более 5%. При этом наилучшие результаты (практически такие же, как и полученные на основе оригинальной версии) показал алгоритм, основанный на использовании расчета оптического потока с применением фильтра энтропии.

Представленные результаты позволяют говорить о применимости данного подхода к задачам построения детекторов обнаружения объектов, имеющих ярко выраженные структурные характеристики, поскольку это обеспечивает сокращение затрат на поиск и подготовку обучающих данных, а также в условиях невозможности сбора достаточного для начала процесса обучения количества элементов обучающей выборки.

## Заключение

В ходе выполнения диссертационной работы были поставлены и решены следующие научные задачи:

- проведен анализ известных подходов к использованию моделей деформирующих искажений в задачах обработки сигналов и изображений в системах управления и принятия решений;

- предложены, обоснованы и исследованы модели и алгоритмы влияния деформирующих искажений в рамках задач обработки сигналов и изображений;

- осуществлен статистический синтез и анализ алгоритмов распознавания сигналов и изображений в условиях деформирующих искажений и аддитивного шума;

- синтезированы и исследованы модели и алгоритмы аугментации (искусственного размножения) данных путем внесения в них деформирующих искажений в задачах машинного обучения на примере алгоритмов поиска и обнаружения лиц методом Виолы-Джонса.

В ходе исследований получены следующие новые результаты и выводы.

1. По итогам анализа известных подходов к использованию моделей деформирующих искажений в задачах обработки сигналов и изображений в системах управления и принятия решений представлена и описана общая схема проведения исследований в интересах построения алгоритмов обработки информации для синтеза различных типов алгоритмов принятия решений с использованием моделей деформирующих искажений.

2. Предложена математическая модель прямого внесения деформирующих искажений для функции нескольких переменных. Описаны особенности и алгоритмы внесения деформирующих искажений в решетчатые функции дискретных аргументов с использованием различных видов интерполяции, в том числе на основе применения радиально-базисных функций. Сформулировано и доказано

утверждение о верхней границе ошибки интерполяции деформируемой функции, возникающей при внесении деформирующих искажений. Установлено, что использование предложенного подхода внесения деформации в изображения на основе радиально-базисных функций дает сравнимые результаты с другими методами интерполяции, а в ряде случаев превосходит их по точности и является перспективным способом формирования обучающих выборок в ходе решения различных задач распознавания изображений.

3. Для задач синтеза алгоритмов распознавания цифровых сигналов предложена эквивалентная статистическая модель влияния деформирующих искажений на форму деформируемой функции, основанная на приближенном представлении нелинейного оператора деформации как оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. Доказан ряд свойств указанного оператора и даны оценки получающегося количества размещений в зависимости от ограничений, накладываемых на функцию внесения деформирующих искажений.

4. Дана математическая постановка задачи распознавания детерминированных сигналов в присутствии шумов и деформирующих искажений и осуществлен синтез оптимальных алгоритмов распознавания цифровых сигналов на основе использования модели оператора размещения с повторениями элементов исходного дискретного сигнала с добавлением аддитивной помеховой составляющей, вызванной ошибками квантования непрерывной функции деформации. Синтезированы и исследованы оптимальный, основанный на точном вычислении функций правдоподобия, и квазиоптимальный, основанный на использовании гауссовского приближения функций правдоподобия, алгоритмы распознавания. Установлено, что применение оптимального алгоритма, основанного на вычислении функций правдоподобия классов как взвешенной суммы условных плотностей распределения для всех возможных комбинаций перестановок элементов сигналов,

при существенных отличиях формы распознаваемых сигналов дает относительно небольшое преимущество по отношению к алгоритму, основанному на использовании гауссовского приближения функций правдоподобия. В то же время, выявлено, что в ситуациях, для которых характерны малые отличия формы распознаваемых сигналов, оптимальный алгоритм может давать заметное преимущество по сравнению с алгоритмом, основанным на использовании гауссовского приближения.

