

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

На правах рукописи

Саввин Сергей Викторович

**АЛГОРИТМЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ С ДОСТИЖЕНИЕМ  
ЭФФЕКТА СВЕРХРАЗРЕШЕНИЯ НА ОСНОВЕ МЕТОДОВ  
ОПТИМАЛЬНОЙ ФИЛЬТРАЦИИ И МАШИННОГО ОБУЧЕНИЯ**

Специальность 2.3.1 – Системный анализ, управление и обработка информации

Диссертация  
на соискание ученой степени кандидата  
физико-математических наук

Научный руководитель  
доктор технических наук  
профессор А. А. Сирота

Воронеж – 2022

## Содержание

Введение.....	5
1 Анализ существующих методов и алгоритмов повышения разрешения изображений.....	17
1.1 Обзор известных методов повышения разрешения.....	18
1.1.1 Алгоритмы однокадрового сверхразрешения.....	19
1.1.2 Алгоритмы многокадрового сверхразрешения.....	21
1.1.3 Алгоритмы сверхразрешения видеоданных.....	24
1.2 Особенности повышения разрешения в условиях аппликативных помех .	25
1.3 Методы и алгоритмы построения многокадрового сверхразрешения в условиях аппликативных помех.....	29
1.3.1 Алгоритмы, основанные на модели спиновых стёкол.....	30
1.3.2 Алгоритмы, основанные на модели случайных марковских полей.....	32
1.3.3 Алгоритмы, основанные на фильтрации калмановского типа.....	33
1.3.4 Блочная обработка изображений в алгоритмах многокадрового сверхразрешения.....	34
1.4 Постановка задачи и общая схема проведения исследований в интересах построения алгоритмов многокадрового сверхразрешения в условиях аппликативных помех.....	36
Выводы по главе.....	43
2 Синтез и анализ алгоритмов построения сверхразрешения с использованием методов оптимальной фильтрации.....	45
2.1 Общая схема итеративной обработки последовательности изображений и ее реализация методом условно-линейной фильтрации.....	45
2.2 Синтез алгоритма многокадрового сверхразрешения в условиях аппликативных помех на основе методов оптимальной фильтрации.....	49
2.2.1 Модели состояний и наблюдений.....	49
2.2.2 Обновление оценки изображения высокого разрешения при обработке очередного кадра низкого разрешения при помощи условно-линейного фильтра.....	51

2.2.3	Задача оптимальной фильтрации в адаптивной постановке .....	53
2.3	Двухэтапная сегментация областей, подверженных воздействию аппликативных помех.....	57
2.3.1	Суперпиксельное представление изображений .....	58
2.3.2	EM-алгоритм и его применение для сегментации суперпикселей .....	67
2.3.3	Особенности сегментации ложных наблюдения и постобработка результатов кластеризации .....	70
2.3.4	Включение результатов двухэтапной сегментации в процесс построения сверхразрешения .....	72
2.4	Результаты экспериментального исследования эффективности алгоритмов, основанных на оптимальной фильтрации .....	75
2.4.1	Рассматриваемые методы и алгоритмы .....	75
2.4.2	Формирование тестового набора данных .....	76
2.4.3	Анализ результатов экспериментального сравнения алгоритмов .....	79
	Выводы по главе.....	84
3	Синтез и анализ алгоритмов построения многокадрового сверхразрешения на основе использования глубоких нейронных сетей.....	86
3.1	Использование свёрточных нейронных сетей для построения многокадрового сверхразрешения .....	86
3.2	Алгоритм построения многокадрового сверхразрешения на основе свёрточных нейронных сетей без использования независимой сегментации локальных областей закрытия .....	88
3.2.1	Схема обработки входной последовательности кадров.....	88
3.2.2	Оценка оптического потока между изображениями .....	89
3.2.3	Итеративное уточнение оценки изображения высокого разрешения при помощи свёрточной нейронной сети .....	90
3.2.4	Экспериментальное исследование.....	93
3.3	Алгоритм построения многокадрового сверхразрешения на основе глубоких нейронных сетей с сегментации областей, подверженных воздействию аппликативных помех.....	98
3.3.1	Схема обработки входной последовательности кадров.....	98

3.3.2 Семантическая сегментация аппликативных помех при помощи нейронной сети.....	99
3.3.3 Итеративное уточнение оценки при помощи глубокой сети в виде ациклического направленного графа .....	102
3.3.4 Экспериментальное исследование.....	106
Выводы по главе.....	112
4 Синтез и анализ алгоритмов построения сверхразрешения видеоданных .....	114
4.1 Общая схема обработки видеопоследовательности с использованием алгоритмов многокадрового сверхразрешения.....	114
4.2 Синтез алгоритма построения сверхразрешения видеоданных на основе алгоритма многокадрового сверхразрешения с использованием глубоких нейронных сетей.....	116
4.2.1 Схема обработки входной видеопоследовательности.....	116
4.2.2 Экспериментальное исследование.....	117
4.3 Синтез алгоритма построения сверхразрешения видеоданных на основе нейронной сети в виде направленного ациклического графа .....	121
4.3.1 Схема обработки входной видеопоследовательности кадров .....	121
4.3.2 Экспериментальное исследование.....	123
4.4 Применение алгоритмов сверхразрешения видеоданных для обработки реальных видеопоследовательностей .....	129
4.4.1 Сегментация движущихся объектов на изображениях .....	129
4.4.2 Экспериментальное исследование.....	131
Выводы по главе.....	134
Заключение .....	136
Список использованных источников .....	140

## Введение

**Актуальность темы работы.** Качество управления в современных системах в значительной степени обусловлено скоростью и достоверностью анализа данных, поступающих в контуре наблюдения внешних и внутренних параметров объектов управления. В связи с этим создание и развитие автоматизированных систем во многом опирается на внедрение новых, более совершенных алгоритмов обработки измерительной информации, в том числе и информации, представленной в виде изображений или видеоданных. Практически все средства наблюдения, обработки и анализа цифровых изображений требуют обеспечения детализации поступающих данных, отображающих анализируемые сцены, для получения адекватных результатов. Однако входные данные не всегда соответствуют необходимым требованиям, что проявляется, в том числе, в низком разрешении изображений, обусловленном, в основном, ограниченными возможностями средств регистрации. Регистрация более качественных изображений напрямую часто не представляется возможной в виду технических или экономических ограничений, в результате чего возникает необходимость дополнительной алгоритмической обработки изображений с целью повышения их качества. В этом плане значительный интерес представляют алгоритмы построения сверхразрешения (СР), осуществляющие обработку одного или нескольких цифровых изображений с целью получения результирующих изображений, обладающих более высоким разрешением и детализацией по сравнению с исходными.

В ходе обработки изображений, наряду с аддитивным шумом, зачастую возникает специфический вид помех, называемых аппликативными (АП) и проявляющимися в возникновении на изображениях областей аномальных наблюдений (локальных областей закрытия, ЛОЗ), которые носят случайный характер. Такие искажения также могут рассматриваться как фактор неоднородного снижения разрешения на поступающем изображении. В отличие от помех аддитивного характера, АП не могут быть устранены путём линейной

фильтрации и существенно влияют на качество обработки изображений, в том числе и в задачах построения СР.

В настоящее время разработано множество различных подходов и алгоритмов СР, однако лишь небольшая часть данных исследований [52–58] посвящена решению данной задачи в условиях наличия АП на входных изображениях. В этом плане можно утверждать, что на сегодняшний день данная задача не получила окончательного удовлетворительного решения. При этом следует выделить два базовых и принципиально альтернативных в методическом плане подхода к решению задачи построения алгоритмов СР.

Первый подход основан на применении постулируемых математических моделей данных и проведении статистического синтеза оптимальных или квазиоптимальных алгоритмов обработки изображений. При этом вводится ряд допущений и приближений относительно используемых математических моделей. В частности, это касается использования гауссовской модели данных, широко используемой в известных работах, посвященных синтезу и анализу алгоритмов СР на основе методов оптимальной марковской фильтрации [39-42, 55-58], зачастую используемых при реализации контуров наблюдения автоматических систем управления. Однако такая модель во многих случаях не является в полной степени адекватной, особенно в задачах обработки изображений, что ставит вопрос о допустимости ее применения в целом при решении рассматриваемых задач. Также в недостаточной степени рассмотрены вопросы синтеза алгоритмов СР при наличии АП в условиях статистической неопределенности относительно параметров используемых математических моделей наблюдений (межкадровых сдвигов, дисперсии размытия системы формирования изображений и т.п.).

Для решения данного класса задач, как, впрочем, и других актуальных задач обработки изображений, в последние годы всё большее внимание привлекает другой подход, основанный на применении методов глубокого машинного обучения и, в частности, свёрточных нейронных сетей, а также их различных модификаций. Данный подход свободен от задания в явном виде математических моделей данных, но требует использования значительных по объему обучающих

выборки, содержащих примеры анализируемых изображений. В этом плане известны работы [22-25, 43-46, 48-50], в которых глубокие нейронные сети успешно применяются для решения задачи СР в традиционной постановке, но, при этом, без учета факторов воздействия АП.

Очевидно, также, что в настоящее время назрела необходимость сопоставления результатов, получаемых при реализации указанных подходов к построению алгоритмов СР в условиях АП, в интересах формирования объективной оценки их возможностей, преимуществ и недостатков.

В связи с изложенным, в качестве перспективного направления исследований представляется направление, связанное с совершенствованием алгоритмов построения СР изображений в условиях АП, опирающихся на применение методов оптимальной динамической фильтрации при статистической неопределенности относительно параметров используемых математических моделей и методов машинного обучения, включая глубокие нейронные сети. Данные алгоритмы могут использоваться для повышения качества обработки информации в системах управления, наблюдения, анализа изображений и видеоданных.

**Степень разработанности темы диссертации.** Задача обработки последовательности изображений с целью повышения их разрешения рассматривалась в работах И.М. Журавеля, М.И. Савельева, А.В. Ращупкина, Б.А. Михайлова, Н.Н. Ушаковой, H.S Hou, X. Li, R. Timofte, V. De Smet, S. Schuler, J.-B. Huang, G. Freedman, J. Yang, N.K. Bose, G.K. Chantas, J. Tian, M. Irani, S. Peleg, H. Ji, C. Fermüller, C. Newland, A.J. Patti, M.I. Sezan, A.M. Tekalp, D. Gray, D. Gibbins, Z. Wei, F. Tao, W. Jun и др. [1-51]. При этом одним из наиболее распространённых подходов к организации обработки является итеративное уточнение текущей оценки выходного изображения высокого разрешения при обработке очередного изображения низкого разрешения из входной последовательности [36-42].

Большинство работ по данной теме не рассматривают вопрос построения СР в условиях наличия ложных наблюдений, вызванных воздействием АП, на

входных изображениях. Наличие таких областей на обрабатываемых изображениях можно рассматривать как дополнительный фактор понижения разрешения и устранять их воздействие одновременно с построением СР. В этом плане представляют интерес исследования, проведённые W. Fukuda, A. Kanemura, S. Maeda, S. Ishii, C.S. Negi, K. Mandal, R.R. Sahay, M.S. Kankanhalli, А.Ю. Иванковым и А.А. Сиротой [52-58], целью которых являлось устранение негативного воздействия АП с одновременным повышением разрешения входных изображений. В данных работах рассматривались такие подходы к решению данной задачи как модель спиновых стёкол, модель случайных марковских полей и оптимальная линейная и нелинейная динамическая фильтрация калмановского типа. При этом наиболее перспективным подходом с точки зрения дальнейших исследований в данной области представляется использование методов оптимальной и квазиоптимальной динамической фильтрации, как имеющих наиболее адекватную модель обрабатываемых изображений и потенциал для улучшения (например, использование адаптивного подхода к фильтрации). Развитие данного подхода, на наш взгляд, связано, прежде всего, с учётом статистической неопределённости относительно параметров используемых при синтезе алгоритмов динамической фильтрации математических моделей.

В последние годы всё большую актуальность приобретает использование нейросетевых технологий для решения задачи построения СР. Данный подход находит своё отражение в работах C. Dong, C.C. Loy, K. He, X. Tang, J. Kim, J.K. Lee, K.M. Lee, C. Ledig, L. Theis, F. Huszár, M. Sajjadi, R. Vemulapalli, M. Brown, M. Kawulok, P. Benecki, S. Piechaczek, A.B. Molini, D. Valsesia, G. Fracastoro и др. [22-25, 43-46, 48-50]. Высокая точность и быстрое действие синтезированных в ходе этих исследований алгоритмов, а также их независимость от сложных математических моделей входных данных обуславливает перспективность использования нейронных сетей и для решения задачи построения СР в условиях наличия АП на изображениях входной последовательности. При этом наибольший интерес в данной области представляют свёрточные нейронные сети, хорошо зарекомендовавшие себя при



решении задачи построения СР, как, впрочем, и других задач, связанных с обработкой цифровых изображений [75, 83-86, 89-91].

Таким образом, тема работы, посвящённая исследованию и синтезу моделей и алгоритмов обработки изображений с достижением эффекта сверхразрешения на основе методов оптимальной фильтрации и машинного обучения, является актуальной. Тема работы связана с научными направлениями ФГБОУ ВО «ВГУ» «Математическое моделирование, программное и информационное обеспечение, методы вычислительной и прикладной математики и их применение к фундаментальным исследованиям в естественных науках», «Проблемы передачи, получения, обработки и хранения информации».

**Цель и задачи исследования.** Целью работы является совершенствование алгоритмов обработки информации в задачах построения сверхразрешения для последовательностей изображений и видеоданных в условиях аппликативных помех на основе методов оптимальной фильтрации и машинного обучения с использованием глубоких нейронных сетей.

Для достижения цели в работе рассматриваются и решаются следующие задачи.

1. Анализ существующих подходов к построению СР в традиционной постановке задачи и при наличии аномальных наблюдений (АП) на входных изображениях.

2. Синтез и анализ алгоритмов многокадрового СР в условиях АП на основе методов оптимальной условно-линейной и нелинейной (адаптивной) динамической фильтрации и алгоритмов сегментации областей, подверженных воздействию АП на входных изображениях.

3. Построение алгоритмов многокадрового СР в условиях АП на основе комплекса глубоких нейронных сетей, реализующих различные этапы обработки последовательности входных изображений, включая, в том числе, и отдельный этап сегментации локальных областей закрытия на изображениях, вызванных воздействием АП.

4. Проведение сопоставительного анализа качества обработки изображения в задачах построения СР с использованием алгоритмов, синтезированных в рамках методов оптимальной фильтрации и методов глубокого обучения.

5. Синтез и анализ алгоритмов СР видеоданных в условиях АП на основе использования глубоких свёрточных нейронных сетей.

**Объект исследования.** Объектом исследования являются системы обработки информации и реализуемые в них процессы обработки цифровых изображений.

**Предмет исследования.** Модели и алгоритмы обработки цифровых изображений, подверженных воздействию АП, в задачах повышения разрешения и детализации входных изображений.

**Методы исследования.** При решении поставленных в диссертации задач использовались методы математического анализа, теории вероятностей и математической статистики, теории случайных полей, линейной алгебры, теории графов, теории обратных задач, методы оптимизации, методы оптимальной марковской фильтрации, методы цифровой обработки изображений, методы машинного обучения, а также технологии статистического имитационного моделирования.

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

1. Проведена декомпозиция итеративного процесса обработки информации при анализе последовательности изображений, определяющая общую структуру комплексного алгоритма построения многокадрового СР в условиях АП. Поставлена и решена задача синтеза комплексного алгоритма построения СР на основе последовательности искажённых воздействием АП изображений низкого разрешения, отличающегося использованием принципа адаптивной нелинейной фильтрации для учёта статической неопределённости параметров модели наблюдений, а также процедуры независимой суперпиксельной сегментации локальных областей закрытия на входных изображениях для минимизации воздействия АП.

2. Предложен и исследован двухэтапный алгоритм сегментации изображений на области полезных и ложных наблюдений, основанный на проведении суперпиксельной сегментации каждого изображения низкого разрешения и выполнении кластерного анализа полученных суперпикселей с помощью EM-алгоритма.

Получены математические соотношения для вероятностей принадлежности суперпикселей выделяемым областям, обеспечивающие включение результатов сегментации в вычисление весовых коэффициентов оптимального условно-линейного фильтра, используемого в комплексном алгоритме построения СР.

3. В рамках проведенной декомпозиции процесса обработки информации при построении многокадрового СР в условиях АП обоснована структура алгоритмов построения СР, отличающихся использованием ансамбля глубоких нейронных сетей свёрточного типа и сетей в виде направленного ациклического графа (DAG-сетей), выполняющих преобразования информации на всех этапах, предусмотренных общей схемой построения итеративных алгоритмов СР.

Проведено экспериментальное сравнение синтезированных алгоритмов с алгоритмом, базирующимся на применении методов оптимальной фильтрации. Установлено преимущество алгоритмов, реализованных в виде ансамбля предложенных глубоких нейронных сетей по качеству получаемых изображений высокого разрешения и по времени обработки.

4. Предложены и исследованы алгоритмы построения СР видеоданных как последовательности видеок кадров, искажённых воздействием АП. Первый алгоритм основан на адаптации рассмотренного ранее итеративного алгоритма построения многокадрового СР изображений на основе глубоких нейронных сетей с использованием для получения очередного изображения высокого разрешения совокупности кадров, выделяемых в скользящем временном окне. Для повышения быстродействия во втором алгоритме использована многовходовая DAG-сеть специальной архитектуры, обеспечивающая обработку одновременно всех кадров, выделяемых в скользящем окне. Предложен относительно простой и эффективный алгоритм сегментации локальных областей

закрытия на кадрах видеопоследовательности, основанный на использовании априорного неискаженного изображения сцены и выполнении набора морфологических преобразований.

**Тематика работы** соответствует областям исследования специальности 2.3.1 Системный анализ, управление и обработка информации.

**Теоретическая и практическая значимость.** Основные теоретические и экспериментальные результаты работы отвечают потребностям важного направления развития систем обработки цифровых изображений – получения изображений, обладающих более высоким разрешением и детализацией, в условиях наличия на входных изображениях случайных аномальных наблюдений и ограниченных технических возможностей средств регистрации и передачи данных.

Представленные теоретические и экспериментальные результаты, позволяют проводить сопоставительный анализ альтернативных в методическом плане подходов к построению алгоритмов обработки информации рассматриваемого класса и выбор конкретного алгоритма с учетом возникающих на практике ограничений. Впервые дана объективная оценка преимуществ и недостатков подходов, базирующихся на применении методов оптимальной динамической фильтрации и машинного обучения с использованием глубоких нейронных сетей.

Синтезированные модели и комплексные алгоритмы повышения разрешения и качества изображений на основе методов оптимальной фильтрации и машинного обучения позволяют справиться с негативным воздействием АП на изображениях и обладают достаточной универсальностью в плане входных и выходных разрешений, а также размеров обрабатываемой последовательности изображений, что делает возможным их применение в широком спектре задач компьютерного зрения.

Результаты, полученные в ходе выполнения диссертационной работы, имеют практическое значение для разработки программного обеспечения в интересах автоматической обработки цифровых изображений в различных

информационных системах (например, системах видеонаблюдения, аэрофотосъёмки и т.д.).

### **Положения и результаты, выносимые на защиту.**

1. Модели и комплексные алгоритмы многокадрового сверхразрешения в условиях наличия аппликативных помех на изображениях, основанные на использовании методов оптимальной динамической фильтрации и позволяющие учесть неопределенности относительно параметров используемой математической модели данных и результатов независимой суперпиксельной сегментации локальных областей закрытия на поступающих изображениях.

2. При выполнении сегментации локальных областей на изображениях, подверженных воздействию аппликативных помех, на основе предложенного подхода целесообразно выполнение двух этапов обработки суперпиксельной сегментации и кластерного анализа суперпикселей при помощи EM-алгоритма, что обеспечивает адекватное включение вероятностей принадлежности суперпикселей областям полезных или ложных наблюдений на каждом кадре последовательности в весовую матрицу алгоритма динамической фильтрации.

3. Итеративные алгоритмы многокадрового сверхразрешения в условиях наличия аппликативных помех, основанные на использовании ансамбля свёрточных нейронных сетей и предложенных в работе сетей в виде направленного ациклического графа, выполняющих различные этапы общей схемы обработки изображений, включая оценку смещений между изображениями, сегментацию локальных областей закрытия и улучшение качества изображений.

4. Результаты сопоставительного анализа альтернативных подходов к построению алгоритмов сверхразрешения в условиях аппликативных помех, свидетельствующие о наличии преимущества алгоритмов, реализованных в виде ансамбля предложенных глубоких нейронных сетей, по качеству получаемых изображений высокого разрешения и по времени обработки по отношению к алгоритмам, базирующимся на применении методов оптимальной фильтрации.

5. При построении обучаемых алгоритмов сверхразрешения видеоданных в составе используемого ансамбля глубоких свёрточных нейронных сетей для

получения оценки текущего кадра целесообразно с точки зрения повышения быстродействия использовать многоходовые сети в виде направленного ациклического графа, адаптированные для обработки совокупности предшествующих и последующих кадров видеопоследовательности, выделяемых в скользящем временном окне.

**Реализация результатов работы.** Полученные в диссертации результаты в части разработки алгоритмов обработки цифровых изображений использованы при выполнении НИР «Туманность-ВГУ», «Туба-ВГУ» а также в учебном процессе ФГБОУ ВО «Воронежский государственный университет».

**Степень достоверности результатов работы.** Результаты исследований, сформулированные в диссертационной работе и основанные на теоретических и экспериментальных методах исследований, взаимно дополняют друг друга и согласуются между собой. Корректность синтезированных моделей и комплексных алгоритмов повышения качества и разрешения входных изображений подтверждается результатами проведённых экспериментов на реальных изображениях.

Полученные в работе выводы и результаты имеют понятную трактовку, ясное физическое истолкование и совпадают в ряде частных случаев с известными результатами, полученными другими авторами.

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

**Апробация работы.** Приведённые в диссертации выводы и рекомендации были сформулированы в ходе всестороннего обсуждения результатов исследований. Основные положения работы докладывались и обсуждались: на XIX и XX Международных конференциях «Информатика: проблемы, методология, технологии» (г. Воронеж) в 2019 и 2020 гг.; на Международной научно-технической конференции «Актуальные проблемы прикладной математики, информатики и механики» (г. Воронеж) в 2019 г.; на Международной конференции «International Conference on Digital Image & Signal Processing» (Oxford, United Kingdom) в 2019 г.; на 2-ой Международной конференции «2nd

International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency» (г. Липецк) в 2020 г.; на ежегодных научных сессиях и семинарах факультета компьютерных наук ВГУ.

**Публикации.** По теме диссертационной работы опубликовано 11 [59-67, 78, 79] работ, из них 6 – в изданиях, рекомендованных ВАК, в том числе 4 статьи в журналах, индексируемых в БД Scopus и Web of Science.

Основные результаты по теме работы получены лично автором и опубликованы в соавторстве с научным руководителем. В совместных работах последнему принадлежит постановка задачи, определение направлений и общая идеология проведения исследований. Соискателю принадлежит проведение рассуждений, обоснование утверждений и используемых математических моделей, разработка и исследование алгоритмов обработки и изображений, организация и проведение экспериментов по их исследованию, анализ и интерпретация полученных результатов. В работах [59, 60, 62, 65], выполненных совместно с другими соавторами, соискателем выполнены разработка, программная реализация и исследование алгоритмов построения многокадрового СР в условиях аппликативных помех.

**Структура и объём работы.** Диссертационная работа состоит из введения, четырёх глав, заключения и списка литературы из 100 наименований. Объём работы составляет 151 страницу текста, включая 51 рисунок и 16 таблиц.

В первой главе диссертации производится анализ известных методов и алгоритмов обработки цифровых изображений с целью повышения их разрешения и качества, включая случай, когда на входных изображениях присутствуют аномальные наблюдения, вызванные АП. Рассматриваются особенности фильтрации АП на входных изображениях в ходе восстановления изображения высокого разрешения. Предлагается общая схема организации исследования в интересах синтеза алгоритмов многокадрового СР в условиях АП.

Во второй главе предлагается и реализуется общая схема итеративной обработки последовательности изображений с целью построения СР, включающая в себя этап независимой сегментации областей ложных наблюдений

на обрабатываемых изображениях. Описываются и исследуются алгоритмы многокадрового СР на основе оптимальной динамической фильтрации, в том числе в условиях неопределённости относительно параметров математической модели. Рассматривается алгоритм двухэтапной сегментации областей аномальных наблюдений на входных изображениях на основе суперпиксельного представления изображений.

В третьей главе описываются и исследуются комплексные алгоритмы многокадрового СР на основе технологий машинного обучения (глубоких свёрточных сетей) на различных этапах обработки, включая повышение качества изображений, реализующие ранее описанный итеративный подход к получению итогового изображения. Рассматриваются два подхода к синтезу таких алгоритмов: как с включением этапа независимой сегментации входных изображений на полезные и ложные наблюдения, так и без его использования.

В четвёртой главе описываются и исследуются алгоритмы СР видеоданных на основе использования глубоких нейронных сетей для повышения качества изображений. Рассматриваются два подхода к синтезу таких алгоритмов: на основе прямой адаптации ранее рассмотренного алгоритма многокадрового СР и на основе его модификации, учитывающей специфику обработки видеопоследовательностей в скользящем окне наблюдения.

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



## 1 Анализ существующих методов и алгоритмов повышения разрешения изображений

Под задачей построения сверхразрешения (СР) в дальнейшем будет пониматься процесс обработки последовательности входных цифровых изображений низкого разрешения (НР), в результате которого получается одно или несколько изображений более высокого разрешения (ВР). Конечной целью такой обработки является получение изображений, имеющих большую степень детализации и, в целом, лучшего качества по сравнению с исходными.

Другим обстоятельством, влияющим на качество и детализацию изображений, является наличие на них помех различного характера, которые возникают как при регистрации изображений, так и при их передаче по каналам связи. Данные помехи можно рассматривать как дополнительный фактор снижения разрешения. Большинство алгоритмов СР учитывают это и предполагают также и устранение шумов на исходных изображениях, однако в литературе обычно рассматриваются простейшие помехи аддитивного характера, в то время как более сложные модели помех (например, аппликативных помех) не получают должного внимания.

В случае учёта только аддитивных помех, общую модель получаемых наблюдений можно представить следующим образом:

$$y = \mathbf{F}x + \mu,$$

где  $y$  – наблюдаемое изображение НР;  $x$  – неизвестное изображение ВР;  $\mu$  – аддитивный шум;  $\mathbf{F}$  – оператор понижения разрешения, включающий в себя прореживание, смещение камеры, размытие и прочие возможные искажения. Как правило, оператор  $\mathbf{F}$  не является известным с точностью до параметров и алгоритмы СР должны включать в себя оценку этих параметров (например, оценку смещений камеры между кадрами при построении многокадрового СР или при анализе видеоряда).

Стоит отметить, что в данной работе не рассматриваются вопросы в области совершенствования технических средств регистрации, преобразования и передачи

цифровых изображений, в которых также может решаться задача повышения качества изображений и устранения помех. Рассматриваемые далее методы и алгоритмы СР предполагают работу с уже имеющимися техническими системами и реализуют обработку получаемых от них цифровых изображений.

В настоящее время разработано большое количество алгоритмов СР, основывающихся на разнообразных методах и моделях: от классических методов обработки сигналов до использования свёрточных нейронных сетей. Далее эти алгоритмы, представленные в известной литературе, будут рассмотрены более подробно.

### **1.1 Обзор известных методов повышения разрешения**

В общем случае количество входных и выходных изображений для алгоритмов СР может быть произвольным, но на практике можно выделить несколько наиболее распространённых случаев (рисунок 1.1).

1. Однокадровое СР: на вход алгоритма подаётся одно изображение НР, разрешение которого повышается путём специальной обработки.

2. Многокадровое СР: на вход подаётся последовательность изображений НР (называемых кадрами), отображающих одну и ту же сцену с межкадровыми субпиксельными смещениями, а в результате формируется одно выходное изображение ВР, отображающее исходную сцену с большей детализацией.

3. СР видеоданных: на вход подаётся последовательность изображений НР (кадров), каждому из которых соответствует своё выходное изображение ВР.

Алгоритмы многокадрового и видео-СР используют имеющуюся избыточность информации за счёт множества изображений, отображающих одну и ту же сцену, для получения более точного результата и зачастую используют в своей основе методы и алгоритмы однокадрового СР.

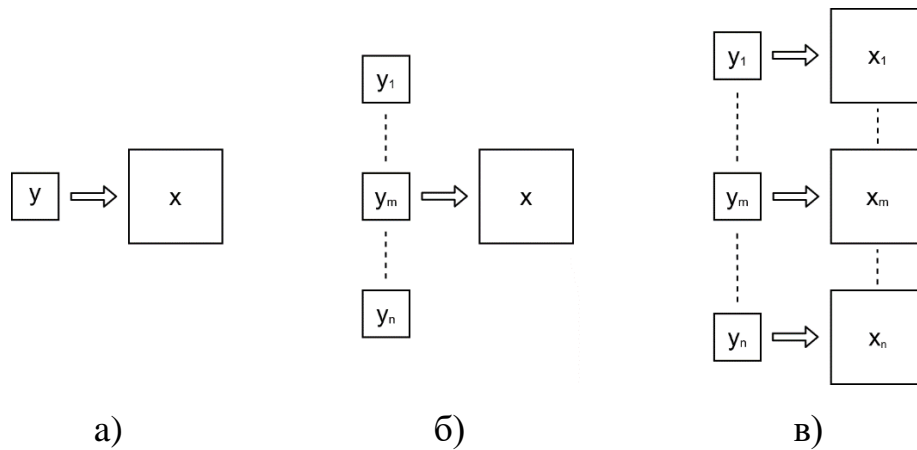


Рисунок 1.1 – Схемы построения СР: а) однокадровое СР; б) многокадровое СР;  
 в) СР видеоданных

### 1.1.1 Алгоритмы однокадрового сверхразрешения

Простейшими способами повышения разрешения изображения являются классические методы интерполяции: метод ближайшего соседа [11], билинейная интерполяция [12], бикубическая интерполяция [13]. Данные методы в виду своей простоты отличаются крайне высокой производительностью, однако не обладают достаточной точностью. Тем не менее, алгоритм бикубической интерполяции широко применяется в более комплексных подходах к СР для получения исходной оценки изображения ВР, которая улучшается в дальнейшем при помощи методов многокадрового СР. Получение более точных результатов при помощи интерполяции возможно за счёт использование дополнительных средств при построении выходного изображения ВР: например, применение вейвлетного преобразования к изображению НР, обнаружение границ между объектами и т.д. [14-16].

Широко распространённым подходом к решению задачи однокадрового СР является использование базы образцов, состоящей из множества пар, где каждому образцу НР сопоставляется образец ВР [17-21]. В этом случае, получение результирующего изображения ВР осуществляется путём разбиения исходного изображения НР на большое количество пересекающихся фрагментов (как правило, квадратных блоков небольшого размера  $3 \times 3$  или  $5 \times 5$ ). Затем каждый

этот фрагмент заменяется соответствующим ему образцом ВР из заранее построенной базы, и полученные таким образом фрагменты ВР объединяются в одно изображение ВР.

В рамках данного подхода можно выделить два способа построения базы образцов:

– внешняя база строится на основе некоторой обучающей выборки, состоящей из множества изображений ВР [17-19];

– внутренняя база строится на основе единственного входного изображения НР (данный процесс базируется на свойстве самоподобия, присущего естественным изображениям) [20, 21].

В последние годы для решения различных задач, связанных с обработкой данных, всё чаще используются методы, основанные на использовании свёрточных нейронных сетей. Задача построения однокадрового СР не является исключением, и в настоящее время существуют различных архитектуры нейронных сетей, осуществляющих преобразования изображения НР в изображение ВР [22-25]. Данные работы являлись закономерным развитием идей исследований, основанных на применении баз образцов.

Впервые такой подход был использован в [22], где глубокая свёрточная нейронная сеть SRCNN осуществляет преобразование входного изображения НР в изображение ВР напрямую. В дальнейшем данная архитектура была улучшена в [23] с целью повышения быстродействия.

В [24] были предложены усовершенствования подхода на основе свёрточной нейронной сети: глубина сети была значительно повышена (количество обучаемых слоёв было увеличено с 3 до 20), а с её помощью оценивается не изображение ВР напрямую, а остаточное изображение ВР, которое необходимо сложить с бикубической интерполяцией исходного изображения НР. Эти изменения позволили значительно увеличить точность получаемых результатов.

В [25] для получения изображений ВР применяется генеративно-сопоставительная (GAN) сеть, состоящая из двух искусственных нейронных сетей,

обучаемых совместно: первая генерирует изображения ВР, а вторая пытается отличить истинные образцы изображений ВР, от изображений, полученных при помощи первой сети, чтобы повысить качество создаваемых её изображений. Данный подход позволяет добиться восстановления высокочастотных деталей, которые могут быть пропущены при использовании других алгоритмов.

### **1.1.2 Алгоритмы многокадрового сверхразрешения**

В случае построения многокадрового СР на вход алгоритма подаётся последовательность изображений НР, отображающих одну и ту же сцену с субпиксельными смещениями между кадрами, в результате чего появляется дополнительная информация об исходных данных. Данные смещения могут быть вызваны как изменением положения камеры с течением времени, так и движением объектов на изображении. В результате этого, в отличие от алгоритмов однокадрового СР, методы обработки последовательности изображений, как правило, включают в себя дополнительный этап анализа изображений НР, который заключается в оценке смещений между кадрами с целью их дальнейшего выравнивания относительно друг друга для получения единственного изображения ВР на выходе алгоритма. Данный шаг может осуществляться различными способами. Например, может оцениваться оптический поток между двумя кадрами или матрицы сдвигов и поворотов.

Алгоритмы многокадрового СР, базирующиеся на обработке изображений в частотной области [26-28], основываются на свойствах используемых преобразований, таких как дискретное преобразование Фурье или дискретное косинусное преобразование, и осуществляют слияние кадров НР в частотной области. С одной стороны, эти методы отличаются быстродействием, однако простота используемых моделей негативно сказывается на точности результатов. С целью улучшения качества получаемых результатов в литературе также рассматривается подход с использованием вейвлет-преобразования исходных кадров НР [29, 30].

Методы многокадрового СР, основанные на интерполяции [31, 32], как правило, состоят из трёх шагов: регистрация кадров НР, получение изображения ВР путём интерполяции и улучшение качества полученного изображения ВР (например, повышение его резкости). Методы данной группы вычислительно эффективны, однако не обладают высокой точностью.

В рамках статистического подхода [33-35] изображение ВР оценивается как максимум апостериорной вероятности по наблюдаемым кадрам НР при помощи методов байесовского вывода. При этом возможно использование априорной информации об изображении и методов регуляризации для наложения ограничений на искомую оценку ВР. Изображения в данном подходе рассматриваются как случайные марковские поля [99].

Распространённым подходом является итеративное уточнение некоторой начальной оценки ВР по мере обработки очередного входного изображения НР. В частности, данный подход реализуется в [36, 37] при помощи алгоритма итеративной обратной проекции (Iterative Back-Projection), а в [38] используется метод проекции на выпуклые множества (Projection onto Convex Sets), позволяющий использовать априорные сведения об изображении ВР.

В [39-42] рассматриваются алгоритмы, основанные на оптимальной линейной фильтрации калмановского типа [68]. В них также реализуется итеративное уточнение начальной оценки ВР по мере обработки нового кадра НР. Алгоритмы СР такого типа дают результаты высокой точности, однако имеют большую вычислительную сложность.