5. Предложена эквивалентная статистическая модель деформирующих искажений в задачах распознавания изображений и осуществлен синтез алгоритмов распознавания с использованием стандартных и смешанных непараметрических ядерных оценок функций правдоподобия классов. Доказаны свойства несмещенности и состоятельности смешанных ядерных оценок многомерных плотностей распределения вероятностей. Представлена обобщенная структура алгоритма обработки информации в интересах синтеза решающих правил при распознавании сигналов и изображений в случае, когда имеются ограничения вычислительной сложности или отсутствуют необходимые исходные данные. Показано, что использование алгоритмов, основанных на смешанных ядерных оценках, эквивалентно реализации процедуры искусственного размножения обучающих данных в соответствии со статистической моделью дополнительных аддитивных искажений, оказывающих влияние на распознаваемые изображения. Для синтезированных и для реальных изображений проведено моделирование и сравнение работы алгоритмов, основанных на использовании стандартных и смешанных непараметрических ядерных оценок, с алгоритмами, основанными на параметрических оценках и гауссовском приближении, и алгоритмом, основанным на использовании нейронной сети прямого распространения.

6. Описана алгоритмическая реализация метода Виолы-Джонса с использованием технологии параллельных вычислений на графических процессорах



CUDA. Получены количественные результаты по сравнению скорости работы параллельной и последовательной версий отдельных этапов работы алгоритма, которые выявили заметные преимущества параллельных версий над последовательными и показали перспективность применения последних как способа увеличения производительности метода Виолы-Джонса.

7. В интересах решения задач машинного обучения, связанных с обработкой изображений, на примере задачи поиска и обнаружения фронтальных лиц методом Виолы-Джонса рассмотрены три модели внесения деформирующих искажений для аугментации обучающей выборки: на основе применения при деформации гармонических функций; на основе смещения контрольных точек; на основе поиска оптического потока по методу Фарнебака при наложении фильтра энтропии. Для исходных и искусственно размноженных в соответствии с описанными моделями обучающих выборок проведены эксперименты по обучению алгоритмов распознавания изображений лиц методом Виолы-Джонса и даны результаты сравнения точности работы полученных детекторов. Полученные результаты показали применимость данного подхода к задачам построения детекторов обнаружения объектов, имеющих ярко выраженные структурные характеристики, для сокращения затрат на поиск и подготовку обучающих данных.

**Рекомендации по использованию.** Предложенные модели и синтезированные на их основе алгоритмы рекомендуется использовать в системах управления и принятия решений при обработке цифровых сигналов и изображений, подверженных действию деформирующих и аддитивных искажений. В качестве конкретной области применения можно указать решение задачи распознавания неоднородного потока биологических объектов по спектральным характеристикам в системах диагностики и экспресс-анализа. В приложении к использованию деформирующих искажений в задачах аугментации полученные результаты применимы в условиях недостаточного объема и несбалансированности обучающих данных при реализации

современных методов машинного обучения, в том числе в рамках парадигмы глубоких нейронных сетей.

**Перспективы дальнейшего развития** связаны с анализом и исследованием возможностей применения смешанных ядерных оценок для искусственного размножения данных и эффективности данного подхода для различных классов алгоритмов машинного обучения, а также с проведением синтеза и анализа алгоритмов распознавания сигналов и изображений в условиях одновременного наличия деформирующих и аппликативных искажений.