Также как и в случае однокадрового СР, в последние годы для решения задачи построения многокадрового СР всё чаще применяются методы, основанные на нейросетевых технологиях [43-46]. При этом можно выделить ряд отличающихся подходов к организации обработки поступающих на вход изображений НР, которые зачастую предполагают использование сразу нескольких нейронных сетей для решения различных задач, возникающих в ходе обработки.

Например, в [43] описывается итеративный алгоритм FRVSR, использующий две глубокие нейронные сети для регистрации изображений НР и улучшения их качества. На вход первой нейронной сети подаются два кадра НР, и она оценивает оптический поток между ними, что позволяет в дальнейшем совместить данные кадры. Вторая нейронная сеть принимает на вход очередной кадр НР, увеличенный до целевого разрешения при помощи бикубической интерполяции, а также текущую оценку изображения ВР. На выходе получается обновлённая оценка ВР, которая будет использоваться на следующем шаге итерации, пока не будут обработаны все кадры НР.

Подход, описанный в [44], основывается на независимом повышении разрешения каждого из исходных кадров НР при помощи одного из алгоритмов однокадрового СР, использующих искусственные нейронные сети [22, 23]. Затем полученные на первом этапе изображения ВР объединяются в одно выходное изображение ВР при помощи другой свёрточной нейронной сети.

В [45] используется схожий подход, в рамках которого осуществляется независимая обработка всех кадров НР при помощи нейронной сети. Однако в её результате получается не изображение ВР, а некоторый набор признаков, соответствующий каждому из кадров НР. Затем все полученные признаки изображений НР объединяются в одно остаточное изображение ВР при помощи другой нейронной сети. Результирующее изображение ВР получается путём сложения начальной оценки ВР, полученной в ходе усреднения входных изображений НР, с вычисленным ранее остаточным изображением.

Алгоритм, рассмотренный в [46], также как и [45] базируется на обработке признаков, полученных в ходе обработки кадров НР. Однако дальнейшая работа с этими признаками осуществляется рекурсивным образом при помощи особым образом сконструированной нейронной сети.

### 1.1.3 Алгоритмы сверхразрешения видеоданных

Алгоритмы СР для видеоряда во многом схожи с алгоритмами многокадрового СР. Основное отличие между данными задачами заключается в том, что при обработке видеопоследовательности для получения очередного кадра ВР используется соответствующий ему кадр НР, а также несколько ( $m$ ) кадров НР до и после него. Таким образом, несмотря на то что, длина всего видео (последовательности кадров) может быть неограниченной, обработка кадров осуществляется при помощи «скользящего окна», обрабатывающего  $2m+1$  изображений на каждом шаге (рисунок 1.2).

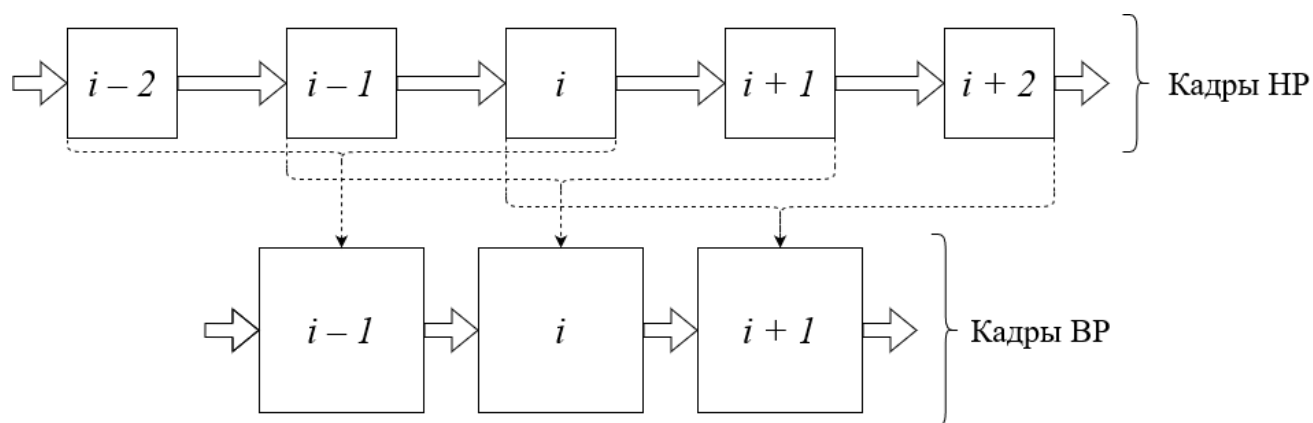


Рисунок 1.2 – Последовательная обработка кадров при построении СР видеоданных в случае  $m=1$

Так, данный подход находит своё отражение в [47], где для получения максимума апостериорной вероятности итогового изображения ВР используется метод наименьших квадратов с итерационным повторным взвешиванием. В процессе работы данного алгоритма в рамках статистического подхода осуществляется оценка оптического потока, аддитивного шума и ядер размытия.

Описанный выше подход реализуется при помощи глубокой свёрточной нейронной сети в [48]. Нейронная сеть, используемая в данной работе, принимает на вход текущий кадр НР, а также некоторое количество кадров НР (1 или 2) до и после него. На выходе данной сети получается одно изображение ВР, соответствующее текущему кадру.



Авторы [49] развивают идеи, предложенные в [48]. Главным отличием данного подхода является совместное обучение нейронных сетей, используемых для повышения разрешения и оценки оптического потока между кадрами, что повысило точность полученных результатов.

Алгоритм, представленный в работе в [50], также использует свёрточные нейронные сети, однако его авторы фокусируются на решении задачи точной регистрации исходных кадров НР и их слиянии в единое изображение, рассматривая различные подходы к этой проблеме.

Специфичной задачей, периодически возникающей при обработке видеопоследовательности, является компенсация размытия изображения, вызванного быстрым движением объектов в кадре. В частности, решению этой проблемы посвящена работа [51], авторы которой предлагают использование статистической модели на основе бинарной скрытой переменной, классифицирующей полезность каждого пикселя рассматриваемого кадра НР в зависимости от его резкости. Это позволяет уменьшить влияние размытых регионов изображения на конечный результат.

## **1.2 Особенности повышения разрешения в условиях аппликативных помех**

Как было отмечено выше, воздействие помех различного вида на входные изображения НР можно рассматривать как дополнительный фактор, локально понижающий разрешение входных изображений. В этом плане особый интерес представляют помехи аппликативного характера, воздействие которых проявляется в замещении некоторых случайных участков изображений, называемых локальными областями закрытия (ЛОЗ), некими ложными наблюдениями, имеющими случайные значения. Примеры изображений, подверженных воздействию аппликативных помех (АП) представлены на рисунке 1.3.



Рисунок 1.3 – Примеры воздействия аппликативных помех

Воздействие АП на одноканальное изображение с разрешением  $n_1 \times n_2$  можно описать с помощью следующей модели [69]:

$$y_{ij} = \lambda_{ij}(x_{ij} + v_{ij}) + (1 - \lambda_{ij})\omega_{ij}, \quad i = \overline{1, n_1}, \quad j = \overline{1, n_2},$$

где  $\mathbf{Y} = \|y_{ij}\|$  – изображение, подверженное воздействию аппликативных помех;  $\mathbf{X} = \|x_{ij}\|$  – оригинальное изображение (полезные наблюдения);  $\mathbf{V} = \|v_{ij}\|$ ,  $M[\mathbf{V}] = 0$  – аддитивный шум;  $\mathbf{\Omega} = \|\omega_{ij}\|$  – аппликативные помехи (ложные наблюдения),  $\mathbf{\Lambda} = \|\lambda_{ij}\|$ ;  $\lambda_{ij} \in \{0, 1\}$  – бинарная матрица, описывающая положение ЛОЗ на изображении ( $\lambda_{ij} = 1$ , если пиксель с координатами  $i$  и  $j$  является полезным наблюдением, и  $\lambda_{ij} = 0$ , если ложным).

В связи с тем, что АП полностью замещают полезные наблюдения в ЛОЗ случайными данными, а форма и расположение ЛОЗ носят случайный характер, фильтрация АП является существенно более сложной задачей по сравнению с устранением воздействия аддитивных помех и не может быть полностью решена методами и алгоритмами линейной фильтрации.

В [69] показано, что в основе фильтрации АП должна лежать сегментация всех пикселей изображения на два класса: полезные и ложные наблюдения, что позволяет выделить границы ЛОЗ. Данная задача может решаться различными методами [69, 55] при условии, что помехи отличаются от полезного изображения цветом, яркостью или текстурой.

Так, в [69] сегментация исходного изображения осуществляется при помощи методов контурного анализа. Изображение представляется в виде

ненаправленного графа  $G = (V, E)$ , где вершины  $V$  представляют из себя интенсивности каждого пикселя изображения, а рёбра  $E$  соединяют соседние по горизонтали и вертикали пиксели. Сегментация границ ЛОЗ в этом случае сводится к нахождению такого разреза графа  $G$ , который минимизирует целевую функцию, включающую в себя в том числе вероятности того, что два соседних пикселя  $i$  и  $j$  принадлежат разным сегментам (т.е. относятся к разным классам). Оценка данной вероятности осуществляется при помощи нейронной сети, принимающей на вход как интенсивности рассматриваемых пикселей  $G_i$  и  $G_j$ , так и интенсивности всех пикселей, лежащих в некоторой окрестности от них.

В [55] используются алгоритмы поиска минимального разреза графа  $G$  [70], требующие предварительного указания вершин, принадлежащих разным сегментам. Для решения этой задачи в автоматическом режиме осуществляется предварительная кластеризация всех пикселей изображения на классы полезных и ложных наблюдений при помощи алгоритма  $k$ -средних [71] с последующей постобработкой результатов.

В последние годы для решения задач сегментации всё большую популярность набирает использование суперпиксельной сегментации изображений [76–78], пример которой представлен на рисунке 1.4. Под суперпикселем понимается группа расположенных рядом и имеющих схожие свойства пикселей, которые рассматриваются при дальнейшей обработке как единое целое. Исходя из свойств суперпикселей, их границы будут достаточно с большой точностью совпадать с границами ЛОЗ на изображении, что делает применение такого подхода перспективным для фильтрации АП. Для завершения сегментации ЛОЗ будет требоваться разбиение полученных суперпикселей на классы полезных и ложных наблюдений, что может быть осуществлено при помощи алгоритмов кластеризации данных.



Рисунок 1.4 – Пример суперпиксельной сегментации изображений

Применение рассмотренных выше подходов к сегментации границ ЛОЗ в автоматическом режиме исходит из предположения, что некоторые свойства аномальных или полезных наблюдений являются известными (например, средняя яркость или размеры), так как эта информация требуется для разбиения на классы. Однако на практике это не всегда может быть возможным, поэтому при наличии достаточно хорошей обучающей выборки более перспективным подходом может являться использование глубоких нейронных сетей для семантической сегментации изображений [83-86].

С точки зрения дальнейшей обработки с целью построения СР для изображений, подверженных воздействию АП, возможны два случая.

1. В случае однокадрового СР восстановление потерянной информации не представляется возможной. Однако для повышения качества выходной изображения ВР можно осуществить закраску ЛОЗ при помощи соответствующих алгоритмов (например, [74, 75]). Схема такой обработки исходного изображения НР представлена на рисунке 1.5.

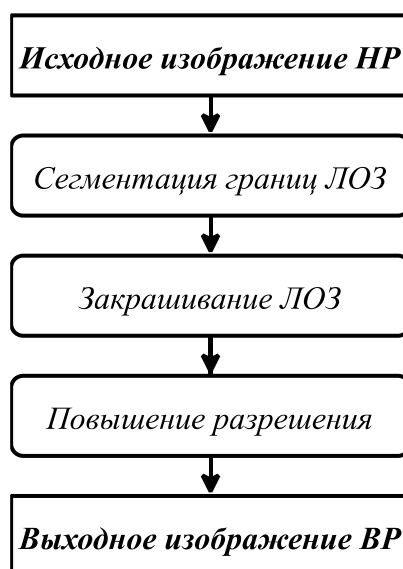


Рисунок 1.5 – Схема однокадрового СР в условиях аппликативных помех

2. В случае обработки последовательности изображений устранение воздействия АП представляется возможным за счёт избыточности входной информации, возникающей как за счёт смещений камеры, так и возможного перемещения ЛОЗ между кадрами. Это позволяет организовать обработку исходных изображений таким образом, чтобы осуществить фильтрацию АП наряду с повышением качества.

### 1.3 Методы и алгоритмы построения многокадрового сверхразрешения в условиях аппликативных помех

По сравнению с традиционными методами многокадрового и СР видеоданных, задача построения СР при наличии ложных наблюдений (АП) на входных изображениях НР рассматривалась в литературе в значительно меньшей степени [52-58].

Описанные ниже алгоритмы рассматривают изображения как векторы, полученные в ходе построчной развёртки двумерных пиксельных матриц. Под  $\{\mathbf{y}^{(t)}\}$  понимается последовательность исходных изображений НР, полученных в моменты времени  $t = \overline{1, T}$ , а под  $\mathbf{x}$  – результирующее изображение ВР, которое требуется получить в ходе построения СР.

### 1.3.1 Алгоритмы, основанные на модели спиновых стёкол

Алгоритмы, описанные в работах [52, 53] используют широко применяемую в физике модель спиновых стекол [97]. При этом она используется не для моделирования изображений, а АП, действующих на них. Данный подход не требует отдельной процедуры сегментации ЛОЗ на изображениях. Вместо этого используемая статистическая модель включает в себя скрытые векторы  $\{\mathbf{z}^{(t)}\}$ , соответствующие каждому изображению НР. Элементы каждого такого скрытого вектора соответствуют пикселю изображения НР и принимают значения  $z_i^{(t)} \in \{-1, +1\}$  в зависимости от того, содержит ли данный пиксель полезное (+1) или ложное (-1) наблюдение.

Вероятностная модель спиновых стёкол  $\mathbf{z}^{(t)}$  описывается как

$$p(\mathbf{z}^{(t)}) = \frac{1}{Z} \exp(-E(\mathbf{z}^{(t)})),$$

$$E(\mathbf{z}^{(t)}) = -J_{self} \sum_i z_i^{(t)} - J_{inner} \sum_{i \sim j} z_i^{(t)} z_j^{(t)},$$

где  $Z$  – нормировочная константа; константы  $J_{self} > 0$  и  $J_{inner} > 0$  задаются исходя из свойств АП;  $i \sim j$  означает, что  $i$ -ый и  $j$ -ый пиксели изображения располагаются по соседству по горизонтали или вертикали.

Условное распределение наблюдаемых изображений НР описывается гауссовским законом

$$p(\mathbf{y}^{(t)} | \mathbf{x}, \mathbf{z}^{(t)}) = N(\mathbf{y}^{(t)}, \mathbf{W}^{(t)}\mathbf{x}, \mathbf{B}^{-1}(\mathbf{z}^{(t)})),$$

где  $\mathbf{W}^{(t)}$  – оператор, характеризующий воздействие системы формирования наблюдаемых изображений (прореживание, размытие и т.д.);  $\mathbf{B}^{-1}(\mathbf{z}^{(t)})$  – матрица, обратная к диагональной матрице ковариации и зависящая от значений скрытого вектора  $\mathbf{z}^{(t)}$ .

Априорное распределение изображения ВР также описывается гауссовским законом

$$p(\mathbf{x}) = N(\mathbf{x}, 0, (\rho \mathbf{A})^{-1}),$$

где  $\rho$  – некоторый коэффициент;  $\mathbf{A}$  – матрица точности, обладающая сглаживающим эффектом.

Изображение ВР  $\mathbf{x}$  можно оценить как математическое ожидание апостериорного распределения  $p(\mathbf{x}|\mathbf{y})$ . Однако это невозможно сделать аналитически из-за наличия скрытых переменных  $\{\mathbf{z}^{(t)}\}$ . Поэтому для получения итогового изображения ВР используется метод вариационного байесовского вывода, в котором для упрощения вычислений апостериорное распределение по набору ненаблюдаемых переменных аппроксимируется некоторым другим распределением, называемым вариационным. В данном случае, такое вариационное распределение для изображения ВР описывается гауссовским законом  $q^*(\mathbf{x}) = N(\mathbf{x}, \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ , параметры  $\boldsymbol{\mu}_x$  и  $\boldsymbol{\Sigma}_x$  которого задаются исходя из соотношений выше.

В конечном счёте, оценка изображения ВР как математического ожидания  $\boldsymbol{\mu}_x$  распределения  $q^*(\mathbf{x})$  сводится к решению системы линейных уравнений с учётом сильной разреженности матрицы  $\boldsymbol{\Sigma}_x$ , что осуществляется путём использования метода сопряжённых градиентов (GNC) [98].

В [53] рассматривается модификация описанного в [52] подхода, которая учитывает в используемой модели инерционное движение областей, подверженных воздействию АП, между кадрами НР. Это осуществляется путём добавления в модель спиновых стёкол для  $\mathbf{z}^{(t)}$  вектора  $\boldsymbol{\theta}^{(t)} = [\theta_1^{(t)}, \theta_2^{(t)}]$ , описывающего смещение ЛОЗ между наблюдениями  $\mathbf{y}^{(t)}$  и  $\mathbf{y}^{(t+1)}$ :

$$E(\mathbf{z}^{(t)}, \mathbf{z}^{(t+1)}, \boldsymbol{\theta}^{(t)}) = -J_{self} \sum_i z_i^{(t+1)} - J_{inner} \sum_{i \sim j} z_i^{(t+1)} z_j^{(t+1)} - J_{move} \mathbf{z}^{(t+1)T} \mathbf{G}(\boldsymbol{\theta}^{(t)}) \mathbf{z}^{(t)},$$

где  $\mathbf{G}(\boldsymbol{\theta}^{(t)})$  – матрица сдвига АП; константа  $J_{move}$  описывает насколько схожи области, подверженные воздействию АП, на смежных кадрах.

Вычисления для получения итоговой оценки ВР проводится таким же образом как в [52], но с учётом дополнительных переменных  $\{\boldsymbol{\theta}^{(t)}\}$ .

### 1.3.2 Алгоритмы, основанные на модели случайных марковских полей

В [54] описывается алгоритм, использующий модель случайных марковских полей [99] для моделирования изображения ВР. Модель наблюдений описывается следующим соотношением:

$$\mathbf{y}^{(t)} = \mathbf{O}^{(t)} \mathbf{D} \mathbf{H}^{(t)} \mathbf{x} + \boldsymbol{\omega}^{(t)},$$

где оператор  $\mathbf{O}^{(t)}$  удаляет поражённые АП пиксели изображения ВР; оператор  $\mathbf{D}$  выполняет прореживание изображения ВР; оператор  $\mathbf{H}^{(t)}$  характеризует воздействие системы формирования изображений;  $\boldsymbol{\omega}^{(t)}$  – вектор аддитивного шума. При этом операторы предполагается, что операторы  $\mathbf{O}^{(t)}$  и  $\mathbf{H}^{(t)}$  либо известны заранее, либо могут быть оценены.

Процедура оценки оператора  $\mathbf{O}^{(t)}$  предполагает явную процедуру проведения сегментации пикселей каждого изображения НР  $\mathbf{y}^{(t)}$  на полезные и ложные наблюдения.

Изображение ВР  $\mathbf{x}$  рассматривается как неоднородное адаптивное случайное марковское случайное поле (модель DAMRF) [99]. Использование такой модели позволяет сохранить неоднородности и высокочастотные детали на изображении ВР.

Плотность распределения  $\mathbf{x}$  задаётся как:

$$p(\mathbf{x}) = \frac{1}{Z} \exp\left(-\sum_{c \in C} g(d_c \mathbf{x})\right),$$

где  $Z$  – нормировочная константа;  $C$  – множество всех клик;  $d_c \mathbf{x}$  – мера локальных пространственных изменений (в данном случае – разница между значениями соседних пикселей).

Используемая модель DAMRF адаптивно оценивает уровень сходства пикселей для сохранений неоднородностей:

$$g(\eta) = -\gamma \cdot \exp\left(-\frac{\eta^2}{\gamma}\right).$$



Максимум апостериорной вероятности изображения ВР  $\tilde{\mathbf{x}}$  может быть оценен при помощи алгоритма постепенной невыпуклой оптимизации (GNC) [99] на основе следующего соотношения:

$$\tilde{\mathbf{x}} = \arg \min \left( \left\| \mathbf{y}^{(t)} - \mathbf{O}^{(t)} \mathbf{D} \mathbf{H}^{(t)} \mathbf{x} \right\|^2 + \beta \sum_{c \in C} g(d_c \mathbf{x}) \right),$$

где  $\beta$  – коэффициент регуляризации.

### 1.3.3 Алгоритмы, основанные на фильтрации калмановского типа

В работах [55-58] рассматриваются алгоритмы многокадрового СР в условиях АП на основе использования оптимальной фильтрации калмановского типа. Данный подход предполагает задание модели состояний для ненаблюдаемых изображений ВР и модели наблюдений для известных изображений НР.

Модель состояний задаётся следующим образом:

$$\mathbf{x}^{(t+1)} = \mathbf{F}^{(t)} \mathbf{x}^{(t)} + \mathbf{u}^{(t)}, \quad (1.1)$$

где  $\mathbf{F}^{(t)}$  – оператор, определяющий значения межкадровых сдвигов объектов на изображениях;  $\mathbf{u}^{(t)}$  – гауссовский вектор.

С учётом специфики АП (наличия ЛОЗ на изображениях НР) модель наблюдений принимает следующий вид:

$$\mathbf{y}^{(t)} = \mathbf{A}^{(t)} (\mathbf{H}^{(t)} \mathbf{x}^{(t)} + \mathbf{v}^{(t)}) + \mathbf{B}^{(t)} (\tilde{\mathbf{y}}^{(t|t-1)} + \mathbf{w}^{(t)}), \quad (1.2)$$

где  $\mathbf{v}^{(t)}$  – гауссовский вектор аддитивного шума;  $\mathbf{H}^{(t)}$  – оператор, характеризующий воздействие системы формирования наблюдаемых изображений, приводящее к снижению разрешения;  $\mathbf{A}^{(t)}$  и  $\mathbf{B}^{(t)}$  – диагональные матрицы со случайными элементами, принимающими значение ноль или единица в случае наличия в соответствующем пикселе полезной или ложной информации;  $\tilde{\mathbf{y}}^{(t|t-1)}$  – прогноз вектора оценки наблюдаемого изображения НР;  $\mathbf{w}^{(t)}$  – вектор с нулевым средним.

На основе данных моделей может быть проведён синтез двух типов алгоритмов СР:

- оптимальный в классе линейных алгоритм фильтрации;
- оптимальный условно-линейный алгоритм фильтрации.

Оптимальный в классе линейных алгоритм фильтрации основывается на использовании стандартных соотношений для весовых матриц, которые не зависят от получаемых наблюдений. При этом было установлено [56], что такой тип алгоритма не обладает достаточной точностью для построения СР в условиях АП.

Оптимальный условно-линейный алгоритм фильтрации, напротив, использует апостериорную информацию о ложных наблюдениях на каждом из входных кадров НР, что позволяет получить большую точность на выходном изображении ВР при построении СР в условиях ложных наблюдений. Получение этой информации предполагает проведение сегментации кадров НР на полезные и ложные наблюдения.

Соотношения для получения оценки ВР в случае оптимальной условно-линейной фильтрации детально приводятся в главе 2, п. 2.2.

### **1.3.4 Блочная обработка изображений в алгоритмах многокадрового сверхразрешения**

Особенностью использования многих описанных выше алгоритмов для построения СР (в частности, алгоритмов на основе оптимальной линейной фильтрации калмановского типа) является большая размерность обрабатываемых данных в случае использования данного подхода напрямую. Так длина векторов НР может достигать порядка  $10^6 - 10^7$ . Возможным решением этой проблемы является организация блочной обработки изображений, позволяющая на порядки снизить размерности обрабатываемых векторов и матриц. Для этого все входные изображения НР делятся на множество небольших перекрывающихся друг друга блоков, каждый из которых обрабатывается независимо (т.е. каждый блок НР рассматривается как отдельное изображение, разрешение которого повышается). После завершения обработки, получившиеся блоки ВР объединяются в одно

изображение ВР на основе их центральных (неперекрывающихся фрагментов). В [57, 58] было показано, что такой подход к обработке практически не сказывается на качестве получаемых изображений, что делает его применение на практике целесообразным.

Пусть изображение разбивается на  $N_1$  и  $N_2$  квадратных блоков по вертикали и горизонтали соответственно. Внешний размер каждого блока  $s_{out} = s_{in} + 2\Delta s$ , где  $s_{in}$  – размер внутренней (неперекрывающийся) области блока,  $\Delta s$  – ширина области перекрытия блоков (рисунок 1.6).

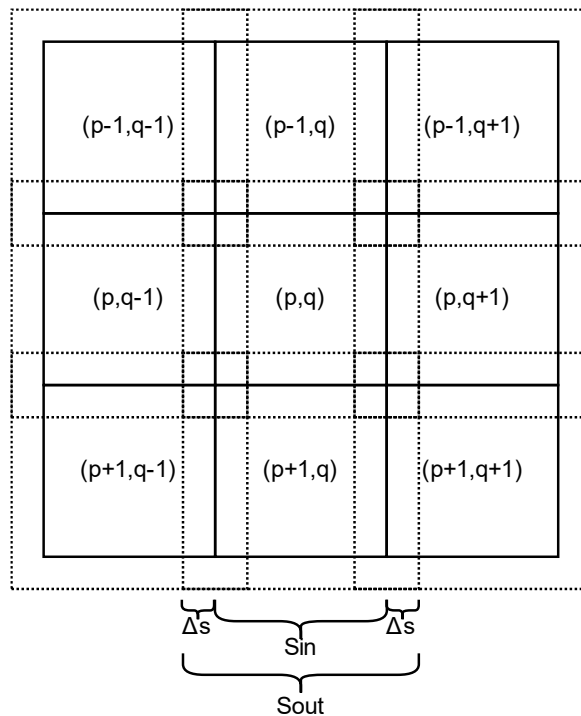


Рисунок 1.6 – Разбиение изображения на пересекающиеся блоки

В случае использования методов оптимальной линейной фильтрации и с учётом блочного представления изображений модели состояний (1.1) и наблюдений (1.2) для блоков имеют следующий вид:

$$\mathbf{x}_{t+1}^{(p,q)} = \mathbf{F}_t^{(p,q)} \mathbf{x}_t^{(p,q)} + \mathbf{u}_t^{(p,q)},$$

$$\mathbf{y}_t^{(p,q)} = \mathbf{A}_t^{(p,q)} (\mathbf{H}_t^{(p,q)} \mathbf{x}_t^{(p,q)} + \mathbf{v}_t^{(p,q)}) + \mathbf{B}_t^{(p,q)} (\tilde{\mathbf{y}}_{t|t-1}^{(p,q)} + \mathbf{w}_t^{(p,q)}),$$

где  $t = \overline{1, T}$ ,  $p = \overline{1, N_1}$ ,  $q = \overline{1, N_2}$ ;  $\mathbf{x}_t^{(p,q)}$  и  $\mathbf{y}_t^{(p,q)}$  – блоки изображений ВР и НР соответственно; все прочие векторы и матрицы имеют тот же смысл, что и раньше, но имеют соответствующие блокам размерности  $s_{out}$  для векторов и  $s_{out} \times s_{out}$  для матриц.

Блочный подход к организации условно-линейной фильтрации не только снижает размерность исходных данных, но и повышает быстродействие алгоритма. Во-первых, в ходе вычисления новой оценки ВР в процессе работы условно-линейного фильтра, коррекция осуществляется только для тех блоков, которые имеют хотя бы частично не подвержены воздействию АП. Во-вторых, независимость блоков друг от друга позволяет организовать их параллельную обработку для обеспечения большего быстродействия алгоритма.

#### **1.4 Постановка задачи и общая схема проведения исследований в интересах построения алгоритмов многокадрового сверхразрешения в условиях аппликативных помех**

С точки зрения постановки задачи СР в условиях наличия АП на входных изображениях, наибольший интерес представляет обработка последовательности кадров НР (многокадровое СР и СР видеоданных), так как появляющаяся в таком случае избыточность информации, позволяет не только получить большую точность при оценке изображений ВР, но и компенсировать воздействие АП, действующих на исходные изображения. При этом важно отметить, что синтезируемый алгоритм должен давать точный результат и в случае отсутствия АП на входных изображениях.

Анализ существующих работ в области многокадрового СР в условиях АП показывает, что данная проблема не получила на данный момент удовлетворительного решения. Такая задача рассматривалась всего в нескольких из числа выше цитируемых работах, причем в них не учитывались в достаточной степени вопрос статистической неопределенности параметров математической модели наблюдений, а также различные подходы к сегментации ЛОЗ. Кроме того,

ни один из рассмотренных методов и алгоритмов не использует широко распространённый сегодня и перспективный подход, основанный на искусственных нейронных сетях. Наконец, в существующей литературе отсутствуют сведения об экспериментальном сравнении результатов, полученных при помощи различных алгоритмов.

Таким образом, на основе проведённого анализа можно выделить два альтернативных подхода к решению данной задачи, являющихся потенциально эффективными.

Первый из них основывается на использовании алгоритмов оптимальной условно-линейной фильтрации последовательности входных изображений НР, дополненных алгоритмами суперпиксельной сегментации для выделения ЛОЗ и адаптированных к условиям статистической неопределенности параметров используемых моделей данных (моделей наблюдения и состояния).

Второй подход основан на отказе от использования математических моделей данных, заданных в явном виде, и использовании методов и алгоритмов машинного обучения и, в частности, глубоких нейронных сетей для обработки кадров НР с обучением по совокупности примеров (прецедентов) изображений на входе и на выходе системы обработки.

Методы, основанные на оптимальной линейной фильтрации последовательности изображений калмановского типа, уже применялись для построения многокадрового СР условиях АП [55-58]. По мнению автора, алгоритмы данного типа представляют наибольший интерес с точки зрения дальнейшего исследования по сравнению с альтернативными, как имеющие в многих случаях адекватную для данной задачи модель. Так, анализ особенностей синтеза алгоритмов СР в условиях наличия ложных наблюдений показывает, что устранение воздействия АП требует организации процедуры независимой сегментации входных изображений НР на полезные и ложные наблюдения с целью выделения границ ЛОЗ. При этом алгоритмы, рассмотренные в существующей литературе, используют методы, основанные на контурном анализе. Отсюда возникает необходимость в рассмотрении альтернативных

подход к сегментации ЛОЗ, основывающихся, например, на использовании суперпиксельного представления изображений. Отдельного рассмотрения заслуживает и ситуация отсутствия АП на изображении, так как их появление носит случайный характер.

Синтез алгоритмов СР на основе искусственных нейронных сетей является одним из наиболее широко используемых современных подходов к решению задачи построения СР в различных постановках задачи. Такой вывод можно сделать на основе проведённого анализа известных работ, посвящённых алгоритмам СР. В частности, в последние годы всё большее распространение получают сети, принадлежащие к классу глубоких свёрточных сетей. Нейронные сети такого типа находят своё применение в решении широкого класса задач, связанных с обработкой изображений: распознавание образов [91], семантическая сегментация [83-86], закраска пропусков [75], оценка оптического потока между кадрами [89, 90] и т.д. Рост популярности этого подхода можно связать с прогрессом в производительности вычислительных систем (например, использование графических процессов для распараллеливания вычислений при помощи технологии CUDA [92]), что в свою очередь привело к появлению новых техник обучения глубоких нейронных сетей.

Основными факторами, которые привели к возможности использования искусственных нейронных сетей для решения задачи построения СР, можно считать определенные трудности в задавании и в использовании явных математических моделей данных. Получение достаточно адекватных результатов при помощи традиционных подходов к СР, таких как методы оптимальной линейной фильтрации, требует использования сложных моделей, что в свою очередь приводит к усложнению дальнейших вычислений. Кроме того, такие алгоритмы, как правило, не обладают высоким быстродействием ввиду необходимости выполнения в них сложных матричных вычислений. Данные проблемы могут быть решены путём использования искусственных нейронных сетей на различных этапах обработки, так как они не предполагают использования моделей для исходных и результирующих данных, а также

обладают относительно высокой производительностью после завершения обучения. Кроме того, нейронные сети обладают возможностью переноса обучения, что позволяет использовать уже обученную нейронную сеть для решения схожей, но несколько отличающейся задачи.

В связи с тем, что в существующей литературе не рассматривается использование искусственных нейронных сетей для построения СР при обработке последовательности кадров НР в сочетании с воздействием АП, применение глубоких свёрточных сетей для решения данной задачи является перспективным направлением исследования. При этом представляется целесообразным синтез двух типов алгоритмов:

- использующих независимую сегментацию ЛОЗ на входных изображениях;
- не реализующих отдельный этап сегментации ЛОЗ на входных изображениях.

Также стоит отметить, что искусственные нейронные сети могут быть использованы для решения как задачи повышения разрешения изображений, так и других задач, возникающих в ходе обработки последовательности изображений (например, регистрация кадров НР и сегментация ЛОЗ на них).

В соответствии с целью данной работы и логикой проведения исследований была предложена и реализована общая схема, раскрывающая основные решаемые задачи и подзадачи и их взаимосвязь друг с другом, представленная на рисунке 1.7.

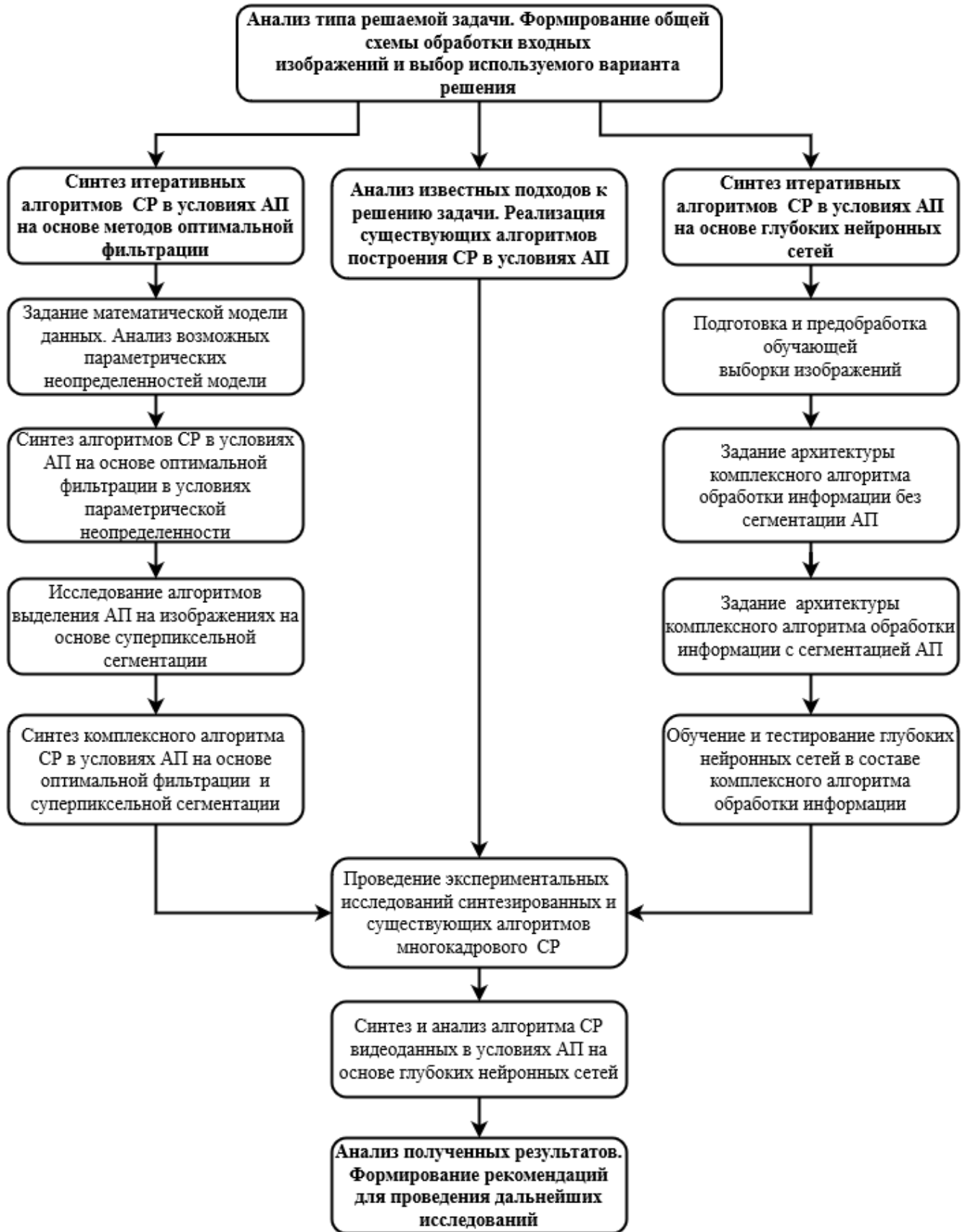


Рисунок 1.7 – Общая схема проведения исследований в интересах построения алгоритмов многокадрового сверхразрешения в условиях аппликативных помех



В рамках данной схемы исследований в первую очередь были изучены особенности организации процесса обработки входной последовательности изображений НР и предложен общий алгоритм функционирования, основанный на итеративном уточнении оценки изображения ВР (глава 2, п. 2.1). В соответствии с разделением используемой методологии на методы статистического синтеза и методы машинного обучения в схеме рисунке 1.7 выделены две ветви, реализующие два принципиально разных подхода к решению рассматриваемой задачи с точки зрения используемых методов и средств.

При реализации первого подхода в рамках данной схемы были рассмотрена математическая модель оптимальной условной фильтрации и основывающийся на ней алгоритм для построения многокадрового СР в условиях АП, использующий результаты сегментации ЛОЗ на входных изображениях (глава 2, п. 2.2). Рассмотрена задача оптимальной фильтрации последовательности изображений в условиях неопределённости относительно параметров математической модели.

Для проведения сегментации областей, подверженных воздействию АП, был предложен и исследован двухэтапный алгоритм сегментации ЛОЗ на основе суперпиксельного представления изображений. В этом алгоритме на первом этапе выполняются суперпиксельная сегментация изображения, затем проводится кластерный анализ полученных суперпикселей для разбиения их на классы полезных и ложных наблюдения (глава 2, п. 2.3). Описаны соотношения, обеспечивающие включение результатов такой сегментации в работу оптимального условно-линейного фильтра.

Затем было проведено экспериментальное сравнение рассмотренного комплексного подхода на основе оптимальной условно-линейной фильтрации и двухэтапного алгоритма сегментации ЛОЗ в традиционной и адаптивной (с учётом параметрической неопределённости используемой модели) постановке с известными альтернативными алгоритмами [52-55] (глава 2, п. 2.4).

В рамках второго подхода, основанного на использовании методологии машинного обучения, как показано на схеме рисунке 1.7, проведён синтез двух

комплексных алгоритмов многокадрового СР на основе глубоких свёрточных сетей, которые реализуют тот же общий алгоритм итеративного уточнения оценки ВР, как и рассмотренный ранее алгоритм на основе модели оптимальной линейной фильтрации (глава 3, п.п. 3.1, 3.2). Первый алгоритм на основе нейронных сетей базируется на основе архитектуры глубокой свёрточной нейронной сети из [43] и не имеет этапа сегментации ЛОЗ. Второй алгоритм использует специальную архитектуру нейронной сети для обновления текущей оценки ВР в условиях АП, организованную в виде направленного ациклического графа (DAG-сеть). Также рассматриваются вопросы регистрации изображений НР путём оценки оптического потока между кадрами и сегментации ЛОЗ при помощи глубоких свёрточных нейронных сетей, включая использование предварительно обученных моделей.

Так как использование искусственных нейронных сетей не предполагает использование специальных моделей входных данных, для проведения обучения предложенных алгоритмов и их экспериментального исследования была сформирована обучающая выборка, состоящая из кадров как подверженных воздействию АП, так и неискажённых ними (глава 3, п. 3.3). Данная выборка использовалась для обучения нейронных сетей для улучшения качества изображений, а также дообучения сети для семантической сегментации областей, имеющих аномальные наблюдения.

Кроме того, на основе подхода, показавшего наибольшую эффективность и использующего комплекс глубоких сетей, включая предложенную DAG-сеть, а также процедуру независимой сегментации АП на входных кадрах НР, были реализованы алгоритмы построения СР для обработки видеоданных (глава 4) в условиях АП. Первый синтезированный алгоритм основывается на использовании ранее рассмотренной DAG-сети и итеративном уточнении оценки, а второй – на модификации DAG-сети с учётом специфики обработки кадров видеопоследовательности.

Завершающий этап исследований в схеме рисунке 1.7 предполагает проведение масштабных компьютерных экспериментов, направленных на

сравнение всех синтезированных тем или иным способом алгоритмов между собой, а также их известных аналогов. На этой основе проводится анализ полученных результатов и формирование рекомендаций для проведения дальнейших исследований в области многокадрового СР в условиях АП.

### **Выводы по главе**

1. Проведён анализ работ, посвящённых алгоритмам и методам построения СР при наличии аддитивного шума на изображениях в различных постановках задачи (однокадровое СР, многокадровое СР и СР видеоданных). Рассмотрены различные модели, использованные в данных алгоритмах, а также возможные подходы к организации обработки исходных изображений. Анализ показывает, что перспективными направлениями исследования является синтез алгоритмов СР на основе двух альтернативных подходов на основе оптимальной линейной фильтрации и искусственных нейронных сетей, включая класс глубоких свёрточных сетей. Причём последний подход получает в последние годы всё большее распространение.

2. Рассмотрена специфика проблемы построения СР в условиях воздействия аппликативных помех на исходных изображениях НР, проявляющихся в возникновении областей ложных наблюдений. Данная задача требует проведения дополнительного этапа фильтрации данных помех, который опирается на сегментацию границ областей изображений, представляющих из себя ложные наблюдения. Проанализированы подходы к решению данной задачи. По сравнению с обработкой единственного входного изображения синтез алгоритма СР для последовательности кадров представляет больший интерес с точки зрения дальнейшего исследования, так как данный подход позволяет компенсировать воздействие АП одновременно с повышением разрешения исходных изображений.

3. Изучены существующие подходы к решению задачи повышения разрешения в условиях АП, которые представлены в литературе в существенно меньшем объёме по сравнению с традиционными алгоритмами СР. Проведён анализ используемых методов и моделей, на основе которого можно сделать заключение, что из существующих работ в данной области наибольший интерес

представляют алгоритмы, основанные на оптимальной условно-линейной фильтрации, как работающие с наиболее общими моделями данных. Потенциальное улучшение алгоритмов такого типа возможно за счёт использования альтернативных подходов к сегментации областей, подверженных воздействию действующих на входные изображения АП. В частности, перспективным направлением исследования является использование суперпиксельного представления изображений. Проведён анализ проблемы большой размерности обрабатываемых векторов и матриц в процессе оптимальной линейной фильтрации, решаемой путём организации блочной обработки изображений.

4. Проведённый анализ также показывает, что задача построения алгоритмов многокадрового СР в условиях АП не получила на данный момент удовлетворительного решения в известных источниках. Существующие работы в недостаточной степени рассматривают вопросы решения представленной задачи в условиях статистической неопределенности относительно параметров математической модели наблюдений. Также недостаточно изучены другие подходы к сегментации областей, подверженных воздействию аппликативных помех, а также использование глубоких свёрточных нейронных сетей для решения различных задач в ходе построения СР. Кроме того, в публикациях отсутствуют количественные и качественные результаты сравнения известных подходов и алгоритмов.

5. Предложена общая схема проведения исследования в интересах построения алгоритмов обработки последовательности изображений, подверженных воздействию АП с целью их фильтрации и одновременного повышения разрешения. Данная схема требует проведения синтеза и анализа комплексных алгоритмов многокадрового СР в условиях аномальных наблюдений, основывающихся на двух альтернативных подходах: оптимальной фильтрации калмановского типа с заданием соответствующих моделей данных и машинного обучения с использованием глубоких нейронных сетей и значительных по объёму обучающих данных.

## **2 Синтез и анализ алгоритмов построения сверхразрешения с использованием методов оптимальной фильтрации**

Среди известных алгоритмов многокадрового СР изображений в условиях АП можно выделить алгоритмы, использующие оптимальную линейную фильтрацию калмановского типа для повышения разрешения входных изображений путём итеративного уточнения оценки выходного изображений ВР [39-42]. Алгоритмы данного типа позволяют эффективно оценить изображение ВР по совокупности наблюдений, используя процедуру независимой сегментации АП на входных изображениях с целью повышения качества получаемых результатов. Далее, в развитие известных результатов, будет проведен синтез и анализ алгоритмов построения многокадрового СР, который использует оптимальную условно-линейную фильтрацию и двухэтапный алгоритм сегментации АП на основе суперпиксельного представления изображений.

### **2.1 Общая схема итеративной обработки последовательности изображений и ее реализация методом условно-линейной фильтрации**

В работах, основанных на использовании таких методов, как итеративная обратная проекция [36, 37], проекция на выпуклые множества [38] и фильтрация калмановского типа [39-42], используется итеративный подход к получению изображения ВР в задаче многокадрового СР. Алгоритмы, реализующие такой подход, осуществляют итеративное уточнение некой начальной оценки изображения ВР по мере обработки (поступления) новых наблюдений (кадров) НР. Общая схема такой обработки, представленная на рисунке 2.1, состоит из двух основных шагов.

1. Регистрация с предобработкой очередного кадра НР, заключающаяся в выравнивании очередного кадра НР относительно другого кадра последовательности (предыдущего или начального).

2. Обновление текущей оценки выходного изображения ВР на основе новых данных, содержащихся в очередном кадре НР.

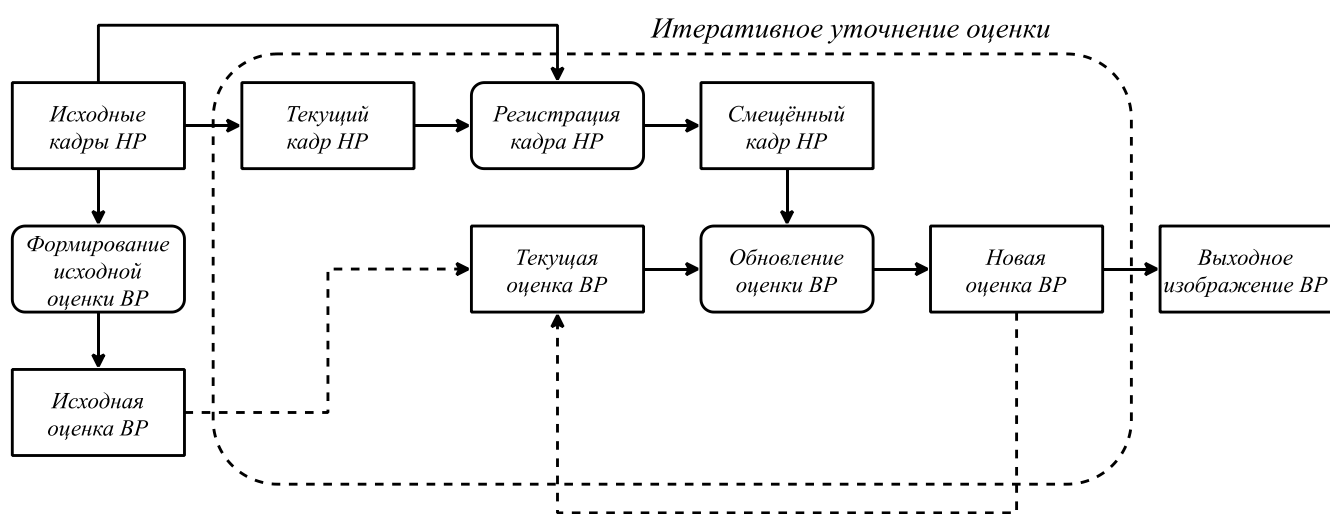


Рисунок 2.1 – Общая схема итеративного подхода к построению СР

К достоинствам данного подхода можно отнести следующее:

- поступающая последовательность кадров НР может иметь произвольную длину;
- каждое входное изображение НР обрабатывается один раз;
- возможность уточнения результата при поступлении новых наблюдений, неизвестных на момент начала обработки.

В рамках задачи многокадрового СР процедура регистрации с предобработкой изображений, заключающейся в оценке смещений и поворотов между двумя кадрами НР, позволяет повысить точность итоговой оценки ВР.

Как было отмечено в главе 1 (п. 1.2), использование дополнительной информации при обработке последовательности кадров НР позволяет устранить воздействие АП, потенциально действующих на эти изображения, одновременно с повышением детализации (разрешения) изображений. В связи с этим общая схема итеративной обработки изображений (рисунке 2.1) в [55] была расширена дополнительным этапом независимой сегментации каждого исходного кадра НР с целью выделения областей, подверженных воздействию АП. Данная информация используется затем на этапе обновления оценки ВР с целью повышения его

точности. Общая схема такой обработки поступающей последовательности изображений приводится на рисунке 2.2.

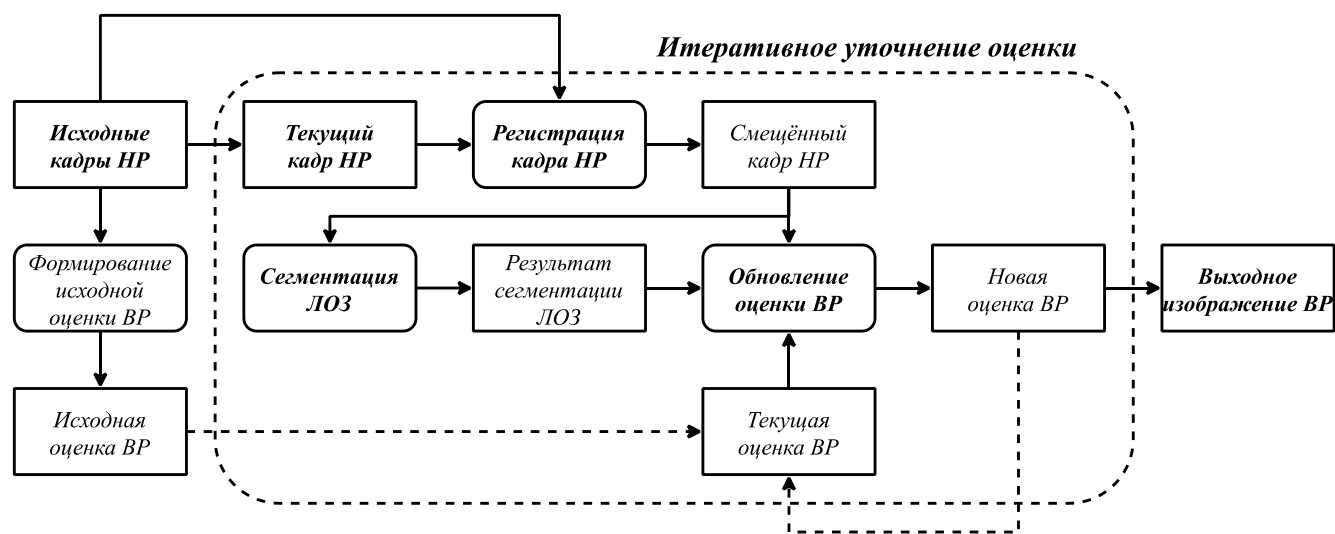


Рисунок 2.2 – Общая схема итеративного уточнения оценки изображения ВР при сегментации ЛОЗ на входных изображениях НР

Схема обработки, представленная на рисунке 2.2, реализуется следующим алгоритмом обработки поступающей последовательности кадров НР.

1. На основе входных кадров НР происходит формирование начальной оценки изображения ВР.

2. Осуществляется регистрация очередного обрабатываемого кадра НР, заключающаяся в оценке его смещения относительно первого кадра с последующим его наложением на первый кадр.

3. Осуществляется сегментация областей, подверженных воздействию ложных наблюдений, на очередном кадре НР.

4. На основе текущей оценки изображения ВР, очередного зарегистрированного кадра НР и результатов его сегментации вычисляется новая оценка изображения ВР.

5. Шаги 2-4 повторяются, пока не будет обработана вся последовательность изображений НР.

6. Последняя полученная оценка ВР принимается за результат восстановления изображения ВР.

Далее в данной главе рассматриваются и исследуются различные варианты алгоритма многокадрового СР в условиях воздействия АП, реализованные в соответствии с общей схемой на рисунке 2.2.

Первым такой алгоритм основан на методе оптимальной условно-линейной фильтрации. Термин условно-линейная фильтрация означает [68], что используемый для обновления оценки алгоритм реализует принцип оптимальной линейной фильтрации относительно вещественных компонент используемых наблюдений. При этом весовые коэффициенты используемого фильтра на каждом шаге зависят от дискретной составляющей наблюдений, отражающих результаты выявления аномальных наблюдений (в нашем случае ЛОЗ АП).

Кроме того, отдельное внимание уделяется решению рассматриваемой задачи в адаптивной постановке для учета параметрической неопределённости используемых при синтезе математических моделей. Получаемые при этом алгоритмы обработки являются нелинейными и их применение существенно усложняет общую схему обработки.

Для решения задачи сегментации АП входных изображений НР предлагается алгоритм двухэтапной сегментации. Он базируется на выполнении сначала суперпиксельной сегментации для выделения областей, поражённых АП, на каждом изображении последовательности и использовании на втором этапе статистического алгоритма кластерного анализа (ЕМ-алгоритма) для суперпикселей. Результаты сегментации, полученные при помощи данного алгоритма, включаются в общий процесс обработки и используются при уточнении оценки изображения ВР на основе оптимальной условно-линейной или адаптивной фильтрации.



## 2.2 Синтез алгоритма многокадрового сверхразрешения в условиях аппликативных помех на основе методов оптимальной фильтрации

### 2.2.1 Модели состояний и наблюдений

Как было отмечено в главе 1, алгоритмы оптимальной линейной фильтрации калмановского типа требуют задания модели состояний, описывающей переход между ненаблюдаемыми оцениваемыми состояниями системы, и модели наблюдения, описывающей зависимость регистрируемых наблюдений от состояний системы. Ниже приводятся данные модели для задачи построения многокадрового СР с учётом наличия ложных наблюдений на входных полутонных изображениях НР. При этом изображения рассматриваются как векторы, полученные в ходе построчной развёртки матрицы пикселей, а их элементы принимают значения в диапазоне  $[0,1]$ . В случае работы с цветными изображениями каждый из каналов обрабатывается отдельно.

Модель состояний, описывающая изображений ВР разрешения  $L_1 \times L_2$ , задаётся следующим образом:

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{u}_k, \quad (2.1)$$

где  $k = \overline{1, K}$ ;  $K$  – количество кадров в последовательности НР;  $\mathbf{x}_{k+1}$  и  $\mathbf{x}_k$  –  $L$ -мерные векторы изображений ВР в моменты наблюдений  $k+1$  и  $k$  соответственно,  $L = L_1 L_2$ ;  $\mathbf{F}_k$  – оператор размера  $L \times L$ , описывающий межкадровые сдвиги объектов на изображениях ВР;  $\mathbf{u}_k$  –  $L$ -мерный гауссовский случайный вектор с нулевым математическим ожиданием матрицей ковариаций  $\mathbf{Q}_k$ , описывающий аддитивный шум. В случае наличия отсутствия движения на исходном изображении ВР можно положить, что  $\mathbf{F}_k = \mathbf{I}$ ,  $k = \overline{1, K}$ .

Исходное состояние системы задаётся исходя из предположения, что оно подчиняется гауссовскому закону распределения  $\mathbf{x}_1 \sim N(\mathbf{x}_1, \tilde{\mathbf{x}}_1, \mathbf{P}_{10})$ , где  $\tilde{\mathbf{x}}_1$  – исходная оценка изображения ВР;  $\mathbf{P}_{10}$  – исходная матрица ковариации ошибки.

Каждому неизвестному изображению ВР  $\mathbf{x}_k$  соответствует наблюдаемое изображение НР  $\mathbf{y}_k$  разрешения  $M_1 \times M_2$ , причём  $M = M_1 M_2 < L$ . Модель системы формирования изображений, преобразующая исходные изображения ВР  $\mathbf{x}_k$  в наблюдения НР  $\mathbf{y}_k$ , должна учитывать следующее:

- сдвиги и повороты камеры или объектов относительно сцены;
- прореживание, приводящее к понижению разрешения наблюдаемых изображений;
- размытие и прочие эффекты фотоприёмников;
- воздействие аддитивных помех;
- воздействие аппликативных помех, проявляющихся в замещении отдельных областей изображений случайными ложными наблюдениями.

Модель наблюдений, учитывающая перечисленные выше факторы, задаётся следующим образом [55].

$$\mathbf{y}_k = \mathbf{A}_k (\mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k) + \mathbf{B}_k (\tilde{\mathbf{y}}_{k|k-1} + \mathbf{w}_k), \quad (2.2)$$

$$\mathbf{A}_k = \begin{pmatrix} a_{k1} & 0 & \cdots & 0 \\ 0 & a_{k2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & a_{kM^2} \end{pmatrix}, \quad \mathbf{B}_k = \begin{pmatrix} b_{k1} & 0 & \cdots & 0 \\ 0 & b_{k2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & b_{kM^2} \end{pmatrix}, \quad \mathbf{A}_k + \mathbf{B}_k = \mathbf{I},$$

где  $k = \overline{1, K}$ ;  $\mathbf{H}_k$  – оператор, характеризующий воздействие системы формирования наблюдаемых изображений с учётом перечисленных выше факторов (смещение, прореживание и размытие);  $\mathbf{v}_k$  – гауссовский вектор аддитивного шума с нулевым математическим ожиданием и матрицей ковариаций  $\mathbf{R}_k$ ;  $\mathbf{A}_k$  и  $\mathbf{B}_k$  – диагональные матрицы со случайными элементами, принимающими значение ноль или единица в случае наличия в соответствующем  $m$ -ом ( $m = \overline{1, M}$ ) пикселе изображения НР полезной ( $a_{km} = 1, b_{km} = 0$ ) или ложной ( $a_{km} = 0, b_{km} = 1$ ) информации;  $\tilde{\mathbf{y}}_{k|k-1}$  – прогноз вектора оценки наблюдаемого изображения НР, полученный на основе априорных сведений об изображениях с учётом обработки  $k-1$  кадра;  $\mathbf{w}_k$  – случайный вектор с нулевым математическим ожиданием и матрицей ковариаций  $\mathbf{S}_k$ , описывающий

отклонение возникающих ложных наблюдений на очередном кадре  $\mathbf{y}_k$  относительно вектора  $\tilde{\mathbf{y}}_{k|k-1}$ ;  $\mathbf{I}$  – единичная матрица.

Следует также отметить, что в рассмотренной модели наблюдений может присутствовать параметрическая неопределенность. В частности, могут быть неизвестными следующие параметры: значения параметров межкадровых сдвигов, параметр дисперсии размытия, вызванного функцией рассеяния фотоприёмников и т.д.

### 2.2.2 Обновление оценки изображения высокого разрешения при обработке очередного кадра низкого разрешения при помощи условно-линейного фильтра

Рассмотрим сначала ситуацию, когда параметры модели наблюдений известны. Как было показано в [55], наибольшую точность при использовании моделей (2.1) и (2.2) в этих условиях можно получить, синтезируя на их основе оптимальный условно-линейный фильтр, который используют апостериорную информацию в виде дискретной компоненты наблюдений относительно локализации аномальных участков на обрабатываемых кадрах НР. Получение этой апостериорной информации, как следует из общей схемы обработки, представленной на рисунке 2.2, предполагает проведение сегментации пикселей изображения на два класса (полезные и ложные наблюдения) с целью локализации ЛОЗ.

Соотношения для обновления текущей оценки ВР  $\tilde{\mathbf{x}}_{k|k}$  при обработке очередного кадра НР  $\mathbf{y}_k$  принимают следующий вид [55]:

$$\begin{aligned}\tilde{\mathbf{x}}_{k+1|k} &= \mathbf{F}_k \mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{W}_k(\boldsymbol{\theta}^k)(\mathbf{y}_k - \mathbf{y}_{k|k-1}) \\ \tilde{\mathbf{y}}_{k|k-1} &= \mathbf{H}_k \mathbf{x}_{k|k-1}, \quad \mathbf{F}_k = \mathbf{I}, \quad k = \overline{1, K}, \\ \mathbf{W}_k(\boldsymbol{\theta}^k) &= \mathbf{V}_{k\theta} \mathbf{U}_{k\theta}^{-1}, \quad \mathbf{V}_{k\theta} = \mathbf{P}_{k|k-1}(\boldsymbol{\theta}^{k-1}) \mathbf{H}_k^T \mathbf{P}_{Ak}^T(\boldsymbol{\theta}^k), \\ \mathbf{U}_{k\theta} &= \mathbf{P}_{AAk\theta} \circ [\mathbf{H}_k \mathbf{P}_{k|k-1}(\boldsymbol{\theta}^{k-1}) \mathbf{H}_k^T + \mathbf{R}_k] + \mathbf{P}_{BBk\theta} \circ \mathbf{S}_k,\end{aligned}\tag{2.3}$$

$$\mathbf{P}_{k+1|k}(\boldsymbol{\theta}^k) = \mathbf{F}_k \mathbf{P}_{k|k}(\boldsymbol{\theta}^k) \mathbf{F}_k^T + \mathbf{Q}_k = \mathbf{P}_{k|k-1}(\boldsymbol{\theta}^{k-1}) - \mathbf{W}_k(\boldsymbol{\theta}^k) \mathbf{U}_{k\theta} \mathbf{W}_k^T(\boldsymbol{\theta}^k) + \mathbf{Q}_k,$$

где  $\tilde{\mathbf{x}}_{k+1|k}$  – прогноз оценки ВР для следующего кадра;  $\mathbf{W}_k$  – весовая матрица, характеризующая, как сильно изменится оценка ВР;  $\mathbf{P}_{k|k}$  – матрица ковариаций оценки;  $\mathbf{P}_{k|k-1}$  – экстраполяция матрицы ковариаций оценки на следующий кадр;  $\boldsymbol{\theta}^k = \{\boldsymbol{\theta}_k\}$ ,  $k = \overline{1, K}$  – последовательность бинарных векторов, получаемых в результате построчной развёрстки результатов независимой сегментации кадра НР  $\mathbf{y}_k$  на полезные ( $\theta_{kl} = 1$ ) и ложные ( $\theta_{kl} = 0$ ) наблюдения;  $\mathbf{P}_{Ak}$  и  $\mathbf{P}_{Bk}$  – диагональные матрицы, содержащие вероятности того, являются ли пиксели кадра  $\mathbf{y}_k$  полезными ( $p_{akl} = p(a_{kl} = 1) = p(b_{kl} = 0)$ ,  $l = \overline{1, M}$ ) или ложными ( $p_{bkl} = p(a_{kl} = 0) = p(b_{kl} = 1)$ ,  $l = \overline{1, M}$ ) наблюдениями (отсюда  $\mathbf{P}_{Ak} + \mathbf{P}_{Bk} = \mathbf{I}$ );  $\mathbf{P}_{AAk}$  и  $\mathbf{P}_{BBk}$  – матрицы попарных вероятностей совместного появления полезных ( $p_{aklm} = p(a_{kl} = 1, a_{km} = 1)$ ) и ложных ( $p_{bklm} = p(b_{kl} = 1, b_{km} = 1)$ ) наблюдений для каждой пары пикселей кадра  $\mathbf{y}_k$ ;  $\mathbf{A} \circ \mathbf{B}$  обозначает поэлементное умножение операторов  $\mathbf{A}$  и  $\mathbf{B}$ ;  $l = \overline{1, M}$ ,  $m = \overline{1, M}$ .

Можно отметить, что соотношения (2.3) реализуют обновление оценки изображения ВР с «пульсирующим» характером изменения элементов весовой матрицы  $\mathbf{W}_k(\boldsymbol{\theta}^k)$ , значения которых резко увеличиваются или уменьшаются в зависимости от вероятностей того, являются ли соответствующие пиксели изображения полезными или ложными наблюдениями. Данные вероятности определяются как диагональные матрицы элементов  $\mathbf{P}_{Ak}(\boldsymbol{\theta}^k)$  и  $\mathbf{P}_{Bk}(\boldsymbol{\theta}^k)$ , которые зависят от результатов сегментации ЛОЗ на поступающих изображениях.

Пусть  $Y_k = \{y_{kl}\}$ ,  $l = \overline{1, M}$  – множество всех пикселей кадра  $\mathbf{y}_k$ . Для получения вектора  $\boldsymbol{\theta}_k$  необходимо разбить  $Y_k$  на два непересекающихся подмножества  $O_k$  – пиксели, представляющие полезные наблюдения, и  $H_k$  – пиксели, представляющий ложные наблюдения (АП). Таким образом, определяются элементы вектора  $\boldsymbol{\theta}_k$ :

$$\theta_{kl} = \begin{cases} 1, & y_{kl} \in O_k, \\ 0, & y_{kl} \in H_k, \end{cases} \quad k = \overline{1, K}, \quad l = \overline{1, M}.$$

Полученная в результате сегментации информация об АП используется для расчёта априорных вероятностей наличия  $p_{akl}$  или отсутствия  $p_{bkl}$  полезной информации в пикселях входных изображений:

$$p_{akl} = \frac{|O_k|}{|Y_k|}, \quad p_{bkl} = \frac{|H_k|}{|Y_k|}, \quad k = \overline{1, K}, \quad l = \overline{1, M}.$$

Соответствующие апостериорные вероятности оцениваются на основе известных или предполагаемых распределений полезных и ложных наблюдений согласно следующим соотношениям:

$$p_{bkl}(\theta_k) = \begin{cases} \frac{p_{akl} \cdot p(\theta_{kl} = 1 | a_{kl})}{p_{akl} \cdot p(\theta_{kl} = 1 | a_{kl}) + p_{bkl} \cdot p(\theta_{kl} = 1 | b_{kl})}, & \theta_{kl} = 1, \\ \frac{p_{akl} \cdot p(\theta_{kl} = 0 | a_{kl})}{p_{akl} \cdot p(\theta_{kl} = 0 | a_{kl}) + p_{bkl} \cdot p(\theta_{kl} = 0 | b_{kl})}, & \theta_{kl} = 0, \end{cases}$$

$$p_{bkl}(\theta_k) = \begin{cases} \frac{p_{bkl} \cdot p(\theta_{kl} = 1 | b_{kl})}{p_{akl} \cdot p(\theta_{kl} = 1 | a_{kl}) + p_{bkl} \cdot p(\theta_{kl} = 1 | b_{kl})}, & \theta_{kl} = 1, \\ \frac{p_{bkl} \cdot p(\theta_{kl} = 0 | b_{kl})}{p_{akl} \cdot p(\theta_{kl} = 0 | a_{kl}) + p_{bkl} \cdot p(\theta_{kl} = 0 | b_{kl})}, & \theta_{kl} = 0, \end{cases}$$

где вероятности  $p(\theta_{kl} = 1 | a_{kl})$ ,  $p(\theta_{kl} = 0 | a_{kl})$ ,  $p(\theta_{kl} = 1 | b_{kl})$  и  $p(\theta_{kl} = 0 | b_{kl})$  вычисляются на основе параметров, соответствующих классам полезных и ложных наблюдений. Эти параметры могут быть вычислены на основе результатов сегментации и зависят от моделей, используемых для её проведения.

Для выполнения данной процедуры автором предлагается использование двухэтапного алгоритма сегментации [78], основанного на суперпиксельном представлении изображений, который рассматривается далее в разделе 2.2.

### 2.2.3 Задача оптимальной фильтрации в адаптивной постановке

Как уже отмечалось ранее, модели наблюдений могут иметь неизвестные параметры, которые подлежат оценке в процессе обработки наблюдений, что

приводит к необходимости решения задачи в адаптивной постановке. В рамках такого адаптивного подхода в модели (2.3) вводится зависимость оператора связи  $\mathbf{H}_k$  относительно случайного вектора  $\delta$ , компоненты которого являются параметрами наблюдения очередного кадра. Так, данный вектор может содержать значения параметров межкадровых сдвигов, параметр дисперсии размытия, вызванного функцией рассеяния фотоприёмников, отдельные параметры алгоритма сегментации ЛОЗ на очередном кадре НР и т.д.

Рассмотрим в качестве примера оценку параметров межкадровых сдвигов. В этом случае вектор  $\delta$  будет содержать параметры аффинного преобразования, которые аппроксимируют межкадровые сдвиги в ВР:  $\delta = (\Delta x, \Delta y, \Delta \theta)^T$ , где  $\Delta x$  и  $\Delta y$  – значения сдвигов по горизонтали и вертикали;  $\Delta \theta$  – угол поворота. При этом области возможных значений данных параметров являются известными и задаются на дискретной решётке отсчётов, состоящей из  $N_\delta$  узлов. Для каждого узла задаётся априорная вероятность его реализации  $p_{\delta_i}$ ,  $i = \overline{1, N_\delta}$ .

При обработке очередного кадра НР  $\mathbf{y}_k$  для каждого возможного значения вектора неизвестных параметров  $\delta = \mathbf{d}_i^{(k)}$ ,  $i = \overline{1, N_\delta}$  вычисляются соответствующие оценки изображения ВР  $\tilde{\mathbf{x}}_{k|k}(\mathbf{d}_i^{(k)})$  на основе соотношений (2.3). При этом используется условный оператор  $\mathbf{H}_k = \mathbf{H}_k(\mathbf{d}_i^{(k)})$ , зависящий от конкретной реализации вектора  $\delta$ .