**Список использованных источников**

1. **Ларичев, О.И.** Теория и методы принятия решений. / О.И. Ларичев. – М.: Логос, 2000. – 296 с.
2. **Юдин, Д.Б.** Вычислительные методы теории принятия решений / Д.Б. Юдин. – М.: Либроком, 2014. – 320 с.
3. **Fukunaga, K.** Introduction to Statistical Pattern Recognition / K. Fukunaga. – 2nd ed. – San Diego, CA: Academic Press, 1990. – 592 p.
4. **Duda, R.O.** Pattern Classification / R.O. Duda, P.E. Hart, D.G. Stork. – 2nd ed. – Hoboken, NJ: Wiley, 2000. – 680 p.
5. **Пупков, К.А.** Методы классической и современной теории автоматического управления: Учебник в 5-и тт. Т.5: Методы современной теории автоматического управления / Под ред. К.А. Пупкова, Н.Д. Егупова. – 2-е изд., перераб. и доп. – М.: Издательство МГТУ им. Н.Э. Баумана, 2004. – 784 с.
6. **Haykin, S.** Neural Networks and Learning Machines / S. Haykin. – 3rd ed. – Upper Saddle River, NJ: Pearson Education, 2009. – 906 p.
7. **Begum, N.** Accelerating Dynamic Time Warping Clustering with a Novel Admissible Pruning Strategy / N. Begum, L. Ulanova, J. Wang, E. Keogh // Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2015. – Vol. 7(3). – P. 49-58. – DOI: 10.1145/2783258.2783286.
8. **Keogh, E.** Exact indexing of dynamic time warping / E. Keogh, C.A. Ratanamahatana // Knowledge and Information Systems. – 2004. – Vol. 7(3). – P. 358-386. – DOI: 10.1007/s10115-004-0154-9.
9. **Liao, T.W.** Clustering of Time Series Data — A Survey / T.W. Liao // Pattern Recognition. – 2005. – Vol. 38. – P. 1857-1874. – DOI: 10.1016/j.patcog.2005.01.025.
10. **Ding, H.** Querying and mining of time series data: experimental comparison of representations and distance measures / H. Ding, G. Trajcevski, P. Scheuermann, X.

- Wang, E. Keogh // Proceedings of the VLDB Endowment. – 2008. – Vol. 1(2). – P. 1542-1552. – DOI: 10.14778/1454159.1454226.
11. **Aghabozorgi, S.** Time-series Clustering – A Decade Review / S. Aghabozorgi, A.S. Shirkhorshidi, T. Ying Wah // Information Systems. – 2015. – Vol. 53. – P. 16-38. – DOI: 10.1016/j.is.2015.04.007.
  12. **Антипов, С.Г.** Исследование и разработка методов и алгоритмов обобщения знаний для систем поддержки принятия решений реального времени : дис. ... канд. тех. наук : 05.13.17 / С.Г. Антипов. – М., 2016. – 152 с.
  13. **Le Guennec, A.** Data Augmentation for Time Series Classification Using Convolutional Neural Networks / A. Le Guennec, S. Malinowski, R. Tavenard // ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Sep. 2016. – 2016. – P. 1-8.
  14. **Rabiner, L.** Fundamentals of Speech Recognition / L. Rabiner, B.H. Juang. – Englewood Cliffs, NJ: Prentice Hall PTR, 1993. – 507 p.
  15. **Corradini, A.** Dynamic time warping for offline recognition of a small gesture vocabulary / A. Corradini. // Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems. – 2001. – P. 82-89. – DOI: 10.1.1.200.2035.
  16. **Пономарев, Д.И.** Использование алгоритмов обнаружения паттернов для идентификации жестов оператора в записях управляющего сигнала манипулятора / Д.И. Пономарев // Труды МФТИ. – 2012. – Т. 4, № 3. – С. 187-197.
  17. **Vemulapalli, R.** Human Action Recognition by Representing 3D Skeletons as Points in a Lie Group / R. Vemulapalli, F. Arrate, R. Chellappa. // IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014 (CVPR 2014). – 2014. – P. 588-595. – DOI: 10.1109/CVPR.2014.82.
  18. **Kampouraki A.** Heartbeat Time Series Classification With Support Vector Machines / A. Kampouraki, G. Manis, C. Nikou // IEEE Transactions on Information Technology in Biomedicine, July 2009. – 2009. – P. 512-518. – DOI: 10.1109/TITB.2008.2003323.