При поступлении очередного кадра для каждого возможного значения вектора неизвестных параметров  $\delta = \mathbf{d}_i^{(k)}$ ,  $i = \overline{1, N_\delta}$ , вычисляются условные оценки изображения ВР  $\tilde{\mathbf{x}}_{k|k}(\mathbf{d}_i^{(k)})$  на основе соотношений (2.1)-(2.3), где используется условный оператор  $\mathbf{H}_k = \mathbf{H}_k(\mathbf{d}_i^{(k)})$ , заданный на базе  $i$ -го возможного значения вектора  $\mathbf{d}_i^{(k)}$ . С помощью полученных условных оценок  $\tilde{\mathbf{x}}_{k|k}(\mathbf{d}_i^{(k)})$  в гауссовском приближении определяются апостериорные вероятности  $P(\mathbf{d}_i^{(k)} | \mathbf{y}_k)$  [57]:

$$P(\mathbf{d}_i^{(k)} | \mathbf{y}_k) = \frac{P(\mathbf{y}_k | \mathbf{d}_i^{(k)}) p_{\delta i}}{\sum_{i=1}^{N_\delta} (P(\mathbf{y}_k | \mathbf{d}_i^{(k)}) p_{\delta i})},$$

$$P(\mathbf{y}_k | \mathbf{d}_i^{(k)}) = \left( (2\pi)^{M^2} |\mathbf{U}_k(\mathbf{d}_i^{(k)})| \right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} e(\mathbf{d}_i^{(k)})\right), \quad (2.4)$$

$$e(\mathbf{d}_i^{(k)}) = \left( \mathbf{y}_{k|k-1} - \mathbf{H}_k(\mathbf{d}_i^{(k)}) \tilde{\mathbf{x}}_{k|k-1} \right)^T \mathbf{U}_k^{-1}(\mathbf{d}_i^{(k)}) \left( \mathbf{y}_{k|k-1} - \mathbf{H}_k(\mathbf{d}_i^{(k)}) \tilde{\mathbf{x}}_{k|k-1} \right).$$

Безусловная относительно вектора  $\delta$  оценка  $\tilde{\mathbf{x}}_{k|k}$  и соответствующая ей ковариационная матрица  $\mathbf{P}_{k|k}$  определяются как взвешенные суммы условных оценок:

$$\begin{aligned} \tilde{\mathbf{x}}_{k+1|k} &= \tilde{\mathbf{x}}_{k|k} = \sum_{i=1}^{N_\delta} \tilde{\mathbf{x}}_{k|k}(\mathbf{d}_i^{(k)}) P(\mathbf{d}_i^{(k)} | \mathbf{y}_k), \\ \mathbf{P}_{k|k} &= \sum_{i=1}^{N_\delta} \mathbf{P}_{k|k}(\mathbf{d}_i^{(k)}) P(\mathbf{d}_i^{(k)} | \mathbf{y}_k) + \sum_{i=1}^{N_\delta} \mathbf{C}_k P(\mathbf{d}_i^{(k)} | \mathbf{y}_k), \\ \mathbf{P}_{k+1|k} &= \mathbf{P}_{k|k} + \mathbf{Q}_k, \quad \mathbf{C}_k = \left( \tilde{\mathbf{x}}_{k|k}(\mathbf{d}_i^{(k)}) - \tilde{\mathbf{x}}_{k|k} \right) \left( \tilde{\mathbf{x}}_{k|k}(\mathbf{d}_i^{(k)}) - \tilde{\mathbf{x}}_{k|k} \right)^T. \end{aligned} \quad (2.5)$$

Функционал правдоподобия  $P(\mathbf{y}_k | \mathbf{d}_i^{(k)})$  в (2.4) рассчитывается с учетом ошибки  $e(\mathbf{d}_i^{(k)})$  между кадром  $\mathbf{y}_k$  и проекциями  $\tilde{\mathbf{x}}_{k|k-1}$  на сетку НР при конкретном значении вектора параметров сдвигов  $\delta = \mathbf{d}_i^{(k)}$ . Поэтому по мере накопления информации те апостериорные вероятности  $P(\mathbf{d}_i^{(k)} | \mathbf{y}_k)$ , которые близки к истинному значению неизвестных параметров  $\delta$ , увеличиваются. В результате, безусловная оценка  $\tilde{\mathbf{x}}_{k|k}$  начинает использовать с большим весом те результаты условной фильтрации, которые отражают истинное значение параметра  $\delta$ .

Как было отмечено в главе 1, п. 1.3.4, использование методов оптимальной фильтрации требует организации блочной обработки изображений с целью понижения размерности обрабатываемых векторов и матриц. Такой подход

обеспечивает дополнительное ускорение вычислений в случае адаптивной постановки задачи. Так, для вычисления апостериорных плотностей адаптивного фильтра  $P(\mathbf{d}_i^{(k)} | \mathbf{y}_k)$  можно использовать лишь несколько ( $N_b$ ) наиболее информативных (наиболее детализированных) и неподверженных воздействию АП блоков очередного кадра. Выбор таких блоков осуществляется на основе результатов сегментации ЛОЗ и, например, по максимуму средней магнитуды градиента пикселей, принадлежащих блоку. При этом значения ошибки  $e(\mathbf{d}_i^{(k)})$  для разных блоков изображения полагаются независимыми [57].

Кроме того, вместо взвешенной суммы условных оценок в качестве  $\tilde{\mathbf{x}}_{k|k}^{pq}$  может быть выбрана условная оценка, которая максимизирует апостериорную вероятность  $P(\mathbf{d}_i^{(k)} | \mathbf{y}_k)$ . Такой подход к адаптивной фильтрации может уступать классической схеме в точности, но при этом обладает существенно лучшим быстродействием за счет применения направленного перебора возможных значений вектора неизвестных параметров  $\delta$ .

С учётом блочной обработки и вышеперечисленных соображений, выражения (2.4) и (2.5) для адаптивной фильтрации принимают следующий вид [65]:

$$P(\mathbf{y}_k | \mathbf{d}_i^{(k)}) = \left( (2\pi)^{N_b s_{LE}^2} \prod_{r=1}^{N_b} | \mathbf{U}_k^r(\mathbf{d}_i^{(k)}) | \right)^{-\frac{1}{2}} \exp \left( -\frac{e(\mathbf{d}_i^{(k)})}{2} \right),$$

$$e(\mathbf{d}_i^{(k)}) = \sum_{r=1}^{N_b} \left( \mathbf{y}_k^r - \mathbf{H}_k^r(\mathbf{d}_i^{(k)}) \tilde{\mathbf{x}}_{k|k-1}^r \right)^T \mathbf{U}_k^{(r)-1}(\mathbf{d}_i^{(k)}) \left( \mathbf{y}_k^r - \mathbf{H}_k^r(\mathbf{d}_i^{(k)}) \tilde{\mathbf{x}}_{k|k-1}^r \right), \quad (2.6)$$

$$\tilde{\mathbf{x}}_{k|k}^{pq} = \tilde{\mathbf{x}}_{k|k}^{pq}(\tilde{\mathbf{d}}_i^{(k)}), \quad \mathbf{P}_{k|k}^{pq} = \mathbf{P}_{k|k}^{pq}(\tilde{\mathbf{d}}_i^{(k)}), \quad \tilde{\mathbf{d}}_i^{(k)} = \arg \max_i P(\mathbf{d}_i^{(k)} | \mathbf{y}_k),$$

где  $r = \overline{1, N_b}$ ,  $N_b$  – количество используемых информативных блоков изображения НР;  $p = \overline{1, N_1}$ ,  $N_1$  – общее количество блоков по вертикали;  $q = \overline{1, N_2}$ ,  $N_2$  – общее количество блоков по горизонтали;  $s_{LE}$  – внешний размер блока НР (см. глава 1, п. 1.3.4).



Следует отметить, что сегментация ЛОЗ при использовании блочного подхода к обработке проводится для всего изображения целиком, так как отдельные блоки изображения не несут достаточно информации для осуществления этого процесса с достаточной точностью.

### **2.3 Двухэтапная сегментация областей, подверженных воздействию аппликативных помех**

Особенностью алгоритмов, основывающихся на оптимальной линейной и условно-линейной фильтрации, является необходимость в выполнении сложных матричных вычислений, что приводит, в том числе, к большим временным затратам на обработку. Для ускорения работы алгоритмы этапы регистрации кадров НР и сегментации ЛОЗ на них проводятся независимо от этапа обновления оценки изображения ВР, используя лишь информацию об изображениях НР.

Предлагаемый подход к проведению сегментации ЛОЗ и, как следствие, разбиению всех пикселей изображения на полезные и ложные наблюдения, состоит из двух основных этапов (рисунок 2.3):

- суперпиксельная сегментация изображения;
- кластерный анализ полученных суперпикселей.

После завершения основных этапов производится постобработка результатов кластеризации с целью определения кластеров, соответствующих ложным наблюдениям.

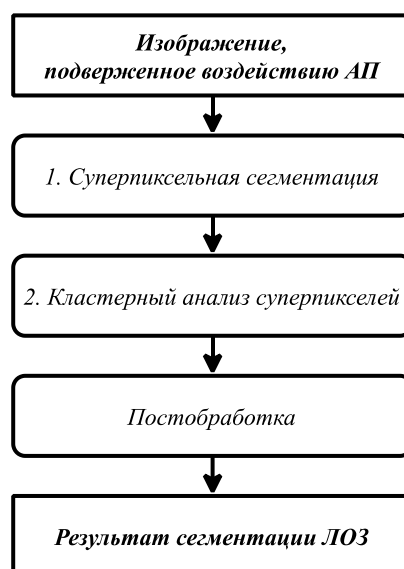


Рисунок 2.3 – Схема двухэтапной сегментации ложных наблюдений на изображениях НР

### 2.3.1 Суперпиксельное представление изображений

На первом этапе обработки кадра НР проводится его суперпиксельная сегментация (СС). Задача СС сводится к разбиению всех пикселей изображения на множество однородных компактных регионов, называемых суперпикселями. В отличие от алгоритмов традиционной сегментации, пиксели, принадлежащие смежным суперпикселям, могут иметь схожие свойства, т.е. СС изображений является избыточной. При дальнейшей обработке изображения каждый суперпиксель рассматривается как единое целое. Если положить, что  $N_P$  – количество пикселей на изображении,  $N_S$  – количество полученных суперпикселей,  $N_O$  – количество реальных объектов на изображении, то выполняется следующее соотношение:  $N_P \gg N_S \gg N_O$ .

Можно выделить следующие преимущества суперпиксельного представления изображений по сравнению с традиционной матрицей пикселей:

- вычислительная эффективность: количество суперпикселей на несколько порядков меньше количества пикселей;

- учёт связей между отдалёнными пикселями, входящими в один суперпиксель;
- большее смысловое значение суперпикселей по сравнению с отдельными пикселями;
- лучшее соответствие границам объектов на изображении по сравнению со скользящим окном прямоугольной формы.

На рисунке 2.5 представлены примеры СС, полученной как для неискажённого изображения, так и для того же изображения, но подверженного воздействию АП. В соответствии с постановкой задачи СС, в случае наличия ложных наблюдений на изображениях, границы ЛОЗ будут совпадать с границами некоторых суперпикселей. Таким образом, задача разбиения пикселей изображения на классы полезных и аномальных наблюдений сводится к кластеризации полученных суперпикселей на данные классы.

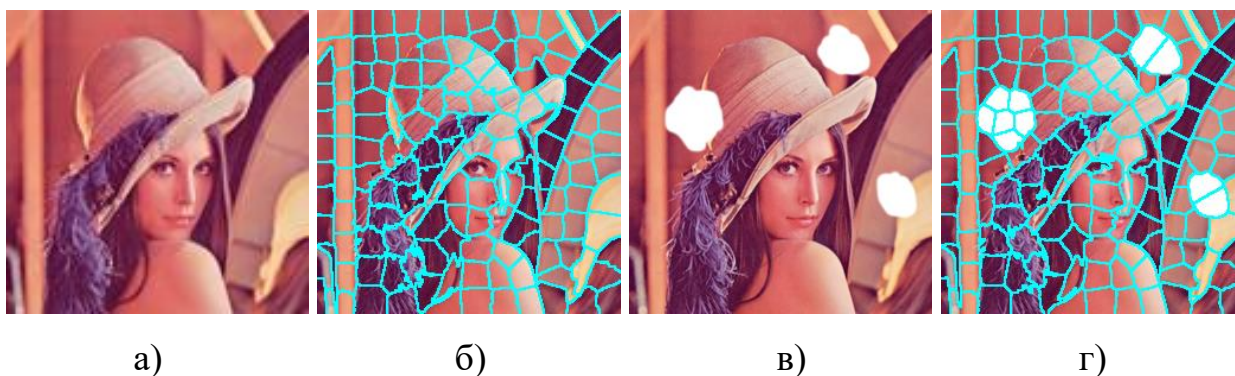


Рисунок 2.4 – Суперпиксельное представление изображений: а) оригинальное изображение; б) СС оригинального изображения; в) изображение подверженное воздействию АП; г) СС изображения, подверженного воздействию АП

В настоящее время существует большое количество алгоритмов СС [76-82], отличающихся по точности, быстродействию, компактности получаемых суперпикселей и устойчивости к шуму и аффинным преобразованиям [78]. В рамках данного исследования использовался алгоритм простой линейной итеративной кластеризации (Simple Linear Iterative Clustering, SLIC) [77, 78],

обладающий хорошим соотношением между быстродействием и точностью, а также дающий на выходе компактные суперпиксели.

**Алгоритм суперпиксельной сегментации SLIC.** Алгоритм SLIC представляет из себя пространственно-ориентированную модификацию алгоритма кластеризации k-средних [71]. Основное отличие от k-средних заключается в ограничении области поиска: пиксели для каждого сегмента ищутся не во всём изображении, а в небольшой области, пропорциональной среднему размеру сегмента  $S^2$  (рисунок 2.5),  $S = \sqrt{N_p / N_s}$ , где  $N_p$  – количество пикселей в изображении,  $N_s$  – заданное желаемое количество суперпикселей. Данное изменение повышает скорость работы алгоритма, а также уменьшает вероятность разбиения кластера на несколько разъединённых компонентов.

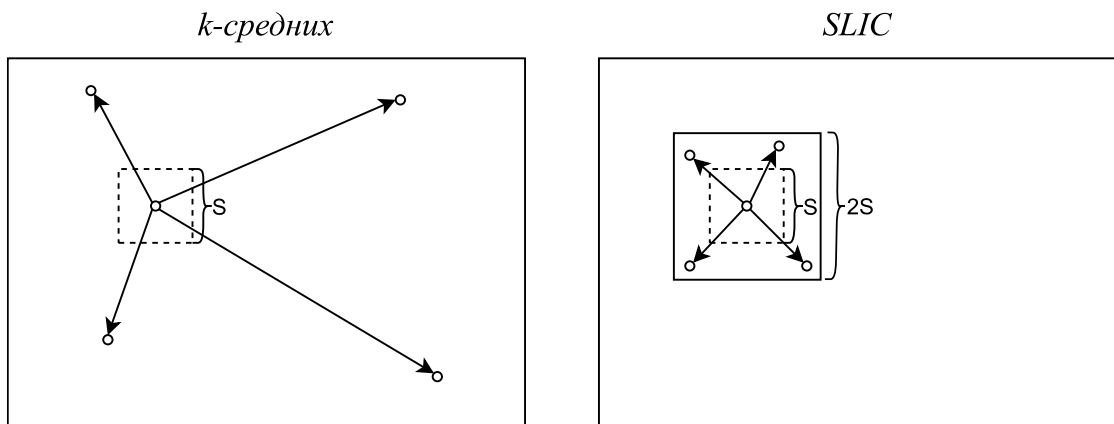


Рисунок 2.5 – Пространственное ограничение области поиска элементов относительно центра кластера в алгоритме SLIC

При проведении кластеризации пиксели цветного изображения характеризуются тремя значениями своих цветовых компонент в цветовом пространстве CIE LAB и координатами в матрице пикселей. Таким образом каждый пиксель описывается вектором из пяти признаков:  $\mathbf{p}_i = (l_i \ a_i \ b_i \ x_i \ y_i)^T$ ,  $i = \overline{1, N_p}$ .

В данном случае для вычисления расстояния между двумя пикселями нельзя использовать евклидово расстояние между векторами  $\mathbf{p}_i$  и  $\mathbf{p}_j$ ,

характеризующими эти пиксели, так как при достаточно большом размере суперпикселя пространственное расстояние будет вносить в результат больший вклад по сравнению с цветом. Поэтому используемое расстояние  $d^{(i,j)}$  (2.7) между двумя пикселями комбинирует нормализованные евклидовы расстояния в цветовом пространстве  $d_c^{(i,j)}$  (2.8) и в пиксельной матрице  $d_s^{(i,j)}$  (2.9):

$$d^{(i,j)} = \sqrt{\left(\frac{d_c^{(i,j)}}{D_c}\right)^2 + \left(\frac{d_s^{(i,j)}}{D_s}\right)^2}, \quad (2.7)$$

$$d_c^{(i,j)} = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}, \quad (2.8)$$

$$d_s^{(i,j)} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (2.9)$$

где  $D_c$  – максимальная дистанция в кластере по цвету,  $D_s$  – максимальная дистанция в кластере по расстоянию.

Из-за пространственного ограничения на область поиска можно положить  $D_s = S$ . Значение  $D_c$  будет варьироваться для разных суперпикселей (кластеров) и изображений, однако для упрощения вычислений его можно заменить некоторым константным значением. С учётом этих замен (2.7) можно привести к следующему виду:

$$d^{(i,j)} = \sqrt{d_c^{(i,j)2} + \left(\frac{d_s^{(i,j)}m}{S}\right)^2}, \quad (2.10)$$

где значение константы  $m$  позволяет контролировать веса расстояний  $d_c^{(i,j)}$  и  $d_s^{(i,j)}$ . Чем больше  $m$ , тем больший вес имеет пространственная близость, т. е. суперпиксели будут более компактными. Чем меньше  $m$ , тем больший вес имеет сходство по цвету, т.е. суперпиксели будут больше соответствовать границам объектов, но их форма и размеры будут варьироваться в большей степени.

Выбор начальных центров кластеров осуществляется не случайным образом, а исходя из требований к СС: начальные центры располагаются на сетке с шагом  $S$  (рисунок 2.6). Благодаря этому суперпиксели будут иметь средний размер  $S^2$

и будут равномерно распределены на изображении. В качестве начального центра кластера, авторы алгоритма SLIC предлагают выбирать пиксель с наименьшим градиентом из области размером  $3 \times 3$ , что уменьшает вероятность попадания начального центра на границу между кластерами или в зашумленный пиксель. Таким образом, решается проблема выбора исходных центров кластеров, которая присуща алгоритму  $k$ -средних и может негативно влиять на получаемый результат.

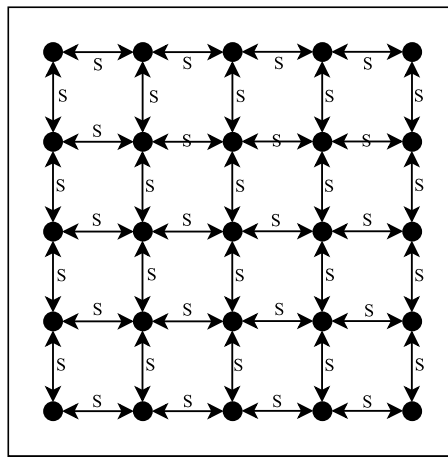


Рисунок 2.6 – Начальные центры кластеров (суперпикселей) алгоритма SLIC

После задания начальных центров кластеров  $\mathbf{c}_k$  и их начальных размеров  $N_k = S^2$ ,  $k = \overline{1, N_s}$ , выполняется итеративное уточнение разбиения на кластеры.

1. Для центра каждого кластера вычисляются расстояния до всех пикселей в его области поиска (рисунок 2.5) и затем сравниваются с расстояниями до центров кластеров, которым пиксели принадлежат в данный момент.

2. Если расстояние  $d^{(i,j)}$  от некоторого пикселя  $\mathbf{p}_i$  до центра  $j$ -го кластера  $\mathbf{c}_j$  меньше, чем расстояние  $d^{(i,k)}$  до центра  $k$ -го кластера  $\mathbf{c}_k$ , то пиксель  $\mathbf{p}_i$  переносится в  $j$ -ый кластер и центры кластеров пересчитываются:

$$N'_j = N_j + 1, N'_k = N_k - 1, \mathbf{c}'_j = \frac{\mathbf{c}_j + \mathbf{p}_i}{N'_j}, \mathbf{c}'_k = \frac{\mathbf{c}_k + \mathbf{p}_i}{N'_k}.$$

3. Если выполнено заданное число итераций осуществляется останов, в противном случае – переход на следующую итерацию (по утверждению авторов

алгоритма, для достижения приемлемой точности достаточно проведения 10 итераций).

После завершения итеративного процесса кластеризации может возникнуть ситуация, при которой некоторые из пикселей не были отнесены ни к одному из суперпикселей из-за пространственного ограничения на область поиска. Кроме того, возможно разделение суперпикселей на несколько изолированных частей или появление суперпикселей очень малого размера. Данные факторы обуславливают необходимость проведения постобработки полученных результатов.

Для этого изображение представляется в виде ненаправленного графа смежности пикселей, а полученные кластеры выделяются при помощи поиска компонентов его связности. Если обнаруженный кластер пикселей не был отнесён ни к одному из суперпикселей или его размер меньше некоторого порогового значения, то он становится частью ближайшего суперпикселя, имеющего наибольшую степень сходства. Иначе, данный кластер считается самостоятельным суперпикселем. Таким образом, после проведения постобработки количество суперпикселей может отличаться от заданного.

Блок-схема алгоритма SLIC представлена на рисунке 2.7.

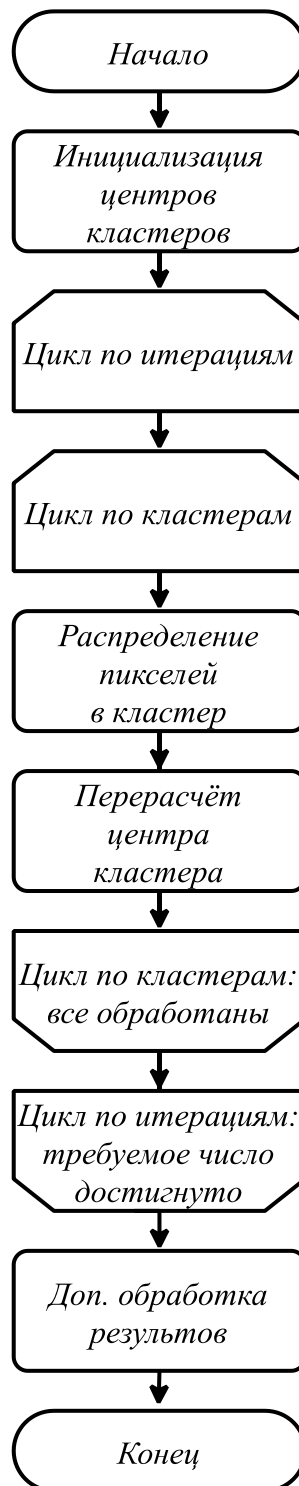


Рисунок 2.7 – Блок-схема алгоритма SLIC

Модификация алгоритма SLIC для учёта текстурных особенностей изображений. При использовании алгоритма SLIC в стандартном варианте в процессе кластеризации пикселей не учитывается текстура отдельных областей изображения. Данное обстоятельство хорошо проявляется на изображениях, у



которых потенциальные суперпиксели имеют примерно одинаковую яркость и цвета, но отличаются по текстуре (рисунок 2.8а). Стандартная реализация SLIC не способна распознать на таких изображениях границы между областями с разной текстурой и просто делит изображение на примерно одинаковые по размеру области (рисунок 2.8б). Такое отсутствие учёта текстурных особенностей сегментов присуще большинству алгоритмов СС.

Потенциальным решением данной проблемы является рассмотрение сразу несколько соседних пикселей одновременно (например, в блоках  $2 \times 2$ ), чтобы учитывать их взаимосвязи и текстурные особенности. В частности, можно учитывать дисперсии цветовых компонент изображения. С одной стороны, применение данного подхода к стандартному алгоритму SLIC позволяет расширить вектор признаков при проведении кластеризации, что потенциально повышает точность сегментации. С другой стороны, данное изменение в алгоритме является достаточно простым и не требующим серьёзных дополнительных вычислительных затрат.

С учётом данных соображений, расстоянием между некоторым блоком пикселей и центром кластера будет взвешенная сумма евклидовых расстояний между средними значениями координат блока, средними значениями цветовых компонент блока и дисперсиями цветовых компонент блока. Такая модификация алгоритма SLIC, включающая в процесс кластеризации дисперсию наблюдений в некоторой окрестности пикселя для учёта информации о текстуре, рассматривалась автором в [78, 79]. С учётом этой дополнительной информации, выражение для расчёта расстояния (2.10) принимает следующий вид:

$$d^{(i,j)} = \sqrt{d_c^{(i,j)2} + \left( \frac{d_s^{(i,j)} m}{S} \right)^2 + \alpha^2 d_v^{(i,j)2}},$$

$$d_v^{(i,j)} = \sqrt{(v_i^{(l)} - v_j^{(l)})^2 + (v_i^{(a)} - v_j^{(a)})^2 + (v_i^{(b)} - v_j^{(b)})^2},$$

где  $v_i^{(l)}$ ,  $v_i^{(a)}$ ,  $v_i^{(b)}$  – дисперсии цветовых компонент;  $\alpha > 0$  – настраиваемый весовой коэффициент.

Данная модификация алгоритма SLIC справляется с сегментацией областей, обладающих разной текстурой намного лучше, чем оригинальный алгоритм (рисунок 2.8в).

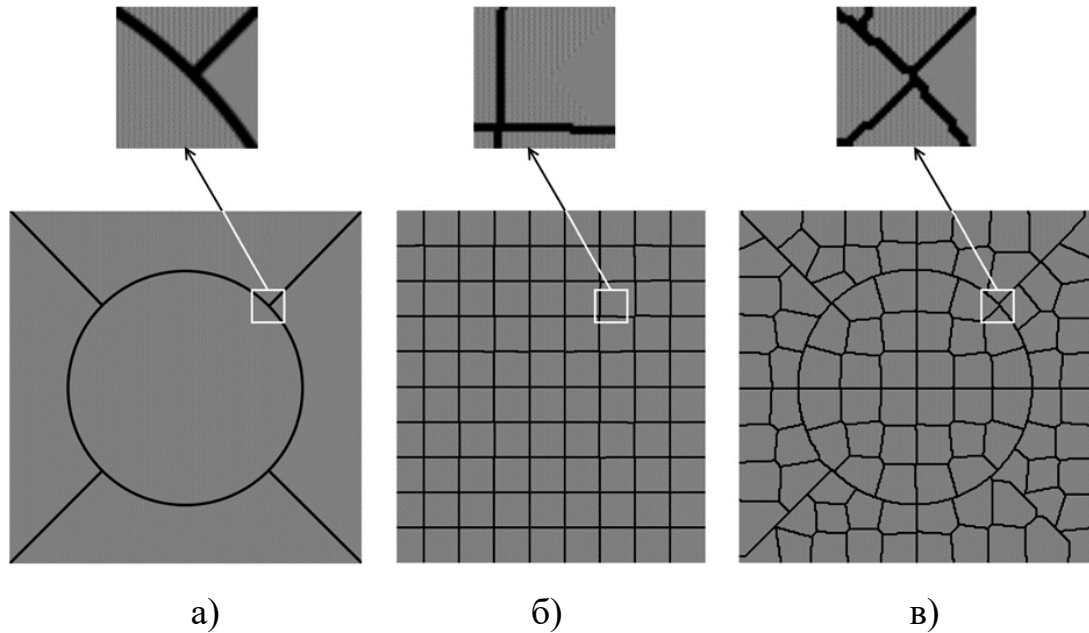


Рисунок 2.8 – Суперпиксельная сегментация изображения с областями примерно одинакового цвета, но с разной текстурой: а) реальные границы областей; б) результат оригинальной алгоритма SLIC; в) результат модифицированного алгоритма SLIC

На рисунке 2.9 представлен примеры СС, полученные при помощи оригинального и модифицированного алгоритма SLIC. Так как модифицированный алгоритм обрабатывает небольшие блоки пикселей, границы между суперпикселями являются менее гладкими по сравнению с оригинальным подходом.

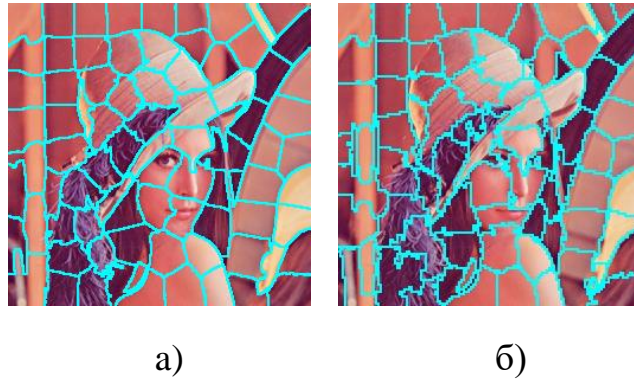


Рисунок 2.9 – Результат СС полученной при помощи алгоритма SLIC:  
а) стандартный вариант; б) модификация с учётом сведений о текстуре

### 2.3.2 EM-алгоритм и его применение для сегментации суперпикселей

Так как каждый суперпиксель состоит из достаточно большого числа пикселей, можно с большой степенью достоверности считать, что оценки параметров (математического ожидания, дисперсии и т.п.) цветовых компонент, получаемые путем усреднения данных по числу пикселей с большой степенью приближения могут рассматриваться как реализации случайных величин, описываемых нормальным распределением. В этом случае для дальнейшего разделения всех суперпикселей на классы полезных и аномальных наблюдений на втором этапе сегментации предлагается применять EM-алгоритм (Expectation-Maximization), широко применяемый для разделения гауссовских смесей [72].

При использовании EM-алгоритма количество классов  $M$  считается известным, но априорные вероятности появления наблюдений каждого из этих классов  $\omega_i$  считаются неизвестными:  $\pi_i = p(\omega_i)$ ,  $\sum_{i=1}^M \pi_i = 1$ ,  $i = \overline{1, M}$ . Вводится предположение относительно вида плотностей распределения классов  $p(\mathbf{x} / \theta_i)$ ,  $i = \overline{1, M}$ , где  $\theta_i$  – вектор неизвестных параметров плотностей распределений классов (в случае нормального распределения  $\theta_i = \{\mathbf{m}_i, \mathbf{C}_i\}$ ). Для разбиения всех наблюдений на кластеры необходимо предварительно найти оценку параметров

всех распределений  $\Theta = (\theta_1^T, \dots, \theta_M^T, \pi_1, \dots, \pi_M)^T$ , используя критерий максимума логарифма функции правдоподобия для выборки  $X^N = \{\mathbf{x}^{(k)}\}$ ,  $k = \overline{1, K}$ :

$$p(X^N / \Theta) = \prod_{k=1}^N p(\mathbf{x}^{(k)} / \Theta) = \prod_{k=1}^N \sum_{i=1}^M p(\mathbf{x}^{(k)} / \theta_i) \pi_i, \quad \tilde{\Theta} = \arg \max \ln P(X^N / \Theta). \quad (2.11)$$

Решение данной задачи напрямую не является возможным ввиду её сложности. Однако, можно ввести вспомогательную совокупность скрытых переменных  $Q = \{q_{ik}\}$ , которые можно вычислить, если известны параметры распределений  $\Theta$ . В качестве таких скрытых переменных удобно рассматривать апостериорные вероятности того, что образ  $\mathbf{x}^{(k)}$  был порождён классом  $\omega_i$ .

При инициализации алгоритма задаются начальные приближения параметров случайной смеси  $\tilde{\Theta} = (\tilde{\theta}_1^T, \dots, \tilde{\theta}_M^T, \pi_1, \dots, \pi_M)^T$ ,  $\tilde{\theta}_i = \{\tilde{\mathbf{m}}_i, \tilde{\mathbf{C}}_i\}$  (например, случайным образом) и априорных вероятностей  $\pi_i = \frac{1}{M}$ ,  $i = \overline{1, M}$ .

Уточнение оценок происходит итеративно в ходе процесса, который состоит из двух шагов, давших название алгоритму.

1. E-шаг (*expectation*) – вычисление условных математических ожиданий скрытых переменных  $q'_{ik}$  (апостериорных вероятностей) на основе текущего приближения параметров гауссовских распределений  $\tilde{\Theta}$ :

$$q'_{ik} = \frac{\tilde{\pi}_i N(\tilde{\mathbf{x}}^{(k)}, \tilde{\mathbf{m}}_i, \tilde{\mathbf{C}}_i)}{\sum_{i=1}^M \tilde{\pi}_i N(\tilde{\mathbf{x}}^{(k)}, \tilde{\mathbf{m}}_i, \tilde{\mathbf{C}}_i)}, \quad i = \overline{1, M}, \quad k = \overline{1, K}.$$

2. M-шаг (*maximization*) – вычисление нового приближения  $\tilde{\Theta}$  для максимизации правдоподобия наблюдений (2.11) на основе текущих значений скрытых переменных  $q'_{ik}$ :

$$\tilde{\pi}_i = \frac{1}{N} \sum_{k=1}^N q'_{ik} = \frac{N_i}{N}, \quad \tilde{\mathbf{m}}'_i = \frac{1}{N} \sum_{k=1}^N q'_{ik} \mathbf{x}^{(k)}, \quad \tilde{\mathbf{C}}'_i = \frac{1}{N} \sum_{k=1}^N q'_{ik} (\mathbf{x}^{(k)} - \tilde{\mathbf{m}}'_i)(\mathbf{x}^{(k)} - \tilde{\mathbf{m}}'_i)^T.$$

Итеративный процесс осуществляется до сходимости алгоритма или до достижения максимального числа итераций. По его завершении каждое

наблюдение зачисляется в тот кластер, вероятность нахождения в котором является для этого элемента максимальной.

Пиксели, входящие в состав суперпикселя  $S_i$ ,  $i = \overline{1, N_s}$ , описываются набором из пяти признаков (три цветовые компоненты в используемом цветовом пространстве и две координаты в матрице пикселей):

$\mathbf{p}_j^{(i)} = (c_1^{(i,j)}, c_2^{(i,j)}, c_3^{(i,j)}, x_1^{(i,j)}, x_2^{(i,j)})^T$ ,  $j = \overline{1, N_s^{(i)}}$ . Рассмотрим возможные значения для формирования вектора признаков  $\mathbf{s}^{(i)}$  (соответствует вектору  $\mathbf{x}^{(k)}$  в соотношениях выше), используемых для кластеризации суперпикселей при помощи EM-алгоритма

Так как положение ЛОЗ на изображениях является случайным, то они могут находиться в произвольных удалённых друг от друга областях изображения. Поэтому использование значений координат пикселей при проведении кластеризации не представляется целесообразным. При рассмотрении цветовых компонент пикселей, можно выделить следующие признаки для использования:

– средние значения цветовых компонент:

$$m_k^{(i)} = \frac{1}{N_i} \sum_{j=1}^{N_i} c_k^{(i,j)}, \quad k = 1, 2, 3;$$

– дисперсии цветовых компонент:

$$d_k^{(i)} = \frac{1}{N_i} \sum_{j=1}^{N_i} \left( c_k^{(i,j)} - m_k^{(i)} \right)^2, \quad k = 1, 2, 3;$$

– коэффициенты корреляции между цветовыми компонентами:

$$r_k^{(i)} = \frac{\sum_{j=1}^{N_i} \left( c_k^{(i,j)} - m_k^{(i)} \right) \left( c_{k'}^{(i,j)} - m_{k'}^{(i)} \right)}{\sqrt{\sum_{j=1}^{N_i} \left( c_k^{(i,j)} - m_k^{(i)} \right)^2} \cdot \sqrt{\sum_{j=1}^{N_i} \left( c_{k'}^{(i,j)} - m_{k'}^{(i)} \right)^2}}, \quad k = 1, 2, 3, \quad k' = (k + 1) \bmod 3 + 1;$$

– коэффициенты автокорреляции цветовых компонент:

$$r_k^{(i)} = \frac{\sum_{j=1}^{N_i} \left( c_k^{(i,j)} - m_k^{(i)} \right) \left( c_k^{(i,j')} - m_k^{(i)} \right)}{\sum_{j=1}^{N_i} \left( c_k^{(i,j)} - m_k^{(i)} \right)^2}, \quad k = 1, 2, 3, \quad j' = (j + 1) \bmod N_i + 1.$$

Так как средние значения цветовых компонент в наибольшей мере характеризуют свойства каждого суперпикселя, представляется целесообразным всегда использовать эти признаки, комбинируя их с остальными при необходимости. В частности, в случае использования средних значений и дисперсий цветовых компонент каждый суперпиксель будет характеризовать следующим вектором признаков, имеющим 6 элементов:

$$\mathbf{s}^{(k)} = (m_1^{(i)}, m_2^{(i)}, m_3^{(i)}, d_1^{(i)}, d_2^{(i)}, d_3^{(i)})^T, k = \overline{1, N_s}.$$

### 2.3.3 Особенности сегментации ложных наблюдения и постобработка результатов кластеризации

Традиционно, цифровые изображения представляются в цветовом пространстве RGB, состоящим из трёх каналов: красный, зелёный и синий цвета. Несмотря на то, что такой выбор каналов в RGB был обусловлен особенностями восприятия цвета человека, и это цветовое пространство получило большое распространение в технике, человеку трудно подобрать соответствующие значения цветовых компонент в RGB для заданного цвета. В частности, это связано с тем обстоятельством, что яркость цвета имеет наибольшее значение для человека, но в этом цветовом пространстве она зависит от всех цветовых компонент одновременно. Поэтому человеку, как правило, проще подобрать правильное значение компонент в цветовых пространствах, одним из компонентов которого является яркость.

Таким образом, с точки зрения проведения сегментации особый интерес представляет цветовое пространство CIE LAB, наиболее точно соответствующее особенностям восприятия цвета человеком, включая в себя яркость и две контрастных компоненты (красный-зелёный и синий-жёлтый). Поэтому использование CIE LAB целесообразно в задачах автоматической сегментации, которая должна повторять сегментацию, выполненную человеком вручную. Отсюда имеет смысл проводить разбиение суперпикселей на полезные и ложные наблюдения, используя это цветовое пространство.

Задача выделения АП на изображениях НР подразумевает разделение всех суперпикселей на два класса: «ложные» и «полезные» наблюдения. Однако на практике полученное множество суперпикселей зачастую лучше описывается смесью более чем двух нормальных распределений (например, соответствующих различным объектам или группам объектов на изображениях). Таким образом, при инициализации EM-алгоритма может быть целесообразным задавать число кластеров большее двух. Пример такой ситуации представлен на рисунке 2.10: разбиение суперпикселей на всего на два кластера (рисунок 2.10б) приводит к ошибочной классификации части полезного изображения как ложных наблюдений, в то время как при разбиении суперпикселей на большее число кластеров (рисунок 2.10в,г) такая проблема не возникает.

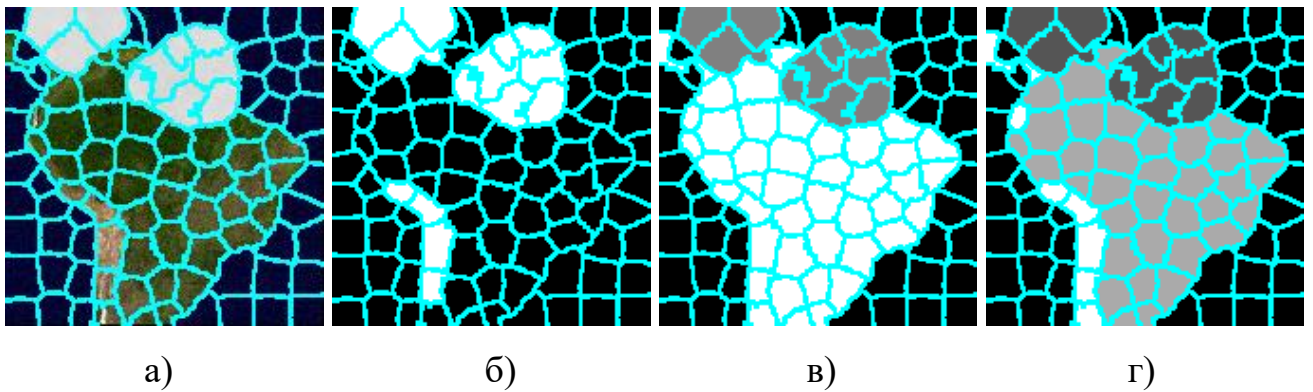


Рисунок 2.10 – Результаты разбиения суперпикселей на различное число классов:

а) СС изображения, подверженного воздействию АП; б-г) результат кластеризации суперпикселей на 2, 3 и 4 класса соответственно

После проведения кластеризации суперпикселей необходимо определить кластеры, соответствующие аномальным наблюдениям. Для этого можно воспользоваться априорной информацией о полезных или аномальных наблюдениях, чтобы понять какие из кластеров в большей мере сходны с ними.

Также возможна ситуация, в которой входное изображение не подвержено воздействию АП. В этом случае можно ввести некоторое пороговое значение при сравнении кластеров с априорной информацией о полезных или ложных

наблюдениях, а также учесть данную вероятность при включении результатов сегментации в оптимальный условно-линейный фильтр.

### 2.3.4 Включение результатов двухэтапной сегментации в процесс построения сверхразрешения

Для включения результатов независимой сегментации кадров НР входной последовательности в процесс оптимальной условно-линейной фильтрации (2.3-2.6) помимо получения векторов  $\theta_k$  также требуется задать матрицы вероятностей  $\mathbf{P}_{Ak\theta}$ ,  $\mathbf{P}_{Bk\theta}$ ,  $\mathbf{P}_{AAk\theta}$  и  $\mathbf{P}_{BBk\theta}$ , исходя из особенностей используемых для сегментации моделей и методов.

Для задания диагональных элементов матриц  $\mathbf{P}_{Ak\theta}$  и  $\mathbf{P}_{Bk\theta}$  требуется определить вероятности  $p(\theta_{kl}=0,1|a_{kl})$  и  $p(\theta_{kl}=0,1|b_{kl})$ ,  $l=\overline{1,M}$ . Одной из особенностей суперпиксельного представления изображений является тот факт, что все пиксели, принадлежащие одному из суперпикселей, рассматриваются при дальнейшей обработке как единое целое и, соответственно, должны иметь одинаковые вероятности принадлежности к полезным или ложным наблюдениям. Таким образом, если  $l$ -ый и  $m$ -ый пиксели изображения  $y_k$  принадлежат  $n$ -ому суперпикселю изображения ( $y_{kl} \in S_n$ ,  $y_{km} \in S_n$ ), то должны выполняться следующие соотношения:

$$\begin{aligned} p(\lambda_{kl}=0,1|a_{kl}) &= p(\lambda_{km}=0,1|a_{km}), \\ p(\lambda_{kl}=0,1|b_{kl}) &= p(\lambda_{km}=0,1|b_{km}), \\ p_{akl}(\theta) &= p_{akm}(\theta), \quad p_{bkl}(\theta) = p_{bkm}(\theta). \end{aligned}$$

Данные вероятности могут быть определены исходя из того, что для кластеризации использовалась гауссовская модель:

$$\begin{aligned} p(y_{kl} / \omega_O) &= N(\mathbf{s}_n, \mathbf{m}_a, \mathbf{C}_a), \\ p(y_{kl} / \omega_H) &= N(\mathbf{s}_n, \mathbf{m}_b, \mathbf{C}_b), \end{aligned}$$



где  $\mathbf{m}_a$  и  $\mathbf{C}_a$  – математическое ожидание и матрица ковариации полезных наблюдений;  $\mathbf{m}_b$  и  $\mathbf{C}_b$  – математическое ожидание и матрица ковариации аномальных наблюдений, оценённые при помощи EM-алгоритма;  $\mathbf{s}_n$  – вектор, описывающий признаки суперпикселя  $S_n$ , которому принадлежит пиксель  $y_{kl}$ .

В качестве апостериорных значений соответствующих вероятностей полезности наблюдений могут быть использованы значения апостериорных вероятностей суперпикселей, полученные на последнем шаге работы EM-алгоритма для соответствующего суперпикселя  $S_n$ :

$$p_{akl}(\theta) = q'_{an}, \quad p_{bkl}(\theta) = 1 - p_{akl}(\theta).$$

Особенностью предлагаемого подхода к сегментации ЛОЗ является тот факт, что суперпиксели всегда будут разбиты на классы полезных или ложных наблюдений вне зависимости от того, подвержено ли изображение воздействию АП. Для уменьшения негативных эффектов, которые могут быть вызваны этим обстоятельством, можно провести корректировку априорных вероятностей возникновения ложных наблюдений:

$$p_{akl} = \frac{|O_k|}{|Y_k|} (1 - p_\varepsilon), \quad p_{bkl} = 1 - p_{akl},$$

где  $p_\varepsilon$  – верхняя граница ошибки классификации. Её получение осуществляется на основе верхней границы Чернова [95].

$$p_\varepsilon = \sqrt{\frac{N_a N_b}{N_s^2}} \exp\left[-\mu\left(\frac{1}{2}\right)\right],$$

$$\mu\left(\frac{1}{2}\right) = \frac{1}{8} (\mathbf{m}_a - \mathbf{m}_b)^T \left(\frac{\mathbf{C}_a + \mathbf{C}_b}{2}\right)^{-1} (\mathbf{m}_a - \mathbf{m}_b) + \frac{1}{2} \ln \frac{|\mathbf{C}_a| + |\mathbf{C}_b|}{2\sqrt{|\mathbf{C}_a| \cdot |\mathbf{C}_b|}},$$

где  $\mu(1/2)$  – расстояние Бхаттачария;  $N_a$  и  $N_b$  – количество суперпикселей, отнесённых в процессе сегментации к полезным и ложным наблюдениям соответственно;  $N_s = N_a + N_b$  – общее количество суперпикселей.

Для численного задания элементов матриц попарных вероятностей полезных  $\mathbf{P}_{AAk}$  или аномальных  $\mathbf{P}_{BBk}$  наблюдений необходимо произвести

усреднение элементов  $\theta_k$ , соответствующих пикселям в окрестностях каждого элемента изображения ИР, для которого в  $\theta_k$  содержатся единичные и нулевые значения соответственно. Пусть матрица  $C_{Aki}$  имеет значения элементов  $\theta_k$  в окрестности  $i$ -го единичного элемента  $\theta_k$ , а матрица  $C_{Bkj}$  – значения  $1-\theta_k$  в окрестности  $j$ -го нулевого элемента  $\theta_k$ . Строки матриц  $P_{AAk\theta}$  и  $P_{BBk\theta}$  соответствуют векторным разверткам матриц  $\bar{C}_{Aki}$  и  $\bar{C}_{Bkj}$ , вычисляемых по всем матрицам  $C_{Aki}$  и  $C_{Bkj}$ :  $\bar{C}_{Aki} = M_i\{C_{Aki}\}$ ,  $\bar{C}_{Bkj} = M_j\{C_{Bkj}\}$ .

Таким образом, мы видим, что при решении задачи в адаптивной постановке структура алгоритма усложняется. Полученный в итоге алгоритм обработки может рассматриваться как комплексный алгоритм обработки информации в интересах построения многокадрового сверхразрешения. Данный алгоритм основан на использовании методов оптимальной фильтрации последовательности кадров низкого разрешения и проведении независимой сегментации участков аномальных наблюдений (участков, пораженных аппликативными помехами) на каждом изображении с целью использования результатов сегментации для вычисления коэффициентов фильтров, используемых при обновлении оценки ВР. Детальная схема такого алгоритма представлена на рисунке 2.11.

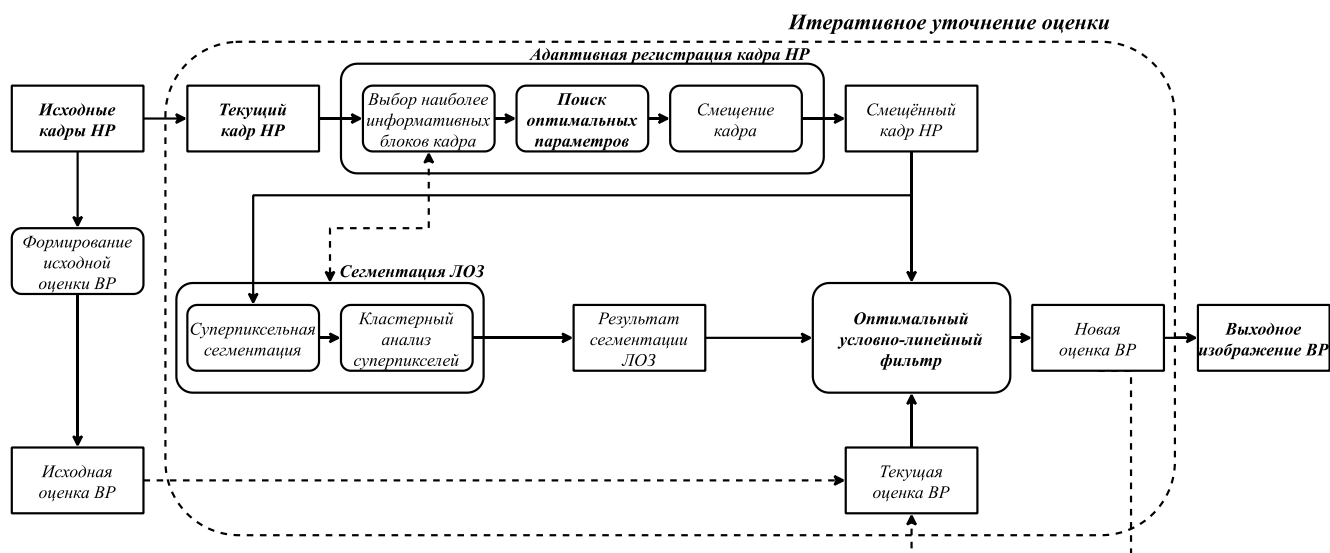


Рисунок 2.11 –Схема итеративного уточнения оценки изображения ВР на основе оптимальной фильтрации в адаптивной постановке с подбором параметров межкадровых сдвигов и реализацией блочной обработки изображений

## 2.4 Результаты экспериментального исследования эффективности алгоритмов, основанных на оптимальной фильтрации

### 2.4.1 Рассматриваемые методы и алгоритмы

Как было отмечено ранее, существующие исследования в области многокадрового СР в условиях ложных наблюдений не содержат сведений об экспериментальных сравнениях таких алгоритмов друг с другом, поэтому в рамках данной работы рассматриваются несколько альтернативных методов [52–54].

В ходе экспериментального исследования проводилось сравнение точности работы следующих алгоритмов:

- алгоритм, основанный на модели спиновых стёкол (АМСС) [52];
- алгоритм, основанный на модели спиновых стёкол с учетом инерционного движения АП (АМССИД) [53];
- алгоритм, основанный на моделях марковских случайных полей (АМСП) [54];

– алгоритм, основанный на использовании методов оптимальной фильтрации (АОФ) [55];

– алгоритм, основанный на использовании методов оптимальной фильтрации в адаптивной постановке для межкадровых смещений (АОФА) [65].

Каждый из данных алгоритмов был реализован в среде Matlab с учётом блочного подхода к обработке изображений, обеспечивающего понижение размерности обрабатываемых векторов и матриц. При этом обработка блоков происходила в однопоточном режиме. Кроме того, реализации данных алгоритмов используют возможности платформы CUDA, использующей возможности видеокарт для ускорения матричных вычислений.

Для всех рассматриваемых выше алгоритмов (кроме АМСС и АМССИД) требуется задание начальной оценки, которая уточняется в ходе работы алгоритма. Было рассмотрено два метода формирования такой оценки:

– первое изображение НР из исходной последовательности (последовательное повышение разрешения при поступлении новых наблюдений);

– усреднённое изображение НР, полученное на основе всей последовательности с учётом межкадровых смещений и локализованных областей, подверженных воздействию АП.

В обоих случаях разрешение начальной оценки повышается до целевого при помощи алгоритма однокадрового СР.

Для локализации ЛОЗ на входных изображениях использовался рассмотренный выше алгоритм двухэтапной сегментации на основе суперпиксельного представления изображений.

#### **2.4.2 Формирование тестового набора данных**

Для экспериментального сравнения рассматриваемых алгоритмов построения многокадрового СР в условиях АП были сформированы 20 наборов цветных изображений НР размера  $128 \times 128$ , состоящие из 20 кадров каждый.

Примеры кадров НР из тестового набора изображений с различной локализацией АП приведены на рисунке 2.12.



Рисунок 2.12 – Примеры изображений из входной последовательности кадров НР, подвергнутых воздействию АП

Каждый набор тестовых изображений создавался из одного исходного изображения ВР разрешения равного или выше  $256 \times 256$  путём последовательного наложения АП и аддитивного шума, применения случайных аффинных преобразований и прореживания искажённого изображения ВР с целью понижения его разрешения. Каждому набору тестовых данных соответствовала своя случайная формирующая бинарная маска, соответствующая областям полезных и ложных наблюдений на изображениях. Ложные наблюдения формировались как реализации случайного поля с заданными параметрами и накладывались на изображение ВР в соответствии с текущей маской ЛОЗ. После этого выполнялось прореживание центральной части полученного изображения ВР размером  $256 \times 256$  для создания очередного кадра НР с разрешением  $128 \times 128$ . При формировании каждого следующего кадра НР указанные действия повторялись, при этом моделировалось перемещение камеры со случайным сдвигом и поворотом относительно ранее полученного изображения, а также перемещение областей, поражённых АП, путем генерации смещения и поворота формирующей бинарной маски. Схема формирования тестовых изображений представлена на рисунке 2.13.

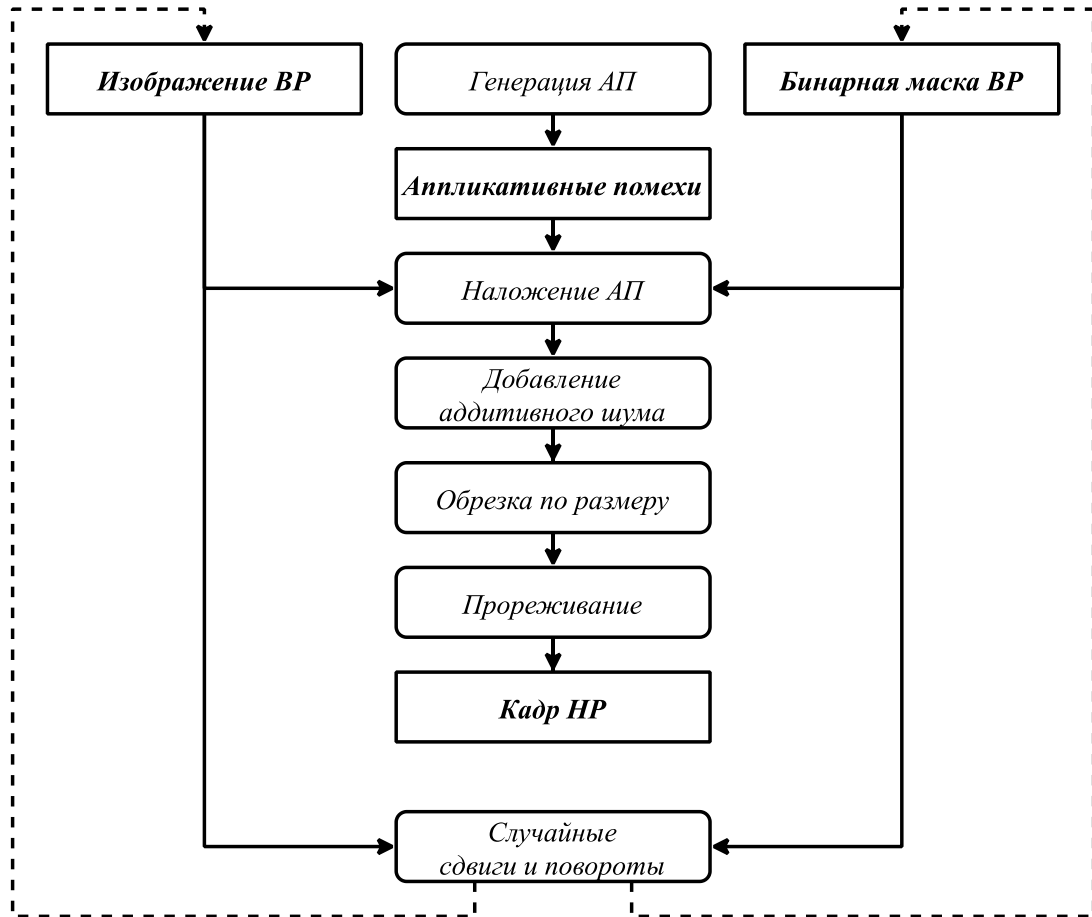


Рисунок 2.13 – Схема формирования тестовых наборов изображений

Для генерации исходной бинарной маски ЛОЗ использовался алгоритм, описанный в [96], позволяющий генерировать пятна случайной формы. На рисунке 2.14 приведены примеры получаемых с его помощью бинарных масок.



Рисунок 2.14 – Примеры сгенерированных бинарных масок ЛОЗ

### 2.4.3 Анализ результатов экспериментального сравнения алгоритмов

В ходе экспериментального исследования разрешение исходных кадров НР повышалось в два раза до  $256 \times 256$ .

Для повышения разрешения исходной оценки использовался алгоритм VDSR, основывающийся на использовании глубокой свёрточной нейронной сети [24]. На рисунке 2.15 представлен пример одного из оригинальных изображений ВР, с которыми будет проводиться сравнение результатов работы алгоритмов, а также примеры начальной оценки, полученной на основе двух рассматриваемых подходов.

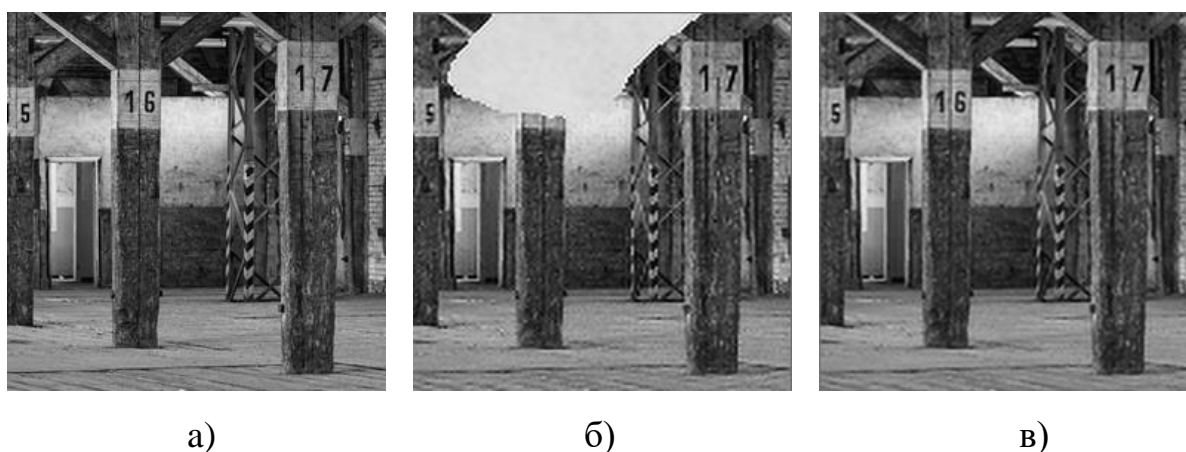


Рисунок 2.15 – Пример начальных оценок: а) оригинальное изображение ВР; б) начальная оценка в виде первого кадра обрабатываемой последовательности; в) начальная оценка в виде усреднённых кадров НР

При реализации блочной обработки изображений использовались следующие параметры:

- количество блоков по вертикали и горизонтали  $s_{in} = 4N_1 = N_2 = 32$ ;
- размер центральной части блоков  $s_{in} = 4$ ;
- ширина перекрывающихся частей блока  $\Delta s = 2$ .

При сегментации ЛОЗ с помощью двухэтапного алгоритма проводилось разбиение пикселей на 100 суперпикселей и кластеризация полученных суперпикселей на 4 первоначальных класса.

На рисунках 2.16 и 2.17 представлены примеры изображений ВР, полученных при помощи рассматриваемых алгоритмов. Качественный анализ полученных результатов показывает, что АМСС и АМССИД хуже справляются с устранением воздействия АП по сравнению с аналогами. АМСП выполняет эту задачу несколько лучше, однако его результаты в большей мере зависят от начальной оценки и менее детализированы по сравнению с результатами АОФ и АОФА, которые справляются с задачей устранения АП наряду с одновременным повышением разрешения лучше остальных.

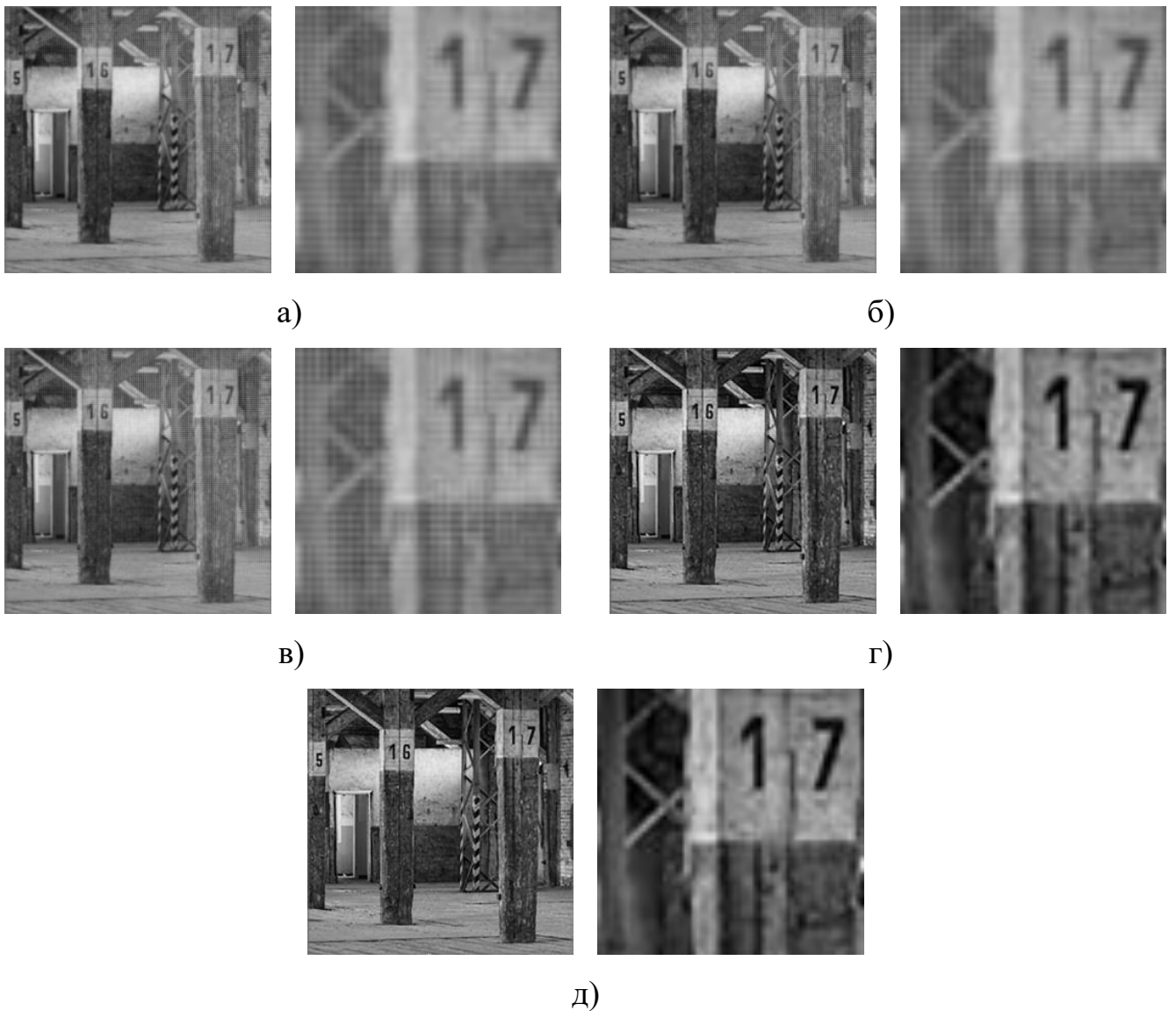


Рисунок 2.16 – Примеры изображений ВР, полученных на основе начальной оценки в виде первого изображения НР с VDSR, и их увеличенные фрагменты:

а) АМСС; б) АМССИД; в) АМСП; г) АОФ; д) АОФА



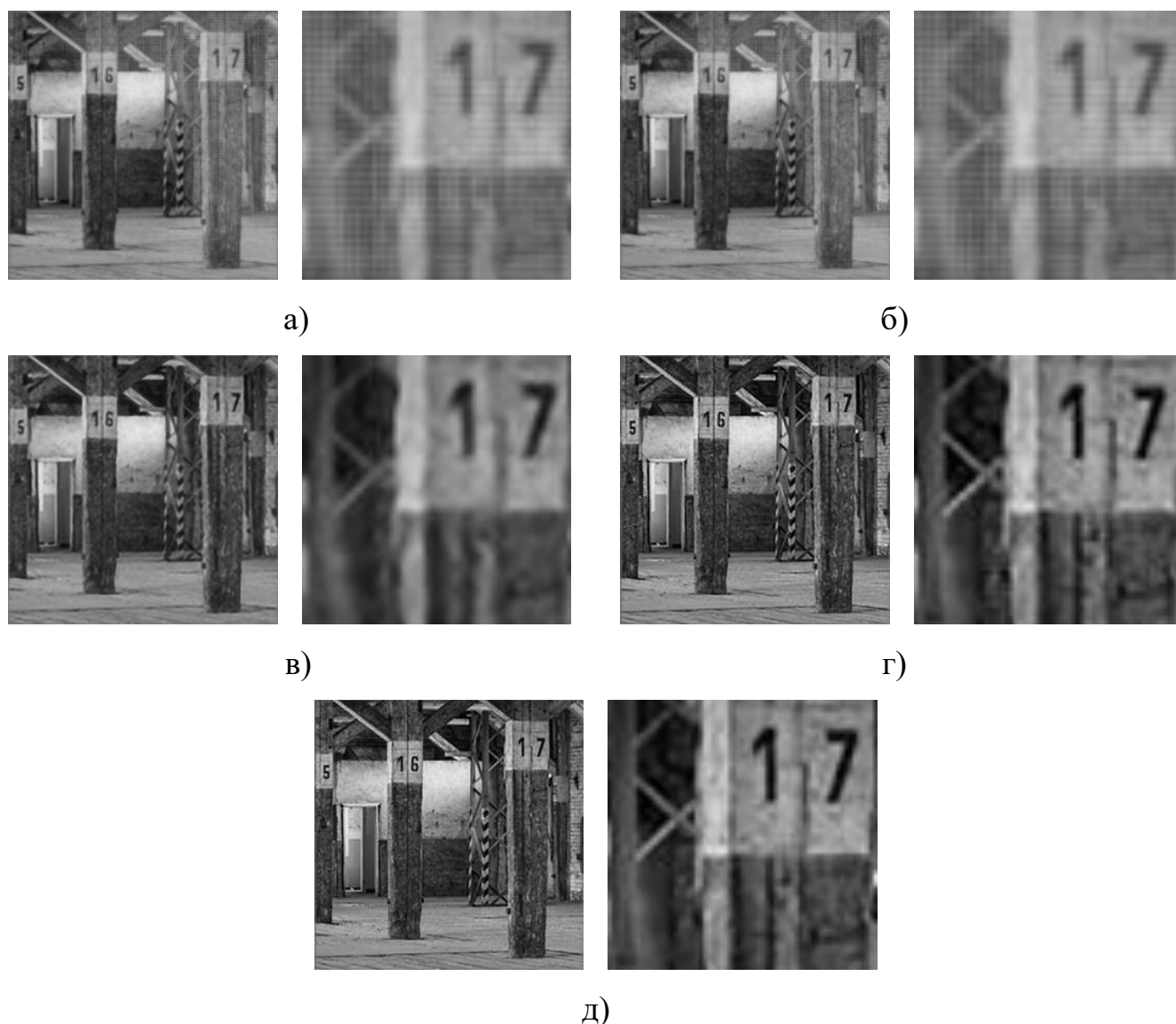


Рисунок 2.17 – Примеры изображений ВР, полученных на основе усредненной начальной оценки с VDSR, и их увеличенные фрагменты: а) АМСС; б) АМССИД; в) АМСП; г) АОФ; д) АОФА

Для численного сравнения результатов с точки зрения качества получаемых изображений ВР применялись следующие показатели.

1. Пиковое отношение сигнала к шуму (Peak Signal-to-Noise Ratio, PSNR; чем больше, тем лучше) измеряется в децибелах и оценивает степень сходства между двумя изображениями. Данная метрика задаётся на основе отношения максимального значения пикселей в изображении к среднеквадратической ошибке (MSE) исходного и восстановленного изображений и не зависит от диапазона значений, которые могут принимать пиксели:

$$PSNR = 20\log_{10}(MAX_I) - 10\log_{10}(MSE),$$

где  $MAX_I$  – максимальное значение пикселя в изображениях;  $MSE$  – среднеквадратическая ошибка между двумя изображениями.

2. Индекс структурного сходства (Structural Similarity Index, SSIM; чем больше, тем лучше) [87] также, как и PSNR и MSE, оценивает насколько два изображения схожи друг с другом, но, в отличие от этих критериев, учитывает «восприятие» различий между изображениями человеком, используя анализ структурных изменений свойств исходного и восстановленного изображений.

3. Индекс качества изображений (Naturalness Image Quality Evaluator, NIQE, чем меньше, тем лучше) [88] оценивает качество изображения без эталонного образца путём анализа его статистических свойств.

Усреднённые значения этих критериев по всей группе обрабатываемых последовательностей кадров для двух подходов к формированию начальной оценки приведены в таблицах 2.1 и 2.2.

Таблица 2.1 – Точность восстановления исходного изображения ВР в случае использования первого изображения НР как начальной оценки

Критерий	Исходная оценка	АМСС	АМССИД	АМСП	АОФ	АОФА
<b>PSNR</b>	16,28	19,22	19,33	19,62	26,46	27,04
<b>SSIM</b>	0,85	0,85	0,88	0,87	0,94	0,95
<b>NIQE</b>	10,13	6,83	6,59	6,53	6,11	6,12

Таблица 2.2 – Точность восстановления исходного изображения ВР в случае использования усредненной начальной оценки

Критерий	Исходная оценка	АМСС	АМССИД	АМСП	АОФ	АОФА
<b>PSNR</b>	26,58	19,22	19,33	22,31	26,80	27,23
<b>SSIM</b>	0,87	0,87	0,89	0,87	0,94	0,96
<b>NIQE</b>	7,58	6,83	6,59	6,72	6,35	6,33

Также проводилось сравнение быстродействия рассматриваемых алгоритмов, результаты которого представлены в таблице 2.3, содержащей среднее время обработки последовательности изображений при помощи рассматриваемых алгоритмов. Вычисления проводились в среде Matlab 2019b на персональном компьютере со следующими системными характеристиками:

- процессор Ryzen 3600 3.6 ГГц;
- видеокарта GTX 1660 6 Гб;
- ОЗУ: 16 Гб.

Таблица 2.3 – Средняя скорость работы рассматриваемых алгоритмов

Критерий	АМСС	АМССИД	АМСП	АОФ	АОФА
<b>Время (сек)</b>	943	967	704	408	517

Анализ полученных результатов показывает, что алгоритмы, основанные на использовании оптимальной фильтрации (АОФ и АОФА), имеют большую точность восстановления исходного изображения ВР по сравнению с аналогами. При этом точность результатов алгоритма, использующего адаптивный перебор параметров (АОФ) ожидаемо превосходит точность алгоритма, использующего традиционную постановку задачи (АОФ).

Кроме того, практически все алгоритмы демонстрируют большую точность при реализации второго способа формирования начальной оценки на основе усреднения всех кадров НР входной последовательности. Тем не менее,

использование начальной оценки на основе только первого кадра в сочетании с применением АОФА позволяет получить сходную точность при сохранении возможности реализации обработки в режиме последовательного накопления информации, что может быть полезным на практике.