19. **Wang, J.** Bag-of-words Representation for Biomedical Time Series Classification / J. Wang, P. Liu, M.F.H. She, S. Nahavandi, A. Kouzanid // *Biomedical Signal Processing and Control*, November 2013. – 2013. – P. 634-644. – DOI: 10.1016/j.bspc.2013.06.004.
20. **Бериков, В.Б.** Метод кластерного анализа разнотипных временных рядов / В.Б. Бериков, И.А. Пестунов, М.К. Герасимов // *Вычислительные технологии*. – 2015. – Т. 20, № 2. – С. 20-28.
21. **Алгазинов, Э.К.** Аппаратно-программный комплекс для анализа неоднородного потока объектов в системах фотосепарации реального времени / Э.К. Алгазинов, М.А. Дрюченко, А.А. Сирота, Д.А. Минаков, В.А. Шульгин // *Измерительная техника*. – 2014. – № 5. – С. 23-29.
22. **Тропченко, А.А.** Методы вторичной обработки и распознавания изображений. Учебное пособие / А.А. Тропченко, А.Ю. Тропченко. – СПб: Университет ИТМО, 2015. – 215 с.
23. **Keyzers, D.** Deformation models for image recognition / D. Keyzers, T. Deselaers, C. Gollan, H. Ney // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2007. – Vol. 29(8). – P. 1422–1435. – DOI: 10.1109/TPAMI.2007.1153.
24. **Felzenswalb, P.** A discriminatively trained, multiscale, deformable part model / P. Felzenswalb, D. McAllester, D. Ramanan // *IEEE Conference on Computer Vision and Pattern Recognition*, June 24-26, 2008 (CVPR 2008). – 2008. – P. 1-8. – DOI: 10.1109/CVPR.2008.4587597.
25. **Wiskott, L.** Face Recognition by Elastic Bunch Graph Matching / L. Wiskott, J.-M. Fellous, N. Kruger, C. von der Malsburg // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1997. – Vol. 19(7). – P. 775-779. – DOI: 10.1109/ICIP.1997.647401.

26. **Zhu, X.** Face Detection Pose Estimation and Landmark Localization in the Wild / X. Zhu, D. Ramanan // Proc. IEEE Conf. Computer Vision and Pattern Recognition. – 2012. – P. 2879-2886. – DOI: 10.1109/CVPR.2012.6248014.
27. **Nastar, Ch.** Generalized Image Matching: Statistical Learning of Physically-Based Deformations / Ch. Nastar, B. Moghadam, A. Pentland // Lecture Notes in Computer Science, Computer Vision, ECCV '96, B. Buxton and R. Cipolla (Eds.), Springer. – 1996. – Vol. 1064. – P. 589-598. – DOI: 10.1007/BFb0015569.
28. **Bichsel, M.** Automatic Interpolation and Recognition of Face Images by Morphing / M. Bichsel // The 2nd International on Automatic Face and Gesture Recognition. Killington, Vermont, IEEE Computer Society Press, Los Alamitos, CA. – 1996. – P. 128-135. – DOI: 10.1109/AFGR.1996.557254.
29. **Заказнов, Н.П.** Теория оптических систем: Учебник для студентов приборостроительных специальностей вузов / Н.П. Заказнов, С.И. Кирюшин, В.Н. Кузичев. – 3-е изд., перераб. и доп. – М.: Машиностроение, 1992. – 448 с.
30. **Brown, D.** Close Range Camera Calibration / D. Brown // Photogrammetric Engineering. – 1971. – Vol. 8. – P. 855-866. – DOI: 10.1.1.14.6358.
31. **Weng, J.** Camera Calibration with Distortion Models and Accuracy Evaluation / J. Weng, P. Cohen, M. Herniou // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1992. – Vol. 14(10). – P. 965-980. – DOI: 10.1109/34.159901.
32. **de Villiers, J.P.** Centi-pixel accurate real-time inverse distortion correction / J.P. de Villiers, F.W. Leuschner, R. Geldenhuys // In Proceedings of the 2008 International Symposium on Optomechatronic Technologies. – 2008. – Vol. 7266. – P. 1-8. – DOI: 10.1117/12.804771.
33. **Жимбуева, Л.Д.** Метод определения суммарной дисторсии цифровых изображений / Л.Д. Жимбуева // Компьютерная оптика. – 2011. – Т. 35, № 3. – С. 347-355.