Также стоит отметить, что все рассматриваемые алгоритмы имеют быстроедействие одного порядка и обработка всей последовательности изображений в однопоточном режиме занимает несколько минут. Данное время может быть снижено за счёт параллельной обработки блоков изображения, однако такая организация такой обработки, хоть и позволит снизить время работы алгоритма в несколько раз, будет требовать большей вычислительной мощности.

### **Выводы по главе**

1. Предложена и проанализирована общая схема итеративного уточнения оценки изображения ВР при обработке поступающей последовательности кадров НР, подверженных воздействию АП. В целях повышения точности получаемых результатов при наличии ложных наблюдений на изображениях, предложены дополнительные этапы формирования начальной оценки на основе исходных изображений НР, а также независимой сегментации областей, подверженных аномальным наблюдениям, на каждом очередном обрабатываемом кадре НР.

2. Рассмотрены модели наблюдений изображений НР и состояний изображений ВР и основывающийся на них алгоритм итеративного обновления текущей оценки изображения ВР на основе оптимальной условно-линейной фильтрации калмановского типа, включающий в себя результаты независимой сегментации пикселей каждого из входных изображений НР на полезные и ложные наблюдения. Приведены математические соотношения для получения обновлённой оценки изображения ВР на основе нового зарегистрированного наблюдения. Рассмотрена адаптивная постановка задачи оптимальной фильтрации с учётом параметрической неопределённости используемой модели. Приведена схема комплексного алгоритма многокадрового СР на основе

оптимальной фильтрации в адаптивной постановке и двухэтапного алгоритма сегментации ЛОЗ на входных кадрах НР.

3. Предложен двухэтапный алгоритм сегментации областей, подверженных воздействию АП, на изображениях, основывающийся на суперпиксельном представлении изображений с последующим кластерным анализом полученных суперпикселей при помощи EM-алгоритма. Рассмотрен алгоритм CC SLIC и предложена его модификация, позволяющая учитывать текстурные различия между отдельными участками изображений при выделении суперпикселей. Проанализированы особенности использования предложенного алгоритма для сегментации ЛОЗ. Приведены математические соотношения для включения получаемых результатов сегментации в оптимальную условно-линейную фильтрацию поступающей последовательности изображений НР.

4. Проведено экспериментальное исследование рассмотренных выше алгоритмов на основе оптимальной фильтрации в условно-линейной и адаптивной (нелинейной) постановке, использующих предложенный двухэтапный алгоритм сегментации. Анализ экспериментального сравнения показывает эффективность предложенного подхода по сравнению с альтернативными методами построения многокадрового СР в условиях наличия АП на входных изображениях, заключающуюся в большей точности получаемой оценки изображения ВР. При этом наибольшую точность имеет алгоритм, основанный на оптимальной фильтрации в адаптивной постановке задачи. Основным недостатком рассмотренных подходов является большое время обработки последовательности кадров.

### **3 Синтез и анализ алгоритмов построения многокадрового сверхразрешения на основе использования глубоких нейронных сетей**

#### **3.1 Использование свёрточных нейронных сетей для построения многокадрового сверхразрешения**

Как было показано в главе 2, алгоритмы многокадрового СР на основе оптимальной фильтрации справляются с задачей устранения АП на входных изображениях и превосходят по точности альтернативные алгоритмы для решения данной задачи в условиях АП. Однако данный класс алгоритмов, как и другие традиционные алгоритмы, основан на задании математических моделей изображений и помех и имеет ряд недостатков.

Во-первых, можно поставить под сомнение адекватность используемых математических моделей. Так, условно-гауссовская модель данных, описывающая изображения, зачастую не соответствует реальности. Во-вторых, сложность используемых моделей и большая размерности обрабатываемых векторов и матриц ведут к длительному времени обработки исходной последовательности кадров НР даже в случае организации параллельной блочной обработки изображений.

Потенциальным решением описанных проблем является использование технологий машинного обучения и, в частности, глубоких нейронных сетей. Данный подход не требует задания сложных математических моделей для описания входных данных, а быстроедействие обученных нейронных сетей потенциально намного выше, чем быстроедействие оптимального фильтра для обработки данных большой размерности. Кроме решения задачи обновления оценки ВР, нейронные сети позволяют решать и другие задачи, возникающие в ходе построения СР (например, оценки смещений между кадрами и сегментации ЛОЗ). Следует также отметить, что данная парадигма решения задач в области обработки изображений свободна от использования математических моделей

данных, но требует существенных по объему обучающих выборок, содержащих примеры анализируемых изображений.

В плане решения задачи построения СР наибольший интерес представляют глубокие свёрточные сети. В [22] было показано, как различные этапы обработки изображений при решении данной задачи могут аппроксимироваться свёрточными слоями. Кроме того, универсальность свёрточных слоёв позволяет использовать уже обученные сети для обработки изображений разных размеров.

Полученные ранее в главе 2 результаты позволяют утверждать, что потенциальные возможности итеративной обработки последовательности входных изображений НР на основе общих схем, представленных на рисунках 2.1 и 2.2, достаточно велики в плане достижения высокого качества получаемых изображений. Поэтому такой подход к обработке изображений представляется разумным и для построения алгоритмов, основанных на использовании глубоких нейронных сетей.

В этом плане можно рассмотреть два подхода к синтезу комплексных алгоритмов СР на основе использования свёрточных нейронных сетей. Во-первых, так как в известной литературе не рассматривался вопрос их использования для построения СР в условиях АП, имеет смысл рассмотреть возможности нейронных сетей такого класса для фильтрации АП на входных изображениях без использования отдельного этапа сегментации ЛОЗ. Во-вторых, высокая точность результатов, получаемых при помощи комплексных алгоритмов на основе оптимальной фильтрации и сегментации ЛОЗ, обуславливает целесообразность синтеза алгоритма, также включающего в себя этап обнаружения областей, подверженных воздействию АП.

## **3.2 Алгоритм построения многокадрового сверхразрешения на основе свёрточных нейронных сетей без использования независимой сегментации локальных областей закрытия**

### **3.2.1 Схема обработки входной последовательности кадров**

Для построения многокадрового СР на основе итеративного подхода (рисунок 2.1) требуется решить следующие основные задачи:

- регистрация с предобработкой очередного кадра НР, заключающаяся в оценке смещения этого кадра относительно других;
- обновление оценки изображения ВР на основе текущей оценки и очередного зарегистрированного (смещённого) кадра;
- формирование начальной оценки изображения ВР.

В рамках комплексного подхода [64], основанного на использовании глубоких нейронных сетей, для регистрации очередного изображения предлагается использовать нейронную сеть LiteFlowNet2 [90], оценивающую оптический поток между двумя изображениями, а для итеративного уточнения оценки ВР – модифицированную сеть SRNet из алгоритма FRVSR [43], повышающую качество поступающих на вход изображений.

В качестве начальной оценки ВР используется усреднённое изображение, полученное на основе зарегистрированных кадров НР, чьё разрешение было повышено до целевого при помощи бикубической интерполяции.

С учётом данных соображений процесс обработки входной последовательности кадров описывается схемой, представленной на рисунке 3.1. Далее основные шаги этого процесса описываются более подробно.



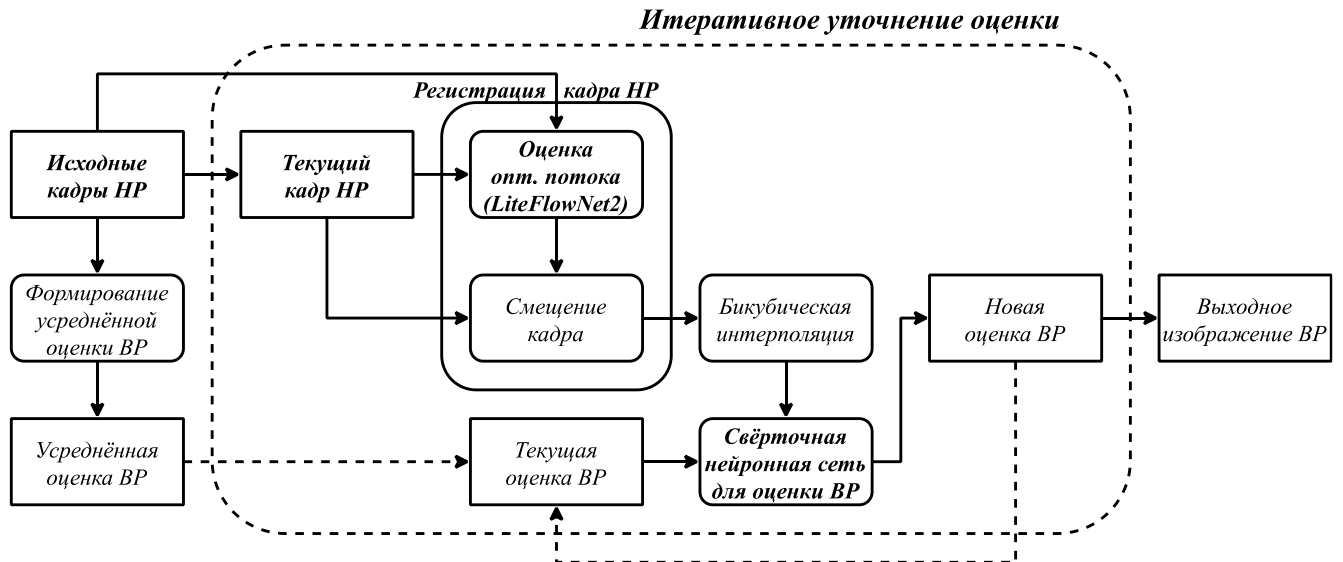


Рисунок 3.1 – Схема обработки входной последовательности кадров НР на основе свёрточных нейронных сетей

### 3.2.2 Оценка оптического потока между изображениями

Для регистрации очередного поступившего на вход изображения НР выполняются следующие шаги:

- оценка оптического потока между очередным и первым кадрами последовательности;
- сдвиг очередного кадра НР на основе полученного оптического потока.

Оптический поток между кадрами оценивается при помощи нейронной сети LiteFlowNet2, предложенной в [90] и реализующей пирамидальный подход к оценке результата, где каждый уровень последовательно повышает детализацию получаемого оптического потока. При этом каждый уровень состоит из двух подсетей (NetC и NetE), имеющих переменное число слоёв. Подсеть NetC выделяет признаки входных изображений, а подсеть NetE последовательно оценивает оптический поток на их основе, а также выполняет регуляризацию результата. Схема оценки оптического потока при помощи LiteFlowNet2 представлена на рисунке 3.2.

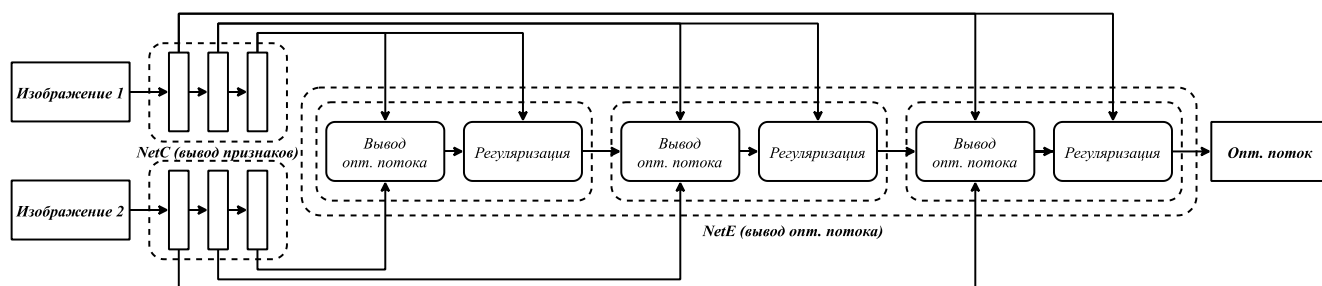


Рисунок 3.2 – Структура LiteFlowNet2 с тремя уровнями для оценки оптического потока между двумя кадрами

### 3.2.3 Итеративное уточнение оценки изображения высокого разрешения при помощи свёрточной нейронной сети

Для уточнения оценки изображения ВР используется модифицированная свёрточная нейронная сети SRNet из [43] (FRVSR), которая принимает на вход два изображения:

- очередной зарегистрированный (смещённый) кадр НР;
- сформированная к текущему шагу оценка изображения ВР.

На выходе данной нейронной сети получается одно изображение ВР, которое является обновлённой оценкой выходного изображения ВР.

Так как входные изображение имеют разные разрешения, необходимо произвести их преобразование, чтобы привести их к одному разрешению. Одной из особенностей алгоритма FRVSR [43] является то обстоятельство, что нейронная сеть SRNet для повышения разрешения принимает на вход два изображения НР. При этом кадр НР, имеющий размер  $(h_L, w_L, c)$ , остаётся без изменений, а текущая оценка изображения ВР размером  $(h_H, w_H, c)$  преобразуется при помощи space-to-width трансформации в изображение размером

$(h_L, w_L, (m^2 + 1)c)$ , где  $m = \frac{h_H}{h_L} = \frac{w_H}{w_L}$ . Повышение разрешения для получения

выходного изображения ВР, имеющего исходный размер  $(h_H, w_H, c)$ , осуществляется при помощи транспонированного свёрточного слоя. С одной стороны, такой подход позволяет оптимизировать вычисления; с другой –

структура сети для повышения разрешения перестаёт быть универсальной, так как позволяет повысить разрешение только ровно в  $m$  раз (например, в [43]  $m = 2$ ).

Альтернативным подходом, используемым в ряде других работ по SR [24, 44], является предварительное повышение разрешения входного кадра до требуемого (BP) методом бикубической интерполяции. Таким образом, задача построения SR сводится к задаче улучшения детализации входного изображения.

Архитектура нейронной сети, решающей данную задачу в этой работе, основывается на сети SRNet из [43] с учётом вышеизложенных соображений и представлена на рисунке 3.3. В отличие от оригинальной сети SRNet в данной работе транспонированные свёрточные слои не используются ввиду отсутствия необходимости увеличивать разрешение входных изображений. При этом остаточные связи между слоями были сохранены.

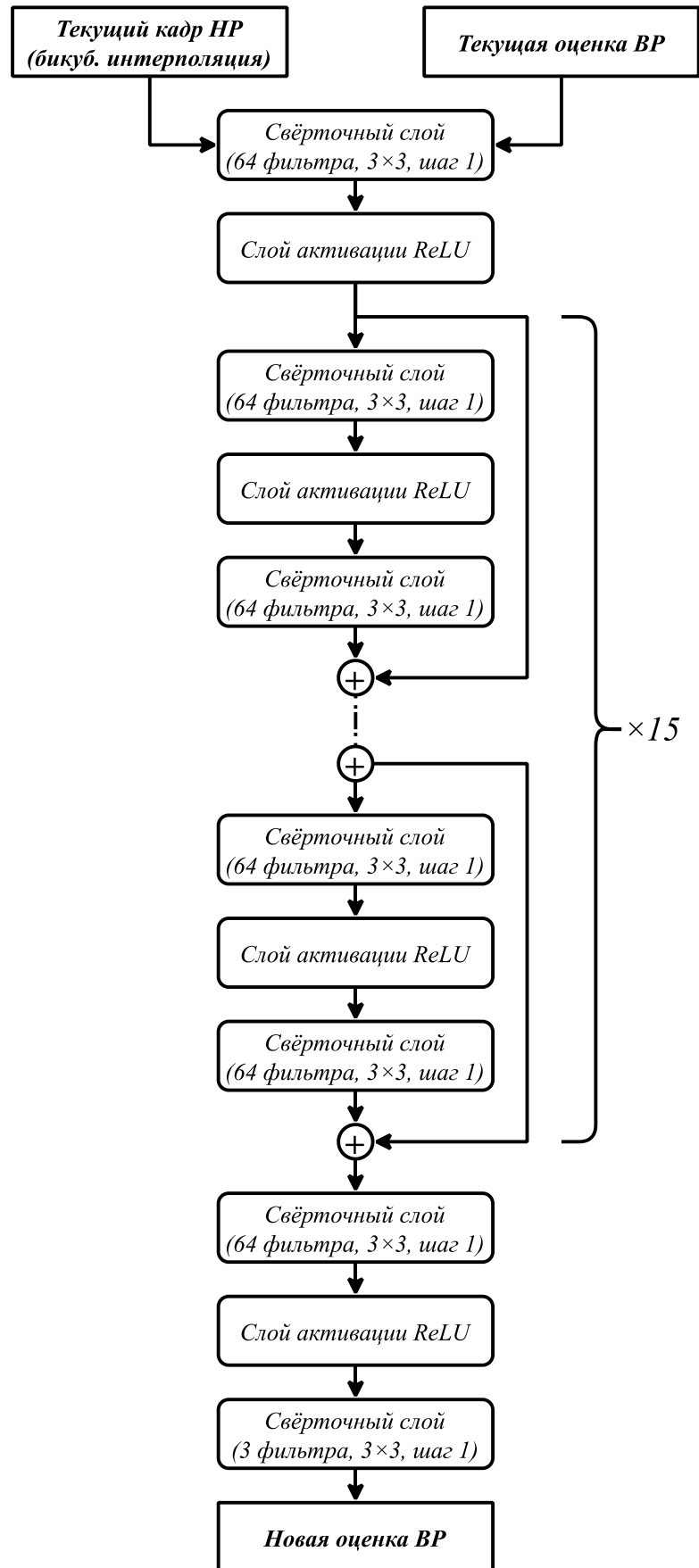


Рисунок 3.3 – Архитектура нейронной сети для обновления оценки изображения  
ВР

При обработке первого изображения текущая оценка выходного изображения ВР неизвестна, поэтому на второй вход нейронной сети подаётся усреднённое изображение ВР, полученное на основе всей последовательности кадров НР и увеличенное до целевого разрешения при помощи бикубической интерполяции.

### **3.2.4 Экспериментальное исследование**

**Обучение используемых нейронных сетей.** В рамках экспериментального исследования осуществлялось повышение разрешения в 2 раза. Обучающая выборка для обучения нейронной сети для обновления оценки ВР была сформирована на основе набора данных SUN [100], который состоит из изображений, отображающих различные категории объектов. Создание выборки осуществлялось путём случайных смещений и поворотов исходных изображений, к которым случайным образом применялись АП. Для этого использовалась схема (рисунок 3.4), схожая со схемой формирования тестовых изображений для экспериментального исследования алгоритмов в главе 2, п. 2.4 (рисунок. 2.13), но с учётом того, что на вход нейронной сети подаётся изображение НР, увеличенное при помощи бикубической интерполяции. При этом некоторые из кадров НР или даже целые их последовательности не были искажены аномальными наблюдениями, так как генерация случайных помех носит случайный характер.

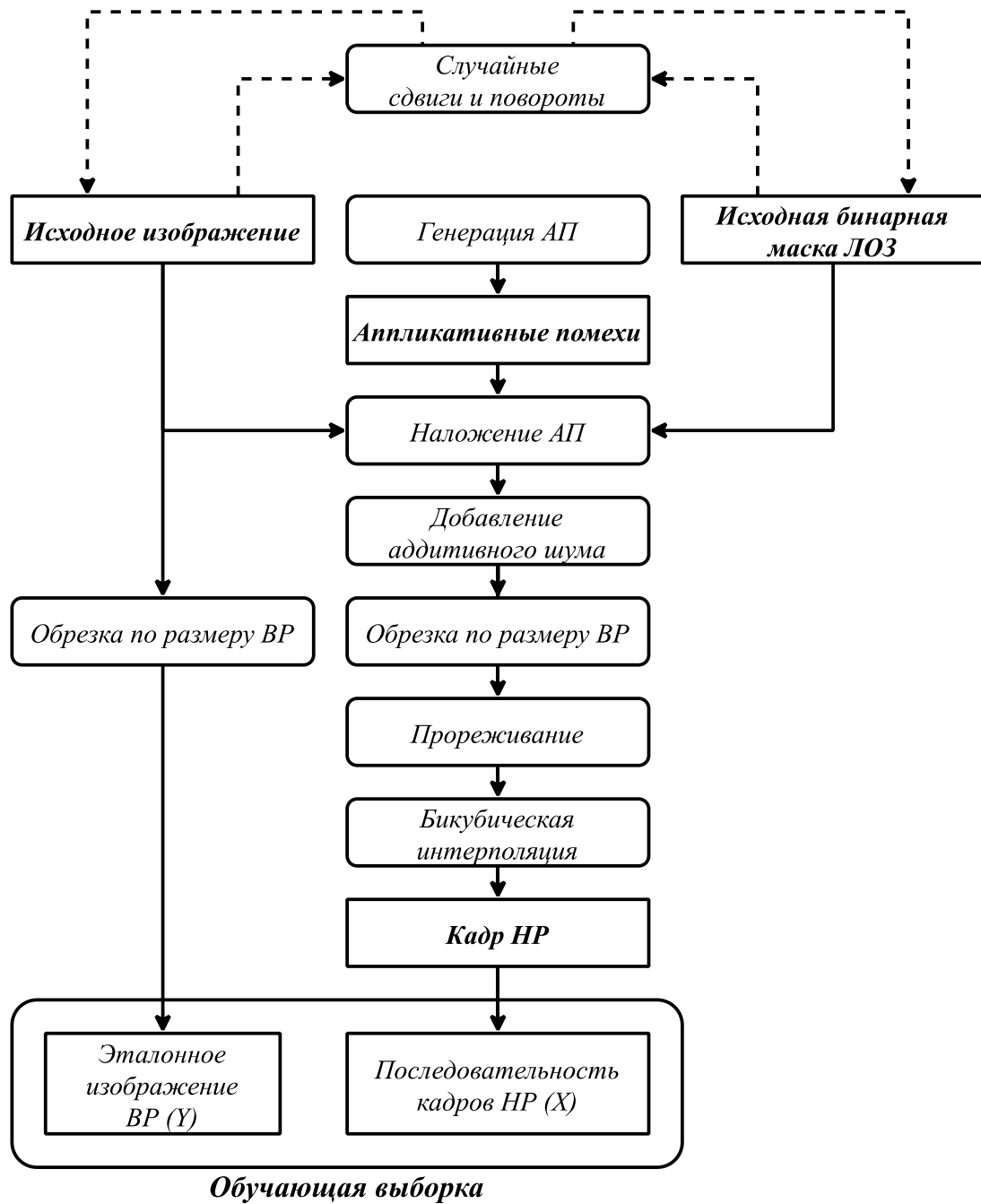


Рисунок 3.4 – Схема формирования обучающей выборки для обучения нейронной сети для обновления оценки ВР

Таким образом, были получены 40000 наборов цветных кадров размером  $64 \times 64$ , состоящих из 10 изображений. Каждому набору кадров соответствует единственное эталонное изображение  $64 \times 64$ , не подверженное воздействию АП. На рисунке 3.5 представлены пример последовательности кадров и соответствующее им эталонное изображений из обучающей выборки.

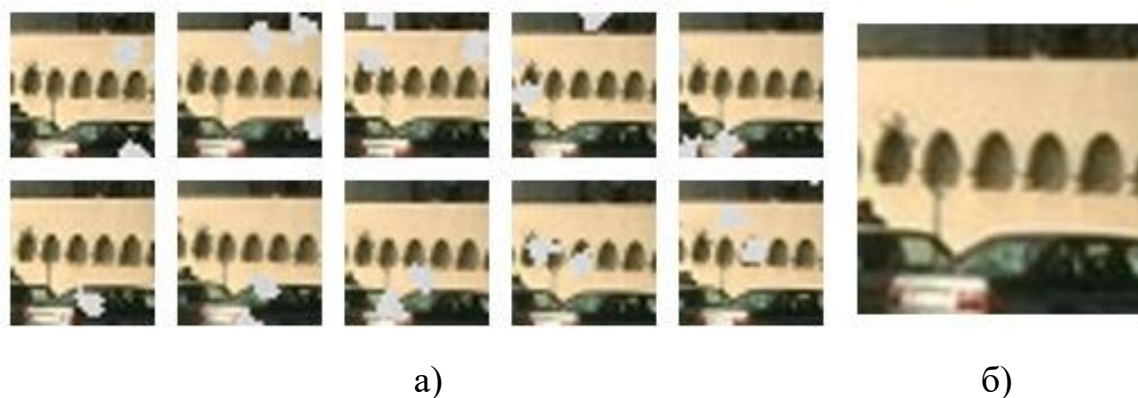


Рисунок 3.5 – Примеры изображений из обучающей выборки: а) входная последовательность изображений; б) эталонное изображение

Обучение нейронной сети для обновления оценки изображения ВР осуществлялось с использованием алгоритма Adam [94] в среде Matlab с постоянной скоростью обучения 0.001 и L2-регуляризацией весов. В качестве функции потерь использовалось значение среднеквадратичной ошибки между полученным и оригинальным изображениями ВР.

В рамках данного исследования использовалась предварительно обученная её авторами нейронная сеть LiteFlowNet2 для оценки оптического потока между кадрами НР.

**Результаты сравнения с другими алгоритмами.** Так как в главе 2, п. 2.4 было показано, что алгоритм, основанный на оптимальной фильтрации в адаптивной постановке (АОФА), показывает наибольшую точность восстановления изображения ВР, экспериментальное сравнение предложенного алгоритма, основанного на свёрточных нейронных сетях (АОСН), было проведено с этим алгоритмом, реализованным с учётом блочной обработки изображений.

Для тестирования предлагаемого алгоритма и его сравнения с альтернативными подходами были созданы 2 выборки, состоящие 100 последовательностей из 10 кадров размером  $128 \times 128$ . Обе этих выборки были сформированы таким же образом, как и обучающая выборка, но без шага бикубической интерполяции кадров НР. При этом изображения в одной из них

могли быть подвержены воздействию АП, а в другой все изображения являлись неискажёнными. В рамках экспериментального сравнения проводилось повышение разрешения в 2 раза (до  $256 \times 256$ ).

На рисунке 3.6 приведён пример оригинального изображения ВР ( $256 \times 256$ ) наряду с изображениями ВР, восстановленными при помощи обоих рассматриваемых алгоритмов (АОФА и АОСНН). Качественный анализ полученных изображений ВР показывает, что АОСНН справляется с устранением АП на входных изображениях результатами. Однако, в некоторых случаях области изображения, изначально подверженные воздействию АП, могут быть несколько искажены графическими артефактами, что несвойственно изображениям, полученным при помощи АОФА.

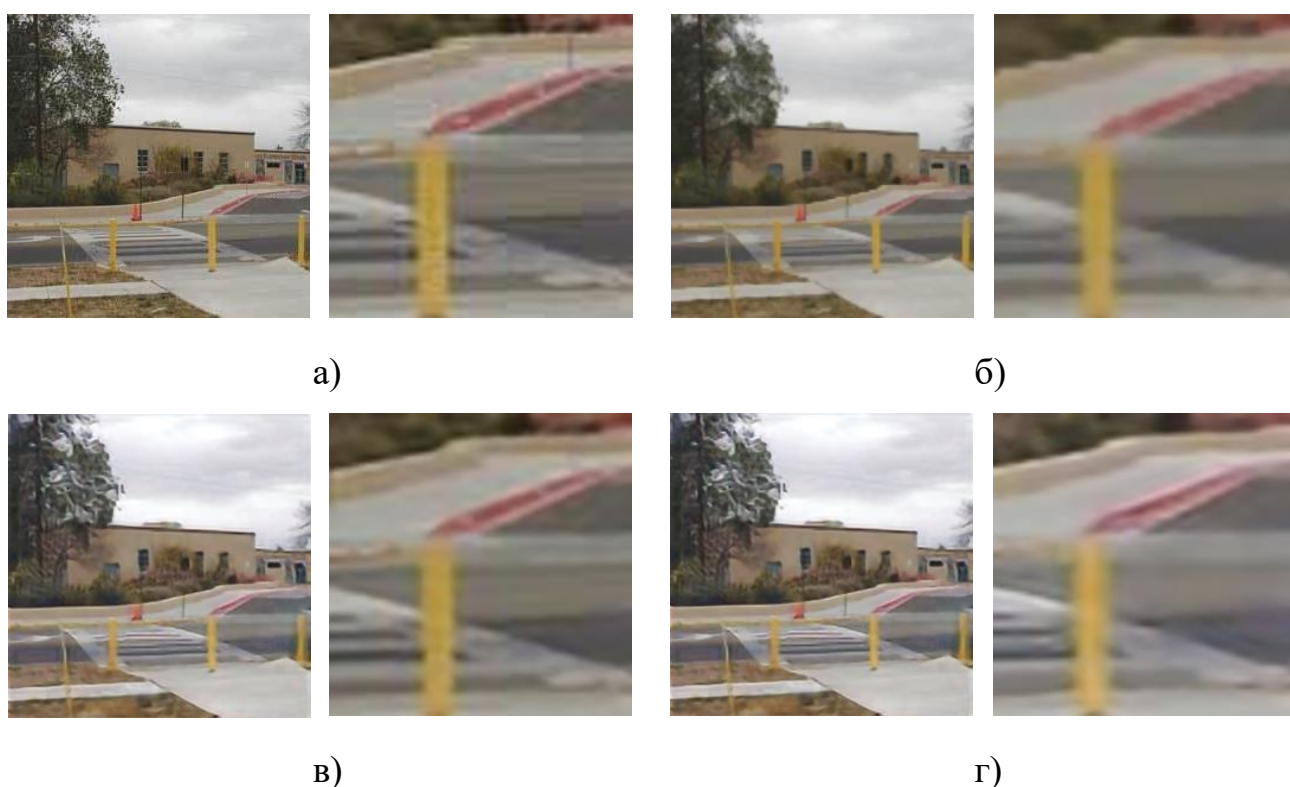


Рисунок 3.6 – Примеры изображений ВР и их увеличенные фрагменты:  
 а) оригинальное изображение ВР; б) усреднённая начальная оценка; в) результат алгоритма, основанного на оптимальной фильтрации в адаптивной постановке (АОФА); г) результат алгоритма, основанного на нейронных сетях без сегментации ЛОЗ (АСНН)



Численное сравнение результатов работы разных алгоритмов проводилось при помощи рассмотренных в главе 2, п. 2.4.1 критериев оценки сходства и качества изображений:

- пиковое отношение сигнала к шуму (PSNR; чем больше, тем лучше);
- индекс структурного сходства (SSIM; чем больше, тем лучше);
- индекс качества изображений (NIQE, чем меньше, тем лучше).

Усреднённые значения этих показателей, полученные в ходе эксперимента, приведены в таблицах 3.1 и 3.2. Их анализ подтверждает результаты качественного визуального сравнения формируемых изображений ВР: рассматриваемый комплексный алгоритм на основе свёрточных нейронных сетей (АСНН) имеет менее точные результаты восстановления исходных изображений по сравнению с АОФА даже в случае отсутствия АП на входных изображениях.

Таблица 3.1 – Точность восстановления исходного изображения ВР в случае наличия АП на входных изображениях

Критерий	АОФА	АСНС
<b>PSNR</b>	25,16	21,63
<b>SSIM</b>	0,93	0,85
<b>NIQE</b>	6,59	6,80

Таблица 3.2 – Точность восстановления исходного изображения ВР в случае отсутствия АП на входных изображениях

Критерий	АОФА	АСНС
<b>PSNR</b>	25,17	22,18
<b>SSIM</b>	0,93	0,87
<b>NIQE</b>	6,58	6,85

Результаты сравнения среднего времени обработки последовательности изображений, приведённые в таблице 3.3, показывают, что АСНН на порядок превосходит АОФА по скорости обработки. Данное преимущество достигается за

счет того, что предлагаемый подход не требует сложных матричных вычислений. Данные результаты были получены в среде Matlab 2020a на персональном компьютере со следующими системными характеристиками:

- процессор Ryzen 3600 3.6 ГГц;
- видеокарта GTX 1660 6 Гб;
- ОЗУ: 16 Гб.

Таблица 3.3 – Среднее время обработки последовательности изображений при помощи рассматриваемых алгоритмов

Критерий	АОФА	АСНС
<b>Время (сек)</b>	223	10

### **3.3 Алгоритм построения многокадрового сверхразрешения на основе глубоких нейронных сетей с сегментации областей, подверженных воздействию аппликативных помех**

#### **3.3.1 Схема обработки входной последовательности кадров**

В целях повышения качества результатов, получаемых при помощи алгоритма, рассмотренного в п. 3.2, в схему обработки изображений можно внести следующие изменения [66]:

- добавить дополнительный этап сегментации ЛОЗ на входных изображениях (рисунок 2.2) аналогично ранее рассмотренным в главе 2 алгоритмам на основе оптимальной фильтрации (при этом данный этап также может выполняться при помощи искусственной нейронной сети);

- рассмотреть другую архитектуру сети для уточнения оценки ВР с целью уменьшить искажения в областях, подверженных воздействию АП (ЛОЗ).

Для сегментации ЛОЗ на входных изображениях НР в рамках данного подхода предлагается использовать нейронную сеть для семантической сегментации изображений FCN-AlexNet [84], а для итеративного обновления оценки выходного изображения ВР – особую свёрточную нейронную сеть,

построенную в виде направленного ациклического графа (DAG-сеть). Регистрация с предобработкой поступающих кадров НР осуществляется также как и ранее в п. 3.2 на основе оптического потока, полученного при помощи LiteFlowNet2 [90].

Схема обработки последовательности кадров НР при помощи предлагаемого комплексного алгоритма представлена на рисунке 3.7. С учётом добавления этапа сегментации ЛОЗ она фактически эквивалентна схеме алгоритма, использующего оптимальную условно-линейную фильтрацию (рисунок 2.3), только все этапы итеративного уточнения оценки изображения ВР (регистрация, сегментация ЛОЗ и обновление текущей оценки) реализуются при помощи нейронных сетей.

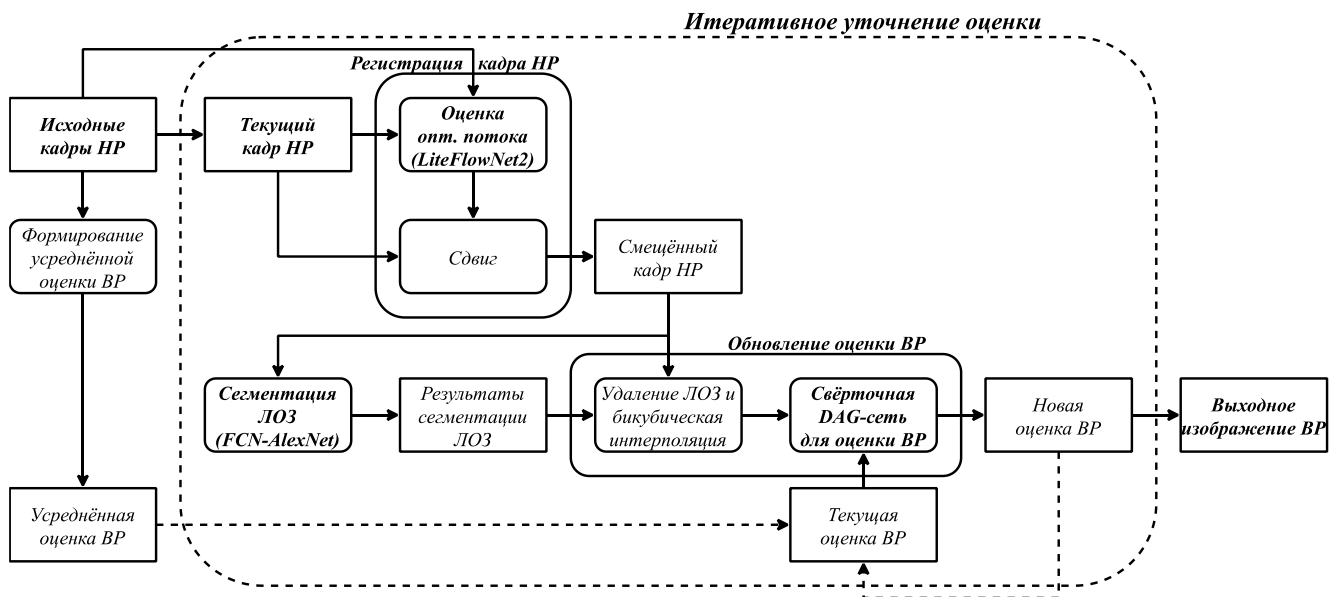


Рисунок 3.7 – Схема итеративного уточнения оценки изображения ВР при сегментации ЛОЗ на входных изображениях НР

### 3.3.2 Семантическая сегментация аппликативных помех при помощи нейронной сети

Для сегментации областей аномальных наблюдений на входных кадрах НР могут использоваться нейронные сети для семантической сегментации [83-86], классифицирующие объекты на изображениях. В общем случае, когда

отсутствуют априорные сведения о входных изображениях, пиксели могут разбиваться на два класса: полезные и ложные наблюдений. Однако, как и в случае использования двухэтапного алгоритма, рассмотренного в главе 2, п 2.3, разбиение пикселей на большее количество классов может потенциально улучшить точность выделения ЛОЗ.