34. **Кульчицкий, А.А.** Исследование моделей калибровки камер технического зрения для программной компенсации искажений в системах контроля геометрических параметров объектов / А.А. Кульчицкий, И.И. Абакумов // *Международ. научный журнал «Инновационная наука»*. – 2015. – № 10. – С. 86-90.
35. **Drap, P.** An Exact Formula for Calculating Inverse Radial Lens Distortions / P. Drap, J. Lefèvre // *Sensors*. – 2016. – Vol. 16(6). – P. 1-18. – DOI: 10.3390/s16060807.
36. **Friston, K.J.** Statistical parametric mapping: the analysis of functional brain images / K.J. Friston, W.D. Penny, J. Ashburner, S.J. Kiebel, T.E. Nichols. – London: Academic Press, 2006. – 656 p.
37. **Brown L.G.** A Survey of Image Registration Techniques / L.G. Brown // *ACM Computing Surveys*. – 1992. – Vol. 24(4). – P. 325-376. – DOI: 10.1145/146370.146374.
38. **Barron, J.** Performance of Optical Flow Techniques / J. Barron, D. Fleet, S. Beauchemin // *International Journal of Computer Vision*. – 1994. – Vol. 12(1). – P. 43-77. – DOI: 10.1007/BF01420984.
39. **Horn, B.** Determining Optical Flow / B. Horn, B. Schunk // *Artificial Intelligence*. – 1981. – Vol. 17. – P. 185-203. – DOI: 10.1016/0004-3702(81)90024-2 6.
40. **Farneback, G.** Two-Frame Motion Estimation Based on Polynomial Expansion / G. Farneback // *Proceedings of the 13th Scandinavian Conference on Image Analysis*. – Gothenburg, Sweden. – 2003. – P. 363-370. – DOI: 10.1007/3-540-45103-X\_50.
41. **Щербаков, В.В.** Алгоритм расчета оптического потока в задачах оценки параметров геометрических преобразований / В.В. Щербаков, А.Г. Гарганеев, И.В. Шакиров // *Доклады ТУСУРа*. – 2012. – № 2 (26), ч. 1. – С. 265-268.
42. **Fortun, D.** Optical flow modeling and computation: A survey / D. Fortun, P. Bouthemy, C. Kervrann // *Computer Vision and Image Understanding, Elsevier*. – 2015. – Vol. 134. – P. 1-21. – DOI: 10.1016/j.cviu.2015.02.008.

43. **Пономарев, Е.С.** Алгоритмы вычисления оптического потока в задаче определения собственного движения / Е.С. Пономарев, А.С. Григорьев // Сборник трудов конференции «Информационные технологии и системы 2015», 7-11 сентября, Олимпийская деревня, Сочи. – 2015. – С. 457-470.
44. **Антоненков, Д.А.** Метод определения скорости движения влекаемых наносов, основанный на программной обработке изображений водной среды / Д.А. Антоненков // Экологическая безопасность прибрежной и шельфовой зон моря. – 2017. – № 1. – С. 87-95.
45. **Wolberg G.** Image Morphing Survey / G. Wolberg // The Visual Computer. – 1998. – Vol. 14. – P. 360-372. – DOI: 10.1007/s003710050148.
46. **Zope, B.** A Survey of Morphing Techniques / B. Zope, S.B. Zope // International Journal of Advanced Engineering, Management and Science (IJAEMS) – 2017. – Vol. 3(2). – P. 81-87. – DOI: 10.24001/ijaems.3.2.15.
47. **Ruprecht, D.** Deformed Cross-dissolves for Interpolation in Scientific Visualization / D. Ruprecht, H. Müller // J. of Visualization and Computer Animation – 1994. – Vol 5(3). – P. 167–181. – DOI: 10.1002/vis.4340050304.
48. **Gao, P.** A Work Minimization Approach to Image Morphing / P. Gao, T.W. Sederberg // The Visual Computer. – 1998. – Vol. 14(8-9). – P. 390-400. – DOI: 10.1007/s003710050150.
49. **Liao, J.** Automating Image Morphing Using Structural Similarity on a Halfway Domain / J. Liao, R.S. Lima, D. Nehab, H. Hoppe, P.V. Sander, J.Yu // ACM Transactions on Graphics (TOG). – 2014. – Vol. 33(5). – P. 1-12. – DOI: 10.1145/2629494.
50. **Хмельнов, А.Е.** Реализация совмещения неточной модели рельефа речного дна с моделью надводного рельефа / А.Е. Хмельнов, А.С. Гаченко // Вестн. Бурят. гос. ун-та. Математика, информатика. – 2016. – № 2. – С. 77-84.
51. **Simard, P.Y.** Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis / P.Y. Simard, D. Steinkraus, J.C. Plat // In Proceedings of the



- Seventh International Conference on Document Analysis and Recognition. – 2003. – Vol. 2. – P. 958-962.
52. **Ciresan, D.C.** Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition / D.C. Ciresan, U. Meier, L.M. Gambardella, J. Schmidhuber // Neural Computation Conference. – 2010. – Vol. 22(12). – P. 3207-3220. – DOI: 10.1162/NECO\_a\_00052.
53. **Wong, S.C.** Understanding Data Augmentation for Vlassification: When to Warp? / S.C. Wong, A. Gatt, V. Stamatescu, M.D. McDonnell // DICTA 2016 Conference, Computer Vision and Pattern Recognition (cs.CV). – 2016. – P. 1-6. – DOI: 10.1109/DICTA.2016.7797091.
54. **Wang, J.** The Effectiveness of Data Augmentation in Image Classification using Deep Learning / J. Wang, L. Perez // arXiv: Computer Vision and Pattern Recognition (cs.CV). – 2017. – P. 1-8.
55. **Небаба, С.Г.** Алгоритм построения деформируемых 3D моделей лица и обоснование его применимости в системах распознавания личности / С.Г. Небаба, А.А. Захарова // Труды СПИИРАН. – 2017. – Вып. 52. – С. 157-179. – DOI: 10.15622/sp.52.8.
56. **Crispell, D.** Dataset Augmentation for Pose and Lighting Invariant Face Recognition / D. Crispell, O. Biris, N. Crosswhite, J. Byrne, J.L. Mundy // 2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR). arXiv: Computer Vision and Pattern Recognition (cs.CV). – 2017. – P. 1-9.
57. **Guennech, A.L.** Data Augmentation for Time Series Classification using Convolutional Neural Networks / A.L. Guennech, S. Malinowski, R. Tavenard // ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data. – Riva Del Garda, Italy, 2016. – P. 1-8.
58. **Жуковский, А.Е.** Синтез обучающей выборки на основе реальных данных в задачах распознавания изображений / А.Е. Жуковский, Н.А. Тарасова, С.А. Усилин, Д.П. Николаев // Информационные технологии и системы (ИТиС'12): сборник трудов конференции. – М., 2012. – С. 377-382.