В данной работе сегментация ЛОЗ осуществляется при помощи нейронной сети FCN-AlexNet [84], которая получается путём трансформации нейронной сети для классификации изображений AlexNet [91]. В ходе этого преобразования осуществляется замена выходных полносвязных слоёв AlexNet, а свёрточные слои, выделяющие признаки объектов, остаются без изменений. Архитектура FCN-AlexNet представлена на рисунке 3.8. На вход данной нейронной сети поступает изображение (рисунок 3.9а), а на выходе получается матрица такого же размера, что и входное изображение, в которой каждому пикселю изображения соответствует номер его класса. В случае использования двух классов, результат будет представлять из себя бинарную маску (рисунок. 3.9б), где 1 соответствует полезному наблюдению, а 0 – ложному.

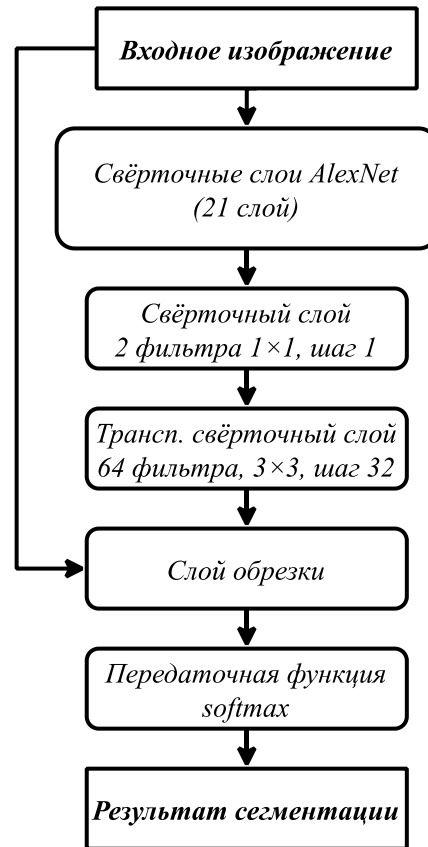


Рисунок 3.8 – Архитектура нейронной сети FCN-AlexNet для семантической сегментации на два класса

Так же, как и в случае алгоритмов, использующих оптимальную условно-линейную фильтрацию изображений, полученная в ходе сегментации и представленная в виде бинарной маски информация о расположении ЛОЗ используется в дальнейшем для минимизации воздействия АП на процесс обновления оценки изображения ВР. Этого можно добиться путём поэлементного умножения пикселей очередного кадра НР на полученную бинарную маску (рисунок 3.9в).



а)



б)



в)

Рисунок 3.9 – Сегментация ЛОЗ на изображении: а) изображение, подверженное воздействию АП; б) результат семантической сегментации изображения; в) изображение с удалёнными ЛОЗ (результат поэлементного умножения)

### **3.3.3 Итеративное уточнение оценки при помощи глубокой сети в виде ациклического направленного графа**

Итеративное уточнение текущей оценки ВР осуществляется при помощи глубокой сети, организованной в виде направленного ациклического графа (DAG-сеть), которая состоит из двух параллельных свёрточных сетей, выходные части которых объединяются при помощи слоя конкатенации по каналам. Архитектура предлагаемой DAG-сети представлена на рисунке. 3.10.

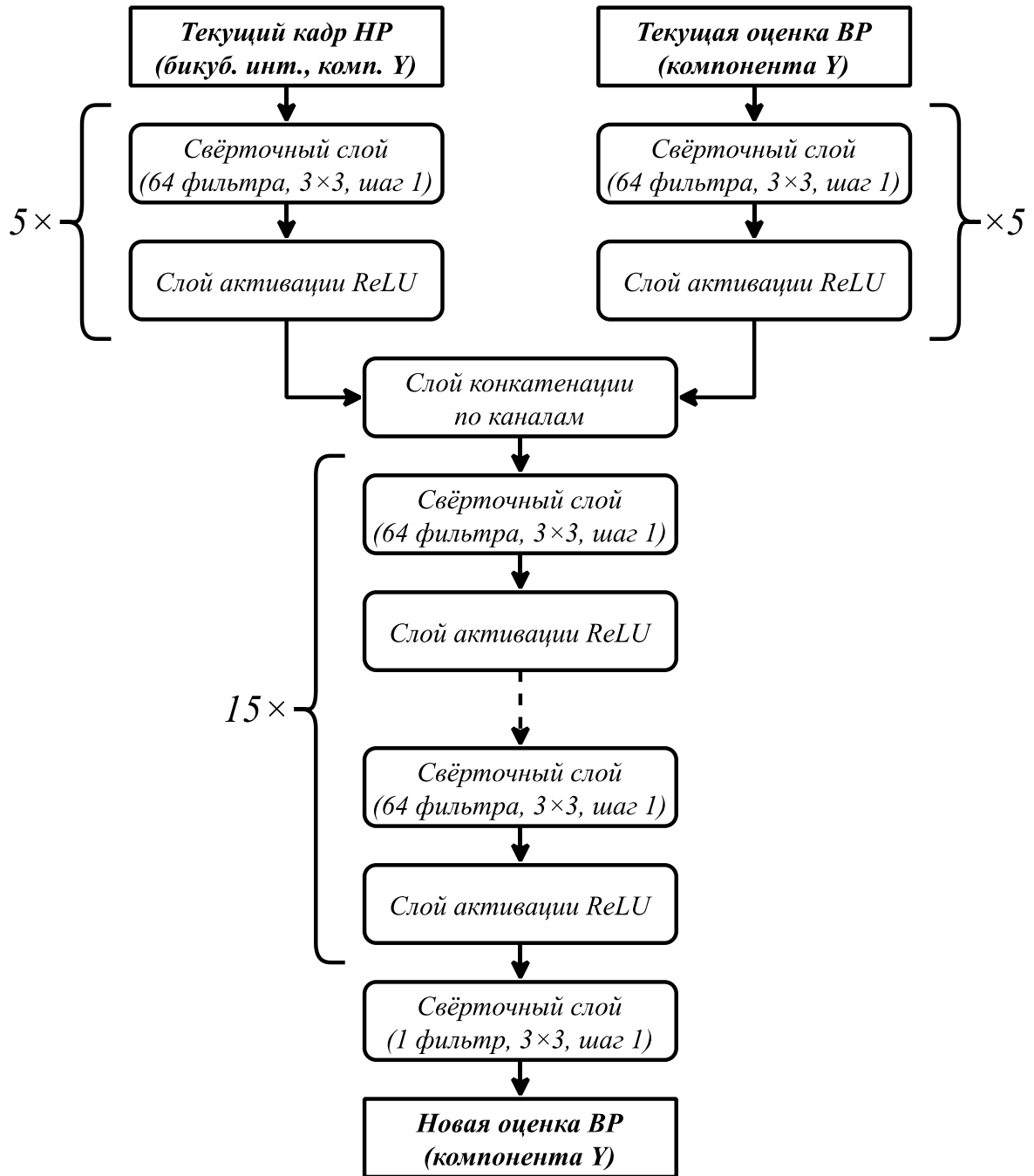


Рисунок 3.10 – Архитектура нейронной DAG-сети для обновления оценки изображения ВР

Как и ранее рассмотренная в 3.2.3 модификация сети SRNet, предлагаемая DAG-сеть принимает на вход очередной зарегистрированный кадр НР, увеличенный до ВР при помощи процедуры бикубической интерполяции, и текущую оценку изображения ВР, а на выходе получается изображение, соответствующее обновлённой оценке ВР. Однако с целью повышения точности получаемых результатов в процесс обработки были внесены некоторые изменения.

Во-первых, на входных кадрах НР области, подверженные воздействию АП были замещены нулевыми значениями (чёрным цветом), на основе результатов сегментации на предыдущем этапе обработки. Это было сделано с целью устранения воздействия аномальных наблюдений на итоговый результат.

Во-вторых, DAG-сеть не занимается обработкой цветных изображений напрямую. Вместо этого на её входы подаются цветовые компоненты яркости  $Y$  изображений, представленных в цветовом пространстве  $YCbCr$ . На выходе получается остаточное одноканальное изображение, которое необходимо сложить с компонентой яркости  $Y$  текущей оценки изображения ВР для её уточнения по результатам обработки нового наблюдения. Данный подход был предложен в [24] с целью повышения точности обработки.

С учётом данных изменений процесс уточнения оценки ВР состоит из следующих этапов, проиллюстрированных схемой на рисунке 3.11.

1. Удаление областей аномальных наблюдений при помощи поэлементного умножения пикселей очередного кадра НР на бинарную маску, полученную в ходе сегментации ЛОЗ.

2. Повышение разрешения полученного на шаге 1 изображения НР до целевого при помощи процедуры бикубической интерполяции.

3. Получение остаточного изображения ВР при помощи DAG-сети.

4. Сложение полученного остаточного изображения с компонентой яркости  $Y$  текущей оценки ВР для получения обновлённой оценки.



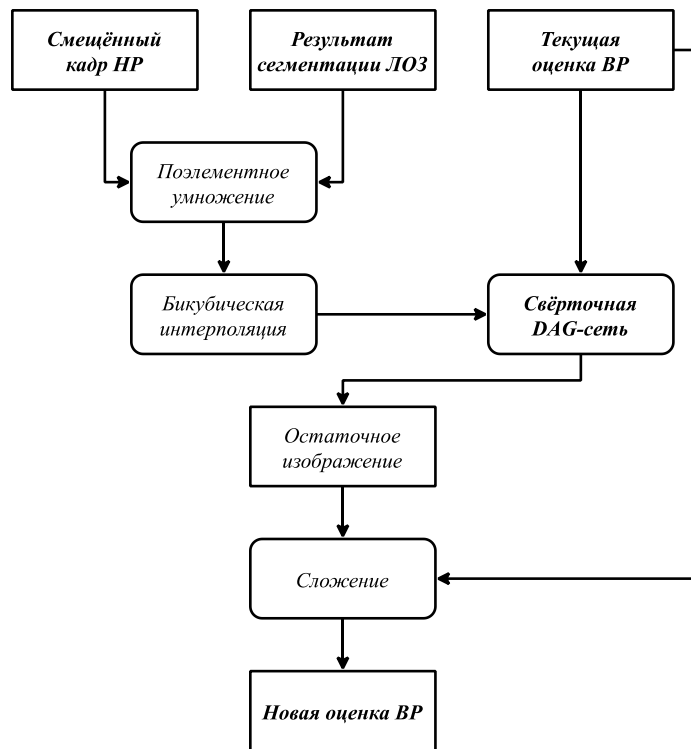


Рисунок 3.11 – Схема расчёта новой оценки изображения ВР

На рисунке 3.12 приведены примеры оценок выходного изображения ВР, полученных на различных этапах обработки последовательности кадров при помощи предложенной DAG-сети.



Рисунок 3.12 – Изображения на различных этапах обработки: а) исходная оценка ВР; б) оценка ВР после обработки первого НР; в) оценка ВР после обработки 10 кадров НР

### 3.3.4 Экспериментальное исследование

**Обучение используемых нейронных сетей.** Главным отличием от ранее рассмотренного подхода к формированию обучающей выборки являлась необходимость создания обучающей выборки для сети сегментации ЛОЗ FCN-AlexNet, состоящей из пар изображений, потенциально подверженных воздействию АП, и соответствующих им бинарных масок локализации ЛОЗ. Это можно осуществлять совместно с формированием обучающей выборки для нейронной сети, осуществляющей обновление оценки ВР, так как данный процесс включает в себя создание формирующей бинарной маски для добавления АП на изображения НР.

Как и при исследовании алгоритма без отдельного этапа сегментации ЛОЗ (п. 3.2), обучающие выборки для предложенной DAG-сети и FCN-AlexNet были сформированы на основе набора данных SUN [100] путём применения случайных аффинных преобразований к исходному изображению и формирующей бинарной маске ЛОЗ. При этом отдельные изображения НР или целые последовательности кадров не подвержены воздействию аномальных наблюдений. В рамках экспериментального исследования проводилось повышение разрешения в 2 раза.

Обновлённая схема совместного формирования обучающих выборок представлена на рисунке 3.13.

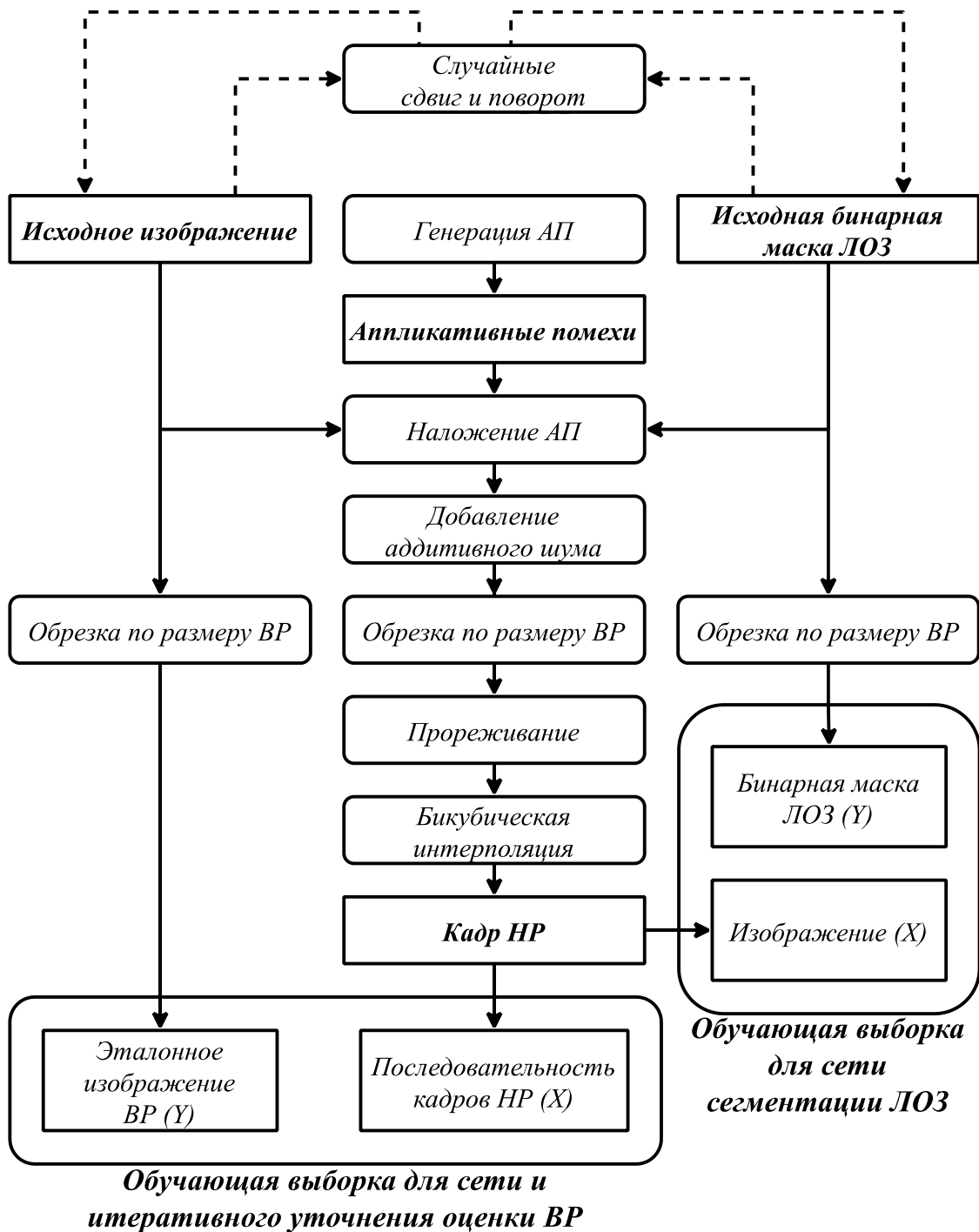


Рисунок 3.13 – Схема формирования обучающих выборок для обучения нейронных сетей обновления оценки ВР и сегментации ЛОЗ

Для обучения DAG-сети были получены 40000 наборов цветных кадров размером  $64 \times 64$ , состоящих из 10 входных изображений и одного эталонного изображения, а для обучения FCN-AlexNet – соответствующие этим наборам

кадров 400000 пар изображений и бинарных масок, примеры которых представлены на рисунке 3.14.

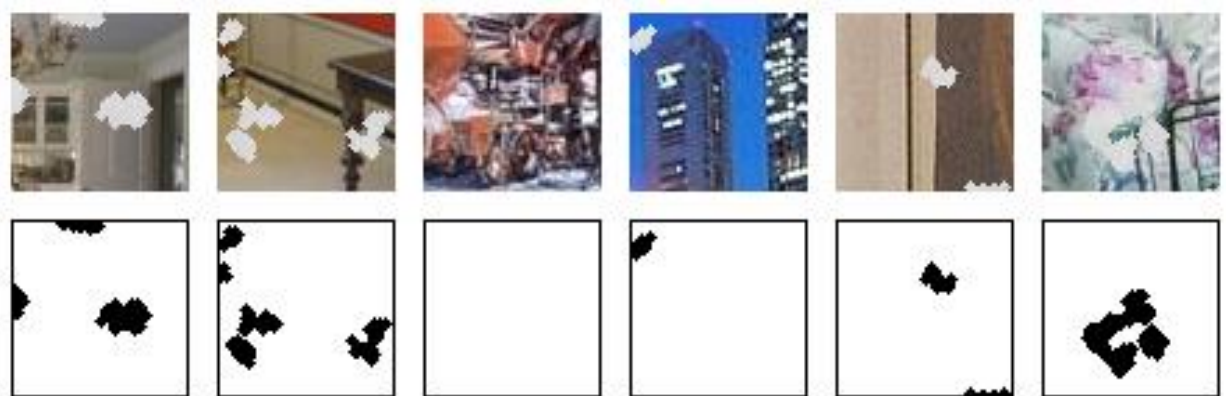


Рисунок 3.14 – Примеры пар изображений из обучающей выборки для обучения сети сегментации ЛОЗ

Обучение DAG-сети для обновления оценки изображения ВР проводилось с использованием алгоритма Adam [94] в среде Matlab с постоянной скоростью обучения 0,0001 и L2-регуляризацией весов. В качестве функции потерь использовалось значение среднеквадратичной ошибки.

Так как сеть FCN-AlexNet получается путём трансформации из сети AlexNet, для её обучения можно использовать уже обученные слои AlexNet, входящие в дистрибутив среды Matlab. Процесс дообучения FCN-AlexNet осуществлялся при помощи алгоритма стохастического градиентного спуска с импульсом (SGDM) [84] с начальной скоростью обучения 0,001, импульсом 0,9 и L2-регуляризацией весов.

Как и ранее, для оценки оптического потока между кадрами НР в процессе регистрации новых наблюдения использовалась предварительно обученная её авторами LiteFlowNet2.

**Результаты сравнения с другими алгоритмами.** Для сравнения предлагаемого комплексного алгоритма на основе свёрточные нейронных сетей и сегментации ЛОЗ на входных изображениях (АЧСС) с альтернативными подходами были сформированы 2 тестовых выборки, состоящие из 100 наборов

из 10 кадров размером  $128 \times 128$ . Данные выборки были созданы по такому же алгоритму, как и обучающая выборка для DAG-сети. При этом изображения в одной выборке были подвержены воздействию АП, а в другой были неискажёнными.

В рамках экспериментального исследования было проведено сравнение предложенного комплексного алгоритма следующими ранее рассмотренными алгоритмами, которые показали наибольшую точность и быстродействие в предыдущих экспериментальных сравнениях:

- алгоритм, основанный на оптимальной фильтрации в адаптивной постановке (АОФА), как дающий наиболее точные результаты;
- алгоритм, основанный на использовании свёрточных нейронных сетей без сегментации областей, искажённых АП (АСНС), как обладающий наилучшим быстродействием.

Разрешение тестовых изображений повышалось в ходе эксперимента в 2 раза. На рисунке 3.15 приведён пример эталонного изображения ВР ( $256 \times 256$ ) с изображениями ВР, восстановленными при помощи рассматриваемых алгоритмов. Визуальный анализ полученных изображений ВР показывает, что АСНСС выполняет устранение АП и одновременное восстановление детализации изображений на одном уровне с результатами, полученными при помощи АОФА, и существенно лучше, чем АСНС.

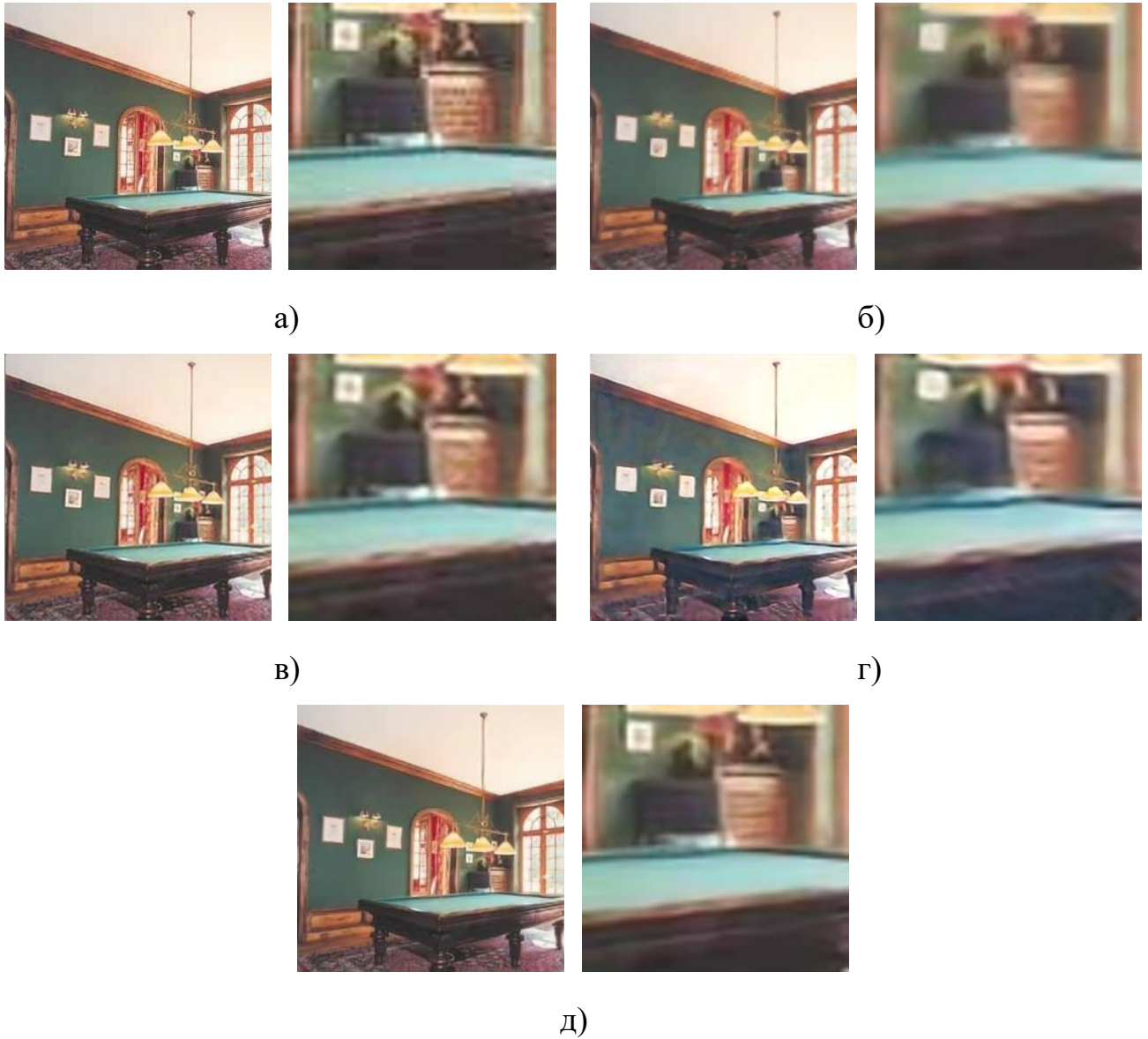


Рисунок 3.15 – Примеры изображений ВР и их увеличенные фрагменты:  
 а) оригинальное изображение ВР; б) усреднённая начальная оценка; в) результат АОФА; г) результат АСНС; д) результат АСНСС

Численное сравнение полученных результатов, проводилось на основе ранее использовавшихся показателей:

- пиковое отношение сигнала к шуму (PSNR; чем больше, тем лучше);
- индекс структурного сходства (SSIM; чем больше, тем лучше);
- индекс качества изображений (NIQE, чем меньше, тем лучше).

Усреднённые значения этих критериев, полученные в ходе эксперимента, приведены для разных наборов тестовых изображений в таблицах 3.4 и 3.5.

Таблица 3.4 – Точность восстановления исходного изображения ВР в случае наличия АП на входных изображениях

Критерий	АОФА	АСНС	АСНСС
PSNR	25,16	21,63	29,11
SSIM	0,93	0,85	0,97
NIQE	6,58	6,80	6,66

Таблица 3.5 – Точность восстановления исходного изображения ВР в случае отсутствия АП на входных изображениях

Критерий	АОФА	АСНС	АСНСС
PSNR	25,17	22,18	30,38
SSIM	0,93	0,87	0,98
NIQE	6,52	6,85	6,74

Также было проведено сравнение среднего времени обработки последовательности кадров при помощи рассматриваемых алгоритмов, результаты которого приведены в таблице 3.6. Вычисления проводились в среде Matlab 2020a на персональном компьютере со следующими системными характеристиками:

- процессор Ryzen 3600 3.6 ГГц;
- видеокарта GTX 1660 6 Гб;
- ОЗУ: 16 Гб.

Таблица 3.6 – Среднее время обработки изображения при помощи рассматриваемых алгоритмов

Критерий	АОФА	АСНС	АСНСС
Время (сек)	223	10	9

Анализ полученных результатов подтверждает результаты визуального сравнения восстановленных изображений ВР: предложенный в данном разделе комплексный алгоритм многокадрового СР АСНСС восстанавливает изображения

ВР с большей точностью по всем введённым показателям, чем АОФА, имея существенно лучшее быстродействие (на одном уровне с АСНСС). При этом данное преимущество в точности проявляется как при наличии АП на входных изображениях НР, так и при обработке неискажённых изображений.

### **Выводы по главе**

1. Проанализированы проблемы традиционных подходов к построению многокадрового СР в условиях АП, опирающиеся на использование математических моделей. Для решения данных проблем предложено использовать методы машинного обучения, включая глубокие свёрточные нейронные сети. Предложены два подхода к синтезу алгоритмов, реализующих итеративное уточнение оценки изображения ВР и опирающихся на использование нейросетевых технологий: без реализации отдельного этапа сегментации АП и применяющий его.

2. Предложен комплексный алгоритм многокадрового СР в условиях АП, реализующий общую схему итеративного уточнения оценки ВР и использующий искусственные нейронные сети для регистрации изображений НР и улучшения их качества. Рассмотрена архитектура глубокой свёрточной сети для повышения качества изображений с остаточными связями между слоями. Проведено экспериментальное исследование синтезированного алгоритма, в рамках которого создана обучающая выборка и проведено обучение предложенной нейронной сети. Анализ результатов эксперимента показывает, что рассматриваемый алгоритм работает на порядок быстрее традиционных алгоритмов и справляется с устранением АП на входных изображениях, однако точность получаемых результатов не всегда является удовлетворительной.

3. Предложен комплексный алгоритм многокадрового СР, реализующий общую схему итеративного уточнения оценки ВР с сегментацией АП на поступающей последовательности изображений НР и использующий свёрточные нейронные сети на всех этапах обработки изображений, включая сегментацию



ЛОЗ. Для сегментации областей, подверженных воздействию АП, предложено использовать сеть семантической сегментации. Предложена новая архитектура свёрточной нейронной сети для уточнения оценки изображения ВР, представленной в виде направленного ациклического графа (DAG-сеть). В ходе экспериментального исследования проведено обучение нейронных сетей для семантической сегментации ЛОЗ и уточнения текущей оценки изображения ВР. Сравнение результатов, полученных при помощи синтезированного комплексного алгоритма, с результатами альтернативных алгоритмов показывает его превосходство как в плане точности получаемых результатов, так и в плане быстродействия.

## **4 Синтез и анализ алгоритмов построения сверхразрешения видеоданных**

### **4.1 Общая схема обработки видеопоследовательности с использованием алгоритмов многокадрового сверхразрешения**

Одним из способов решения задачи построения СР в случае обработки видеопоследовательности является построение многокадрового СР для каждого из исходных кадров НР. Для этого обработка осуществляется методом «скользящего окна», при котором происходит обработка очередного кадра НР, а также некоторой его окрестности по времени, включающей в себя  $m$  кадров до и  $m$  кадров после него (рисунок 1.2). Таким образом, для решения данной задачи потенциально могут использоваться любые алгоритмы многокадрового СР, включая рассмотренные ранее.

В случае, если кадры исходной видеопоследовательности могут быть подвержены воздействию АП, то представляется целесообразным использование одного из алгоритмов многокадрового СР, осуществляющих устранение негативного воздействия АП на качество получаемых изображений ВР с одновременным повышением разрешения исходных изображений.

С учётом вышеизложенных соображений общий алгоритм для обработки видеоданных в интересах построения СР и устранения воздействия АП принимает следующий вид, представленный на рисунке 4.1.

1. Оценка сдвигов между кадрами НР входной видеопоследовательности.
2. Сегментация ЛОЗ на кадрах НР входной видеопоследовательности.
3. Инициализация обработки первого кадра видеопоследовательности.
4. Обработка текущего кадра в целях повышения его разрешения и устранения воздействия АП с учётом предшествующих и предыдущих  $2m$  кадров в скользящем временном окне.
5. Переход к следующему кадру НР, если не все кадры видеопоследовательности были обработаны.

При этом для оценки межкадровых сдвигов и сегментации областей аномальных наблюдений могут использоваться произвольные алгоритмы, включая описанные ранее в главах 2 и 3.

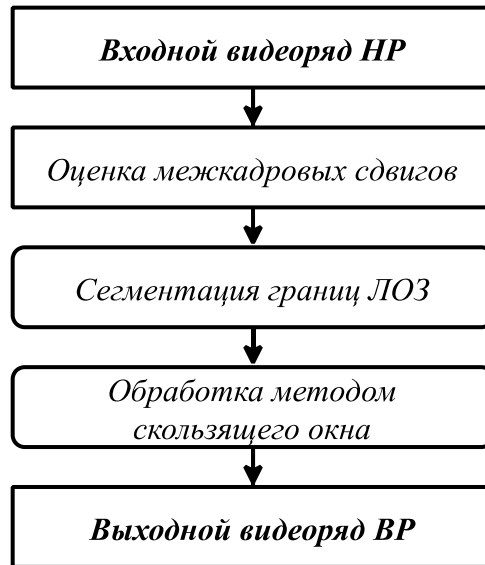


Рисунок 4.1 – Схема обработки входной последовательности изображений НР на основе свёрточных нейронных сетей с целью построения СР для видеопоследовательности

Далее будут рассмотрены два подхода к реализации такой схемы обработки [67].

Первый подход основан на прямом использовании уже рассмотренной выше в п. 3.3.3 DAG-сети для повышения качества изображений в составе алгоритма многокадрового СР АСНСС.

Второй подход основан на применении новой, специально модифицированной с учётом особенностей обработки видеоданных DAG-сети.

## **4.2 Синтез алгоритма построения сверхразрешения видеоданных на основе алгоритма многокадрового сверхразрешения с использованием глубоких нейронных сетей**

### **4.2.1 Схема обработки входной видеопоследовательности**

Применение DAG-сети, синтезированной для ранее описанного в п. 3.3 алгоритма многокадрового СР АСНСС, в ходе обработки текущего положения скользящего окна и получения очередного кадра ВР, предполагает проведение следующих операций:

- наложение всех окрестных кадров на текущий обрабатываемый кадр на основе информации о межкадровых сдвигах;
- замещение областей, подверженных воздействию АП, нулевыми значениями на основе результатов сегментации ЛОЗ;
- перемещение текущего кадра НР в начало последовательности для соответствия процессу обработке в АСНСС;
- вычисление начальной усреднённой оценки кадра ВР, которая будет итеративно уточняться при помощи DAG-сети.

После подготовки входных данных, осуществляется итеративный процесс уточнения текущей оценки выходного кадра ВР при помощи DAG-сети, принимающей на вход текущую оценку и очередной смещённый кадр НР.

На рисунке 4.2 представлена схема обработки текущего кадра с учётом вышеизложенных соображений.

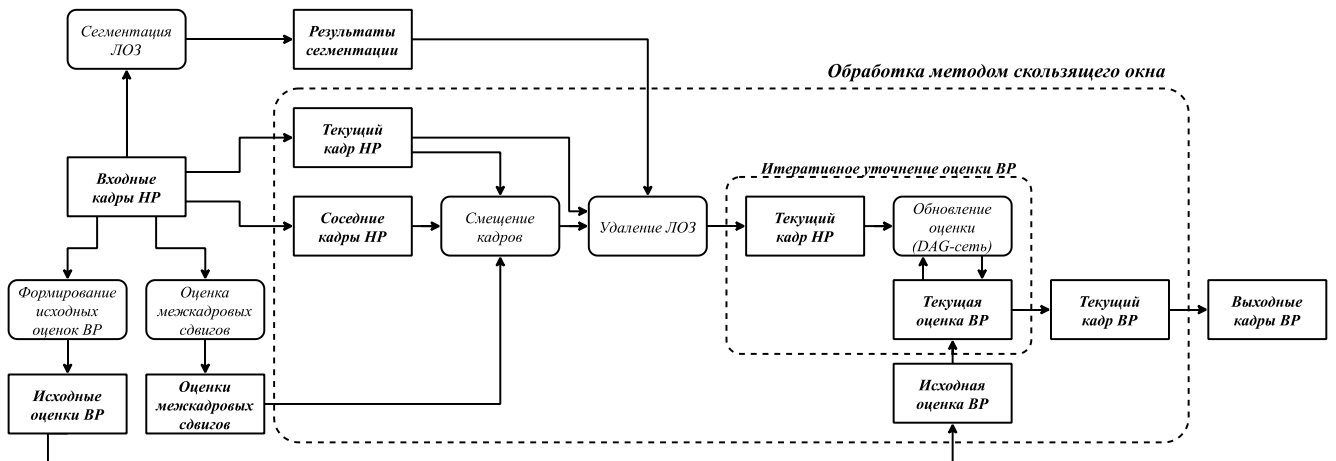


Рисунок 4.2 – Схема обработки видеопоследовательности при помощи DAG-сети для повышения качества изображения в рамках многокадрового СР

#### 4.2.2 Экспериментальное исследование

Так как алгоритм АСНСС позволяет обрабатывать последовательности изображений произвольной длины, представляется целесообразным исследование зависимости точности получаемых результатов от размера скользящего окна  $m$ .