59. **Chawla, N.** SMOTE: Synthetic Minority Over-sampling Technique / N. Chawla, K. Bowyer, L. Hall, W. Kegelmeyer // Journal of Artificial Intelligence Research. – 2002. – № 16. – P. 321-357. – DOI: 10.1613/jair.953.
60. **Chawla, N.V.** SMOTEBoost: Improving Prediction of the Minority Class in Boosting / N.V. Chawla, A. Lazarevic, L.O. Hall, K.W. Bowyer // 7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD). – Dubrovnik, Croatia, 2003. – P. 107-119. – DOI: 10.1007/978-3-540-39804-2\_12.
61. **Bunkhumpornpat, Ch.** DBSMOTE: Density-Based Synthetic Minority Over-sampling Technique / Ch. Bunkhumpornpat, K. Sinapiromsaran, Ch. Lursinsap // Applied Intelligence. – 2012. – Vol. 36(3). – P. 664-684. – DOI: 10.1007/s10489-011-0287-y.
62. **Акимов, А.В.** Моделирование деформирующих искажений на изображениях на основе радиально-базисных функций / А.В. Акимов, М.А. Дрюченко, А.А. Сирота // Мат. XIV Международ. научно-методической конференции «Информатика: проблемы, методология, технологии». – Воронеж: 2014. – Т.2. – С. 14-18.
63. **Акимов, А.В.** Моделирование деформирующих искажений объектов на изображениях с использованием радиально-базисных функций / А.В. Акимов, М.А. Дрюченко, А.А. Сирота // Мат. XV Международ. научно-технической конференции «Кибернетика и технологии XXI века» (С&Т 2014г.). – Воронеж: 2014. – Т.1. – С. 58-67.
64. **Акимов, А.В.** Модели и алгоритмы внесения деформирующих искажений на изображениях с использованием радиально-базисных функций / А.В. Акимов, М.А. Дрюченко, А.А. Сирота // Вестник ВГУ (Системный анализ и информационные технологии). – 2014. – № 1. – С. 130-137.
65. **Akimov, A.V.** Synthesis and Analysis of Algorithms for Digital Signal Recognition in Conditions of Deforming Distortions and Additive Noise / A.V. Akimov, A.A. Sirota // Radioelectronics and Communications Systems. – 2017. – Vol. 60(10). – P. 458-468. – DOI: 10.20535/S0021347017100041.

66. **Акимов, А.В.** Модели и алгоритмы распознавания цифровых изображений в условиях воздействия деформирующих и аддитивных искажений / А.В. Акимов, А.О. Донских, А.А. Сирота // Вестник ВГУ (Системный анализ и информационные технологии). – 2018. – № 1. – С. 104-118.
67. **Акимов, А.В.** Смешанные ядерные оценки многомерных распределений и их применение в задачах машинного обучения / А.В. Акимов, А.О. Донских, Д.А. Минаков, А.А. Сирота // Мат. Международ. науч. конференции «Математическое моделирование и информационные технологии в инженерных и бизнес-приложениях». – Воронеж: 2018. – С. 133-149.
68. **Акимов, А.В.** Анализ свойств смешанных ядерных оценок плотности распределения вероятностей в интересах решения задач машинного обучения / А.В. Акимов, А.О. Донских, А.А. Сирота // Мат. VII Международ. научно-технической конференции «Информационные технологии в науке, образовании и производстве» (ИТНОП-2018). – Белгород: 2018. – С. 23-30.
69. **Акимов, А.В.** Распределенная реализация алгоритма Viola-Jones с использованием графического процессора обработки данных в среде Matlab / А.В. Акимов, А.А. Сирота // Мат. XIII Международ. научно-методической конференции «Информатика: проблемы, методология, технологии». – Воронеж: 2013. – Т. 1. – С. 36-40.
70. **Акимов, А.В.** Разработка и исследование алгоритмов распознавания изображений на основе метода Виолы-Джонса с использованием технологии вычислений на графических процессорах CUDA / А.В. Акимов, А.А. Сирота // Вестник ВГУ (Системный анализ и информационные технологии). – 2014. – № 3. – С. 100-108.
71. **Акимов, А.В.** Искусственное размножение обучающих данных на основе деформирующих искажений при реализации алгоритмов распознавания лиц по методу Виолы-Джонса / А.В. Акимов, М.А. Дрюченко, А.А. Сирота // Мат. XVI Международ. научно-методической конференции «Информатика: проблемы, методология, технологии». – Воронеж: 2016. – С. 27-31.

72. **Акимов, А.В.** Модели и алгоритмы искусственного размножения данных для обучения алгоритмов распознавания лиц методом Виолы–Джонса / А.В. Акимов, А.А. Сирота // Компьютерная оптика. – 2016. – Т. 40, № 6. – С. 911-918. – DOI: 10.18287/2412-6179-2016-40-6-911-918.
73. **Gonzalez R.C.** Digital Image Processing Using MATLAB / R.C. Gonzalez, R.E. Woods, S.L. Eddins. – Second edition. – New Jersey: Prentice Hall. – 2009. – 826 p.
74. **Сизиков В.С.** Устойчивые методы обработки результатов измерений. Учебное пособие / В.С. Сизиков. — СПб.: «СпецЛит», 1999. — 240 с.
75. **Гоноровский, И.С.** Радиотехнические цепи и сигналы: Учебник для вузов / И.С. Гоноровский. – 4-е изд. – М.: Радио и связь, 1986. – 512 с.
76. **Widrow, B.** Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications / B. Widrow, I. Kollar. – Cambridge: Cambridge University Press, 2008. – 751 p.
77. **Сирота, А.А.** Методы и алгоритмы анализа данных и их моделирование в MATLAB / А.А. Сирота. – СПб.: БХВ-Петербург, 2016. – 384 с.
78. **Алгазинов, Э.К.** Анализ и компьютерное моделирование информационных процессов и систем / Э.К. Алгазинов, А.А. Сирота. – М.: Диалог-МИФИ, 2009. – 416 с.
79. **Кривенко, М.П.** Непараметрическое оценивание элементов байесовского классификатора / М.П. Кривенко // Информатика и ее применения. – 2010. – Т. 4, № 2. – С. 13-24.
80. **Абезгауз, Г.Г.** Справочник по вероятностным расчетам / Г.Г. Абезгауз, А.П. Тронь, Ю.Н. Копенкин, И.А. Коровина. – Издание второе. – М.: Воениздат, 1970. – 536 с.
81. **Крянев, А.В.** Математические методы обработки неопределенных данных / А.В. Крянев, Г.В. Лукин. – М.: Физмалит, 2003. – 216 с.

82. **Калиновский, И.А.** Обзор и тестирование детекторов фронтальных лиц / И.А. Калиновский, В.Г. Спицын // Компьютерная оптика. – 2016. – Т. 40, № 1. – С. 99-111. – DOI: 10.18287/2412-6179-2016-40-1-99-111.
83. **Viola, P.** Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M. Jones // Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition 2001. – 2001. – Vol. 1. – P. 511-518.
84. **Viola, P.** Robust Real-Time Face Detection / P. Viola, M. Jones // International Journal of Computer Vision. – Netherlands: Kluwer Academic Publishers. – 2004. – № 57(2). – P. 137-154.
85. **Papageorgiou, C.P.** A general framework for object detection / C.P. Papageorgiou, M. Oren, T. Poggio // International Conference on Computer Vision. – 1998. – P. 555-562.
86. **Freund, Y.** A Short Introduction to Boosting / Y. Freund, R. Schapire // Japan: Journal of Japanese Society for Artificial Intelligence. – 1999. – № 14(5). – P. 771-780.
87. **Jensen, O. H.** Implementing the Viola-Jones Face Detection Algorithm / O. H. Jensen. – DK-2800 Kongens Lyngby, Denmark: Technical University of Denmark Informatics and Mathematical Modelling, 2008. – 36 p.
88. **Steyvers M.** Morphing Techniques for Manipulating Face Images / M. Steyvers // Behavior Research Methods, Instruments, & Computers. – 1999. – Vol. 31(2). – P. 359-369. – DOI: 10.3758/BF03207733.
89. **Arad N.** Image warping by radial basis functions: applications to facial expressions / N. Arad, N. Dyn, D. Reifeld, Y. Yeshurun // CVGIP: Graph Models Image Processing. – 1994. – Vol. 56. – P. 161-172. – DOI: 10.1006/cgip.1994.1015.
90. CMU Image Data Base: face [Электронный ресурс]. – URL: <http://www.cs.cmu.edu/afs/cs/project/vision/vasc/idb/www/html/face/> (дата обращения 29.10.2018).