Для экспериментального исследования рассматриваемого алгоритма СР видеоданных (АСВ1) на основе DAG-сети из алгоритма АСНСС были сформированы 2 тестовых выборки, состоящие из 50 видеопоследовательностей, включающих в себя 48 кадров разрешением  $128 \times 128$ . Изображения в первой выборке были подвержены воздействию АП, в отличие от второй. Обе выборки были созданы по такому же алгоритму, как и обучающая выборка для DAG-сети (п. 3.3.4). Примеры кадров НР представлены на рисунке 4.3.



Рисунок 4.3 – Примеры кадров НР из тестовых выборок

Разрешение тестовых изображений повышалось в ходе эксперимента в 2 раза, а параметр  $m$  принимал значение от 1 до 4. Например, значение  $m=2$  описывает ситуацию, когда для получения очередного кадра ВР видеопоследовательности дополнительно используются 2 предшествующих ему и 2 последующих ему кадра НР. На рисунке 4.4 приведены примеры эталонного изображения ВР ( $256 \times 256$ ) с изображениями ВР, восстановленными при помощи предложенного алгоритма с различными значениями параметра  $m$ .

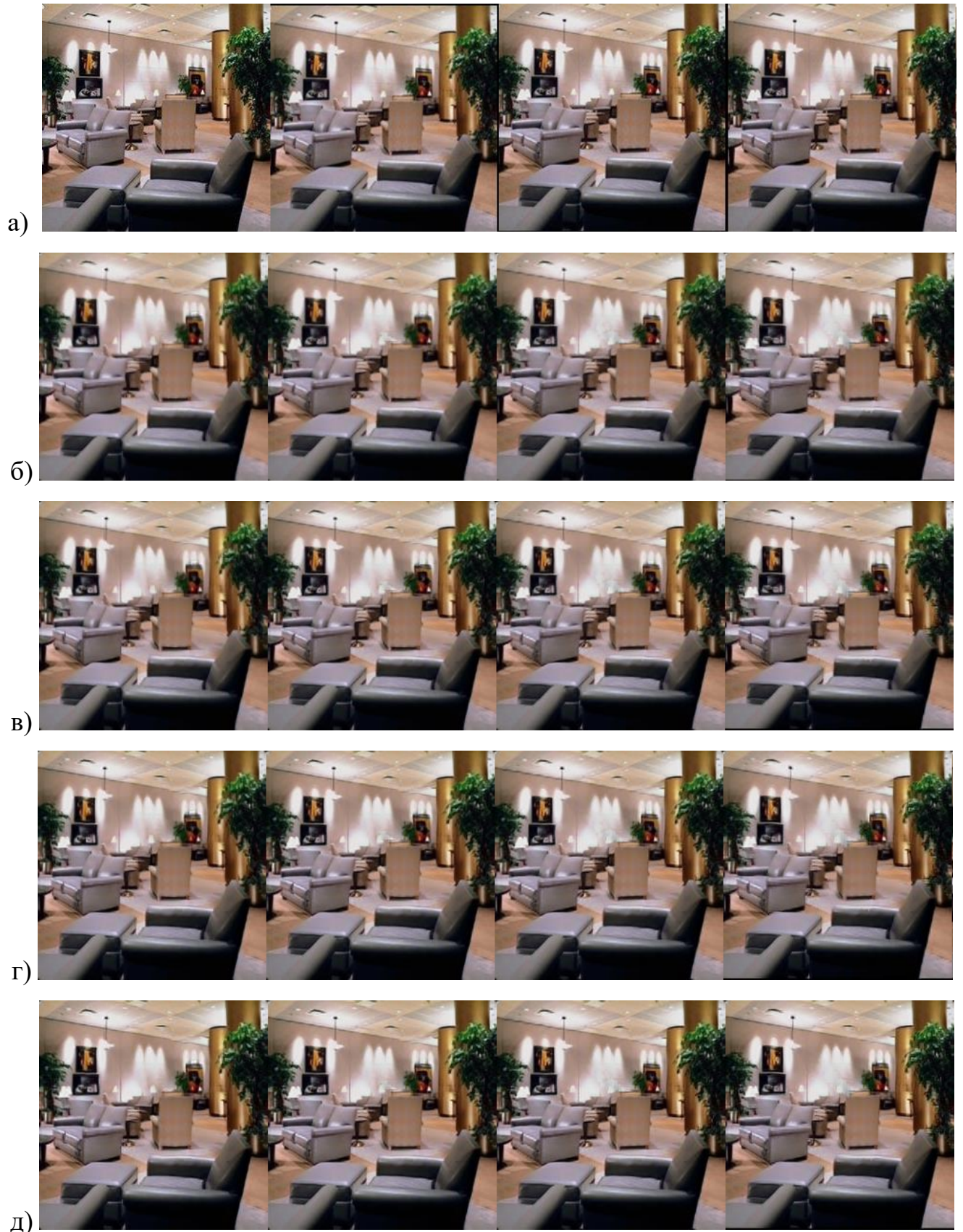


Рисунок 4.4 – Примеры кадров ВР: а) эталонные кадры ВР; б) результат при  $m = 1$ ; в) результат при  $m = 2$ ; г) результат при  $m = 3$ ; д) результат при  $m = 4$

Численное сравнение полученных результатов, проводилось на основе уже использовавшихся ранее в главах 2, 3 показателей:

- пиковое отношение сигнала к шуму (PSNR; чем больше, тем лучше);
- индекс структурного сходства (SSIM; чем больше, тем лучше);
- индекс качества изображений (NIQE, чем меньше, тем лучше).

Усреднённые значения этих показателей, полученные в ходе экспериментального исследования, для разных наборов тестовых изображений и значений параметра  $m$  приведены в таблицах 4.1 и 4.2.

Таблица 4.1 – Точность восстановления исходных кадров ВР в случае наличия АП на входной видеопоследовательности

Критерий	$m = 1$	$m = 2$	$m = 3$	$m = 4$
<b>PSNR</b>	20,96	21,71	22,05	22,93
<b>SSIM</b>	0,88	0,90	0,90	0,91
<b>NIQE</b>	5,88	5,46	5,33	5,31

Таблица 4.2 – Точность восстановления исходных кадров ВР в случае отсутствия АП на входной видеопоследовательности

Критерий	$m = 1$	$m = 2$	$m = 3$	$m = 4$
<b>PSNR</b>	24,35	24,71	25,12	25,81
<b>SSIM</b>	0,93	0,94	0,95	0,95
<b>NIQE</b>	5,72	5,47	5,30	5,82

Также было проведено сравнение среднего времени обработки одного кадра при разных значениях параметра  $m$ . Вычисления проводились в среде Matlab 2020a на персональном компьютере со следующими системными характеристиками:

- процессор Ryzen 3600 3.6 ГГц;
- видеокарта GTX 1660 6 Гб;
- ОЗУ: 16 Гб.



Таблица 4.3 – Среднее время обработки одного кадра в зависимости от размера окрестности  $m$

Критерий	$m = 1$	$m = 2$	$m = 3$	$m = 4$
<b>Время (сек)</b>	3,11	4,61	6,08	8,01

Анализ полученных результатов показывает, что увеличение размера скользящего окна  $m$  ожидаемо повышает качество получаемых результатов. Однако, повышенная точность обработки приводит к большему времени обработки, так как рассмотрение каждого дополнительного кадра в окрестности линейно увеличивает время обработки.

### **4.3 Синтез алгоритма построения сверхразрешения видеоданных на основе нейронной сети в виде направленного ациклического графа**

#### **4.3.1 Схема обработки входной видеопоследовательности кадров**

Одной из особенностей рассмотренного в п. 3.3 алгоритма многокадрового СР (АСНСС) и используемой им DAG-сети является их независимость от числа изображений НР в исходной последовательности. Однако при обработке видеопоследовательностей это обстоятельство можно отнести к недостаткам, так как обработка каждого дополнительного кадра требует использования DAG-сети. Несмотря на высокую точность результатов, получаемых при обработке видеопоследовательностей с помощью алгоритма АСВ1, рассмотренного в п. 4.2, данный процесс занимает большое количество времени, что связано с необходимостью использовать DAG-сеть  $2m + 1$  раз для обработки одного кадра НР.

Потенциальным улучшением такой схемы обработки является модификация данной нейронной сети с целью увеличения количества её входов с 2 до  $2m + 2$ , что позволит передавать ей на вход не только оценку изображения ВР и текущий кадр НР, но и остальные кадры НР, обрабатываемые при текущим положением

скользящего окна и включающие в себя  $t$  предшествующих и  $t$  последующих кадров. Архитектура такой DAG-сети представлена на рисунке 4.5.

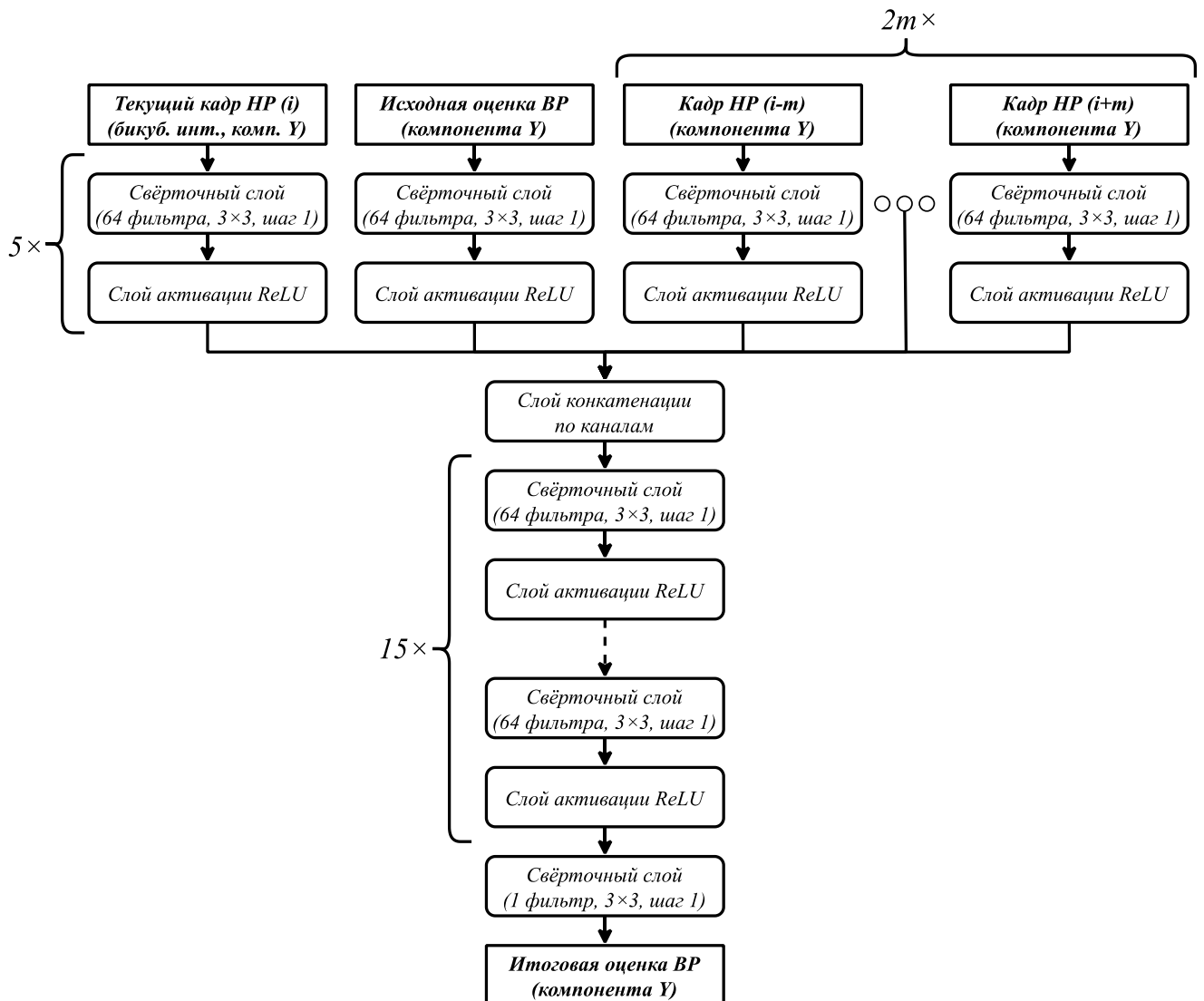


Рисунок 4.5 – Архитектура нейронной DAG-сети для получения кадра ВР на основе соответствующего кадра НР, кадров НР, лежащих в его окрестности, и априорной оценки ВР

Как и в алгоритме многокадрового СР, данная DAG-сеть обрабатывает цветные компоненты  $Y$  изображений в цветовом пространстве  $YCbCr$ , разрешение изображений НР повышается до целевого при помощи процедуры бикубической интерполяции, а обнаруженные области ЛОЗ замещаются нулевыми значениями.

В случае обработки кадров, находящихся в начале или конце видеопоследовательности, количество изображений до или после текущего может быть меньше требуемого количества ( $m$ ). В этом случае недостающие изображения на входе нейронной сети замещаются нулевыми матрицами.

Как и ранее в качестве исходной оценки изображения ВР используется усреднённое изображение, полученное из всех кадров, захваченных скользящим окном.

Схема обработки текущего кадра НР с учётом вышеизложенных соображений представлена на рисунке 4.6. Её ключевым отличием от схемы ранее рассмотренного алгоритма (рисунок 4.2) является отсутствия этапа итеративного уточнения текущей оценки кадра ВР. Вместо этого данная оценка получается после одного использования модифицированной DAG-сети.

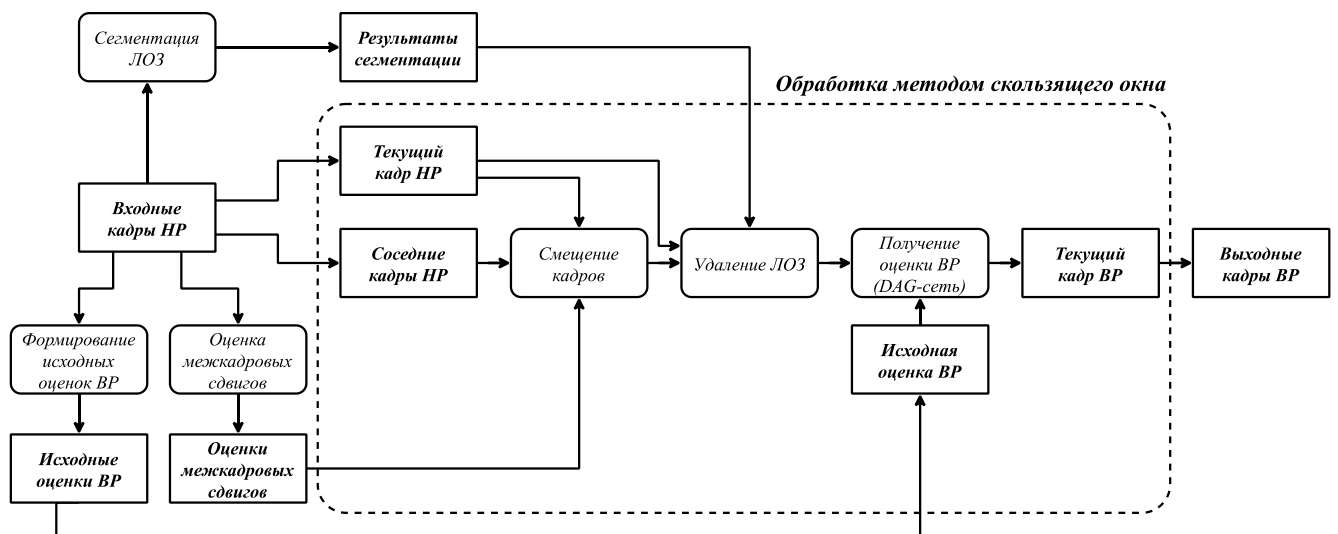


Рисунок 4.6 – Схема обработки видеопоследовательности при помощи модифицированной DAG-сети

### 4.3.2 Экспериментальное исследование

**Обучение используемой DAG-сети.** Как и ранее при исследовании рассмотренных в главе 3 алгоритмов многокадрового СР, обучающая выборка для модифицированной DAG-сети была сформирована на основе изображений из набора SUN [100] путём применения случайных аффинных преобразований к

некоторому исходному изображению ВР и формирующей бинарной маске ЛОЗ. При этом отдельные изображения НР или целые последовательности кадров не были подвержены воздействию АП. В рамках экспериментального исследования проводилось повышение разрешения в 2 раза, а размер скользящего окна  $m$  был фиксированным и равным 3.

Основное отличие процесса создания обучающей выборки от рассмотренных в главе 3 схем заключается в расширении объёма эталонных изображений, так как теперь каждому входному кадру НР соответствует свой выходной образец. Обновлённая с учётом данного требования схема формирования обучающей выборки представлена рисунке 4.7.

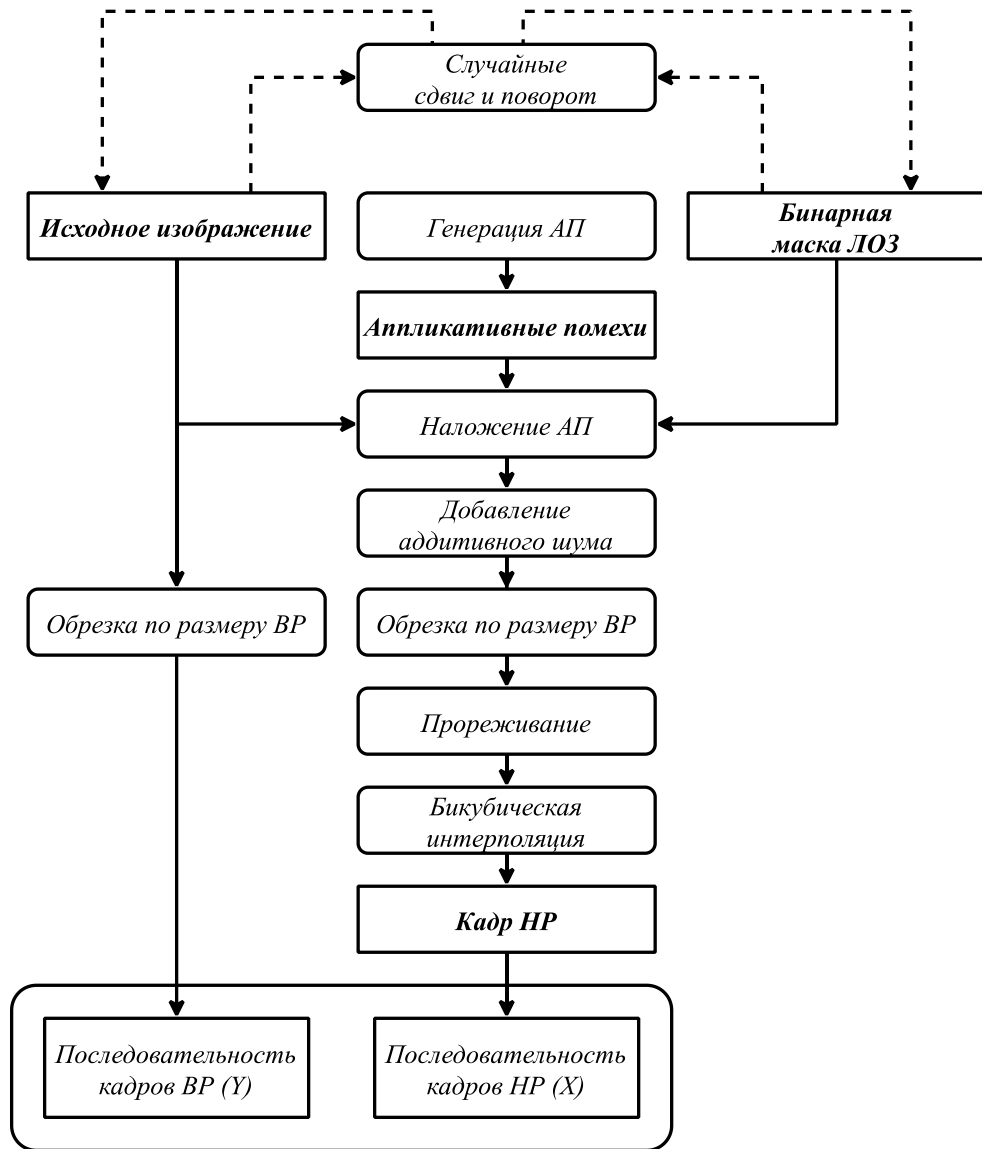


Рисунок 4.7 – Схема формирования обучающих выборок для обучения нейронной сети для повышения качества кадра

Для обучения данной DAG-сети были получены 10000 наборов цветных кадров размером  $64 \times 64$ , состоящих из 48 пар входных и эталонных изображений. Примеры изображения из обучающей выборки представлены на рисунке 4.8.



Рисунок 4.8 – Примеры изображений из обучающей выборки: сверху – кадры НР; снизу – эталонные кадры ВР

**Результаты сравнения с другими алгоритмами.** Для сравнения предлагаемого комплексного алгоритма СР видеоданных (АСВ2) с ранее рассмотренным алгоритмом АСВ1 использовались те же тестовые выборки видеоданных разрешением  $128 \times 128$ , что и в п. 4.2.2.

Разрешение тестовых изображений повышалось в ходе эксперимента в 2 раза. На рисунке 4.9 приведены примеры эталонных кадров ВР ( $256 \times 256$ ) с кадрами ВР, полученными при помощи рассматриваемых алгоритмов АСВ1 и АСВ2. Качественное сравнение полученных после обработки видеоданных показывает, что алгоритм АСВ2, обрабатывающий всю окрестность текущего кадра за один проход DAG-сети, справляется с задачей устранения АП на изображениях, однако получаемые с его помощью изображения имеют меньший уровень детализации по сравнению с результатами ранее рассмотренного алгоритма АСВ1.

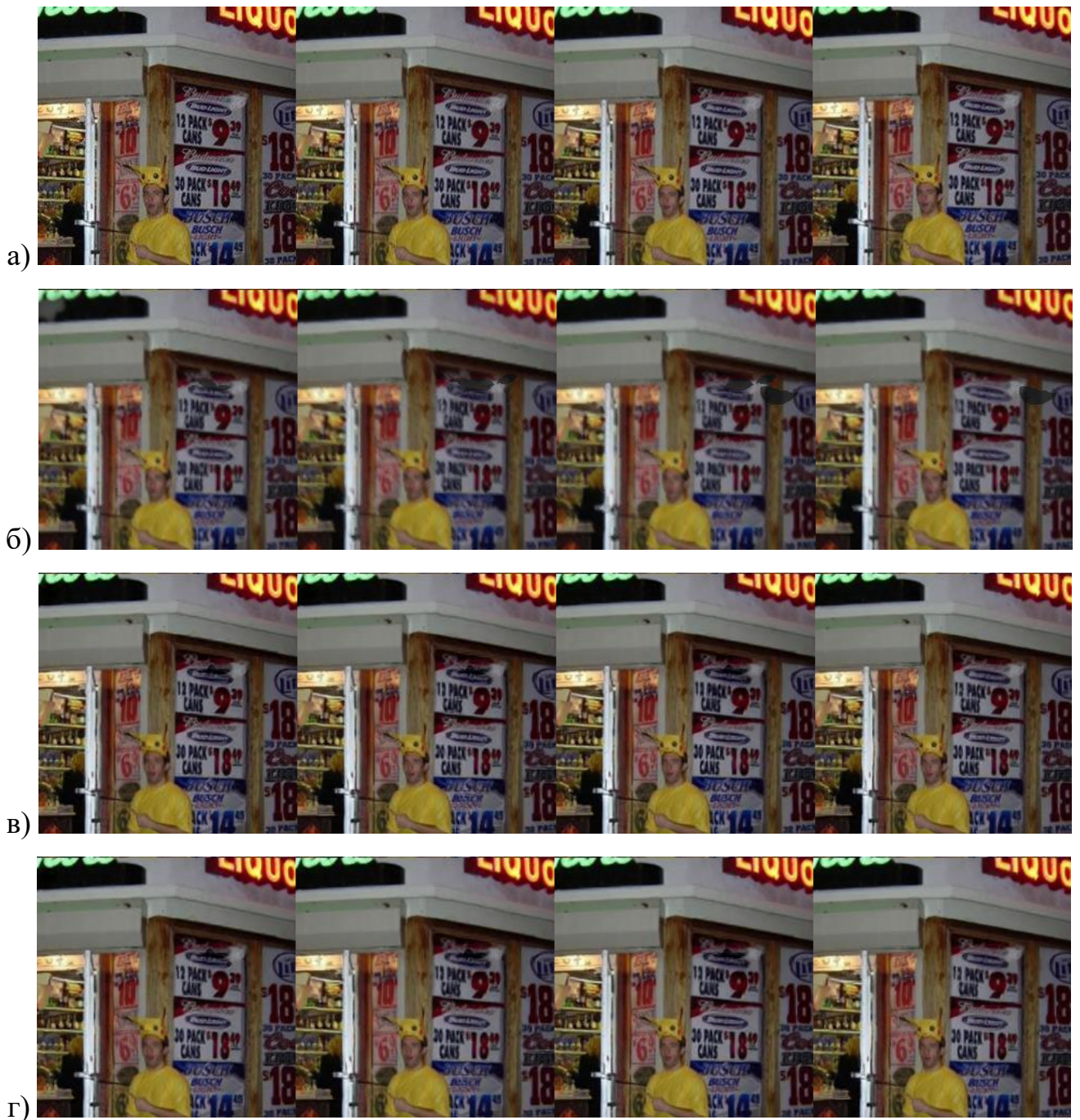


Рисунок 4.9 – Примеры кадров ВР: а) исходные кадры ВР; б) результат алгоритма АСВ2 ( $m = 3$ ); в) результат алгоритма АСВ1 ( $m = 3$ ); г) результат алгоритма АСВ1 ( $m = 4$ )

Численное сравнение полученных результатов, проводилось на основе ранее использовавшихся показателей PSNR, SSIM и NIQE. Их усреднённые значения, полученные в ходе эксперимента, приведены для разных наборов тестовых изображений в таблицах 4.4 и 4.5.

Таблица 4.4 – Точность восстановления исходного кадра ВР в случае наличия АП на входной видеопоследовательности

Критерий	АСВ2 ( $m = 3$ )	АСВ1 ( $m = 3$ )	АСВ1 ( $m = 4$ )
<b>PSNR</b>	22,69	21,97	22,04
<b>SSIM</b>	0,91	0,90	0,91
<b>NIQE</b>	5,71	5,33	5,30

Таблица 4.5 – Точность восстановления исходного кадра ВР в случае отсутствия АП на входной видеопоследовательности

Критерий	АСВ2 ( $m = 3$ )	АСВ1 ( $m = 3$ )	АСВ1 ( $m = 4$ )
<b>PSNR</b>	23,40	25,12	25,82
<b>SSIM</b>	0,92	0,95	0,95
<b>NIQE</b>	5,91	5,30	5,81

Также было проведено сравнение среднего времени обработки последовательности кадров при помощи рассматриваемых алгоритмов, результаты которого приведены в таблице 4.6. Вычисления проводились в тех же условиях, что и при исследовании алгоритма АСВ1 в п. 4.2.2.

Таблица 4.6 – Среднее время обработки одного кадра при помощи рассматриваемых алгоритмов

Критерий	АСВ2 ( $m = 3$ )	АСВ1 ( $m = 3$ )	АСВ1 ( $m = 4$ )
<b>Время (сек)</b>	1,53	6,08	8,01

Анализ полученных результатов показывает, что предложенный алгоритм АСВ2 превосходит ранее рассмотренный алгоритм АСВ1 в скорости обработки и в устранении негативного эффекта АП, действующих на входные изображения НР, несколько уступая ему в качестве получаемых изображений.



#### 4.4 Применение алгоритмов сверхразрешения видеоданных для обработки реальных видеопоследовательностей

Одной из проблем, возникающих при работе с видеоданными, является то обстоятельство, что области с полезной информацией могут быть частично или полностью закрыты от наблюдения движущимися объектами (рисунок 4.10). Например, такая ситуация может возникнуть в системах наружного видеонаблюдения.



Рисунок 4.10 – Пример закрытия полезных областей кадров видеопоследовательности движущимся объектом

При обработке таких видеопоследовательностей можно рассматривать области перекрытий как области, подверженные воздействию АП, и, соответственно, применять алгоритмы СР видеоданных в условиях АП с целью повышения качества входных кадров качества путём устранения ЛОЗ на них. Далее будет рассмотрено использование ранее предложенных в этой главе алгоритмов на основе нейронных сетей для обработки реальных видеофайлов.

##### 4.4.1 Сегментация движущихся объектов на изображениях

Ранее рассмотренные алгоритмы для сегментации областей, подверженных воздействию АП, на входных изображениях, такие как двухэтапный алгоритм на основе суперпиксельного представления изображений (п. 2.3) и свёрточная нейронная сеть для семантической сегментации (п. 3.3.2), требуют априорных

знаний о характере АП на изображениях. При наличии относительно небольшого количества неразмеченных изображений их использование представляется затруднительным.

Тем не менее, используя указанные ниже допущения, можно предложить относительно простой алгоритм сегментации ЛОЗ, основанный на сравнении видеопоследовательности с эталонным изображением (первым кадром) и ряда морфологических операций:

- первый кадр видеопоследовательности не имеет областей перекрытия;
- камера не совершает существенных смещений между кадрами;
- все быстро движущиеся крупные объекты не представляют интерес с точки зрения дальнейшей обработки.

Данный алгоритм состоит из трёх шагов, а обработка всех кадров осуществляется независимо друг от друга.

1. Предварительная оценка областей перекрытий, осуществляемая путём вычитания одного эталонного изображения из текущего кадра и бинаризации результата.

2. Морфологическая эрозия полученного бинарного изображения для устранения узких и небольших областей, которые могут быть вызваны межкадровыми сдвигами и изменениями в полезной информации на обрабатываемом кадре.

3. Морфологическое расширение бинарного изображения с целью возвращения исходного размера оставшимся областям аномальных наблюдений.

В результате такой обработки, на итоговом бинарном изображении останутся только достаточно большие области, которые соответствуют объектам, закрывающим полезную информацию. При этом можно осуществлять настройку данного алгоритма, изменяя порог бинаризации и размер и форму структурного элемента, используемого в морфологических операциях.

На рисунке 4.11 представлен пример сегментации областей перекрытия для одного из кадров.

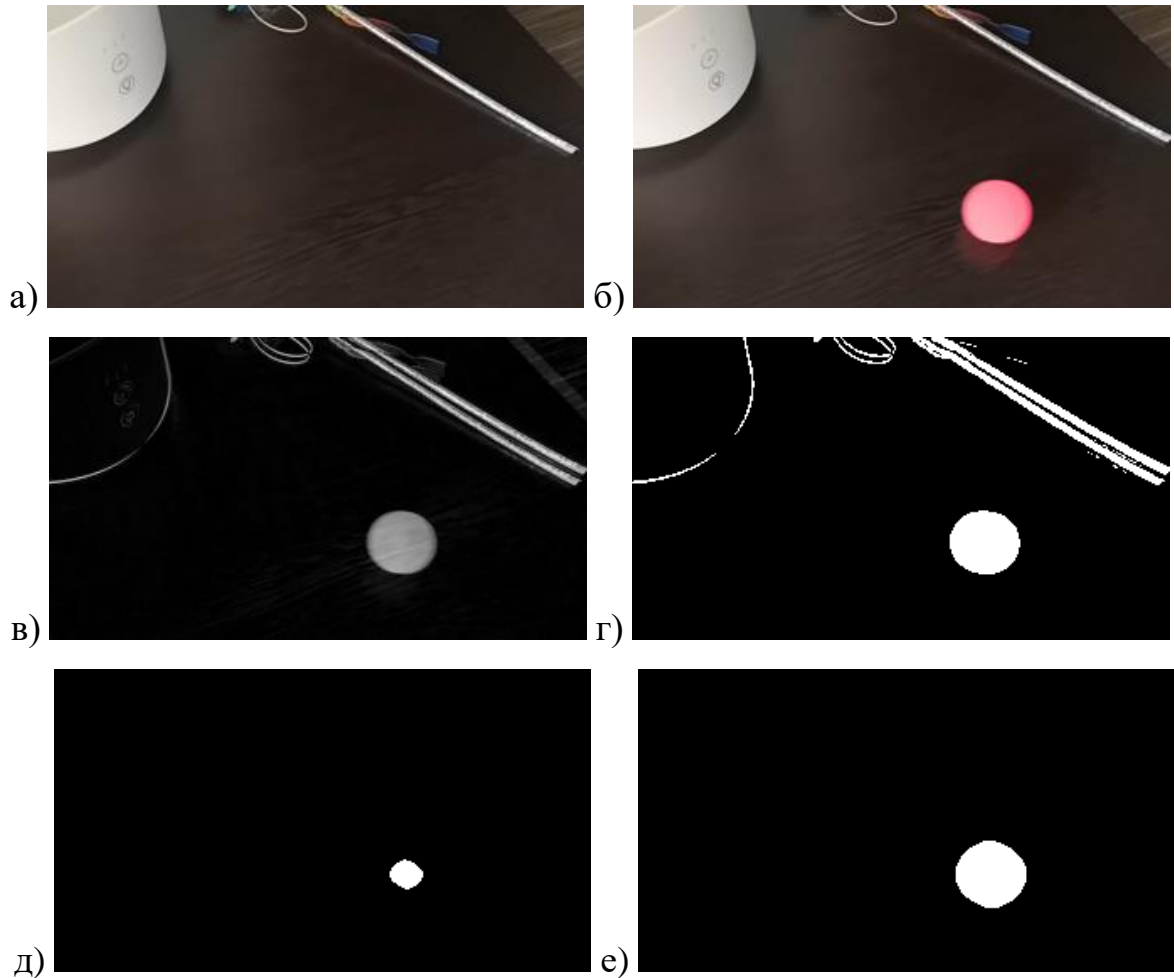


Рисунок 4.11 – Пример сегментации областей нежелательных объектов на изображении: а) эталонный кадр; б) текущий кадр; в) разница между двумя кадрами; г) результат бинаризации; д) результат морфологической эрозии; е) итоговый результат после проведения морфологического расширения

Потенциально, данный алгоритм можно использовать совместно и с другими алгоритмами сегментации ЛОЗ в задачах многокадрового СР с целью повышения точности получаемых результатов, определения свойств АП в процессе обучения или задания вероятностей наличия или отсутствия АП во время процедуры двухэтапной сегментации.

#### 4.4.2 Экспериментальное исследование

Для экспериментального сравнения таких алгоритмов были созданы несколько видеопоследовательностей с движущимися объектами разрешением

320 × 180 длительностью 10 секунд и частотой кадров 30 Гц. Примеры кадров из этих видео представлены выше на рисунке 4.12а.

Для проведения сегментации областей аномальных наблюдений использовался рассмотренный выше алгоритм, оценивающий разницу между изображениями.

В ходе экспериментального исследования осуществлялось повышение разрешения входных кадров в 2 раза. На рисунках 4.12б,в представлены примеры кадров, полученных после обработки видео. Их качественный анализ показывает, что оба рассмотренных ранее алгоритма справляются с устранением движущихся объектов на кадрах видеопоследовательности, сохраняя при этом отдельные изменяющиеся детали.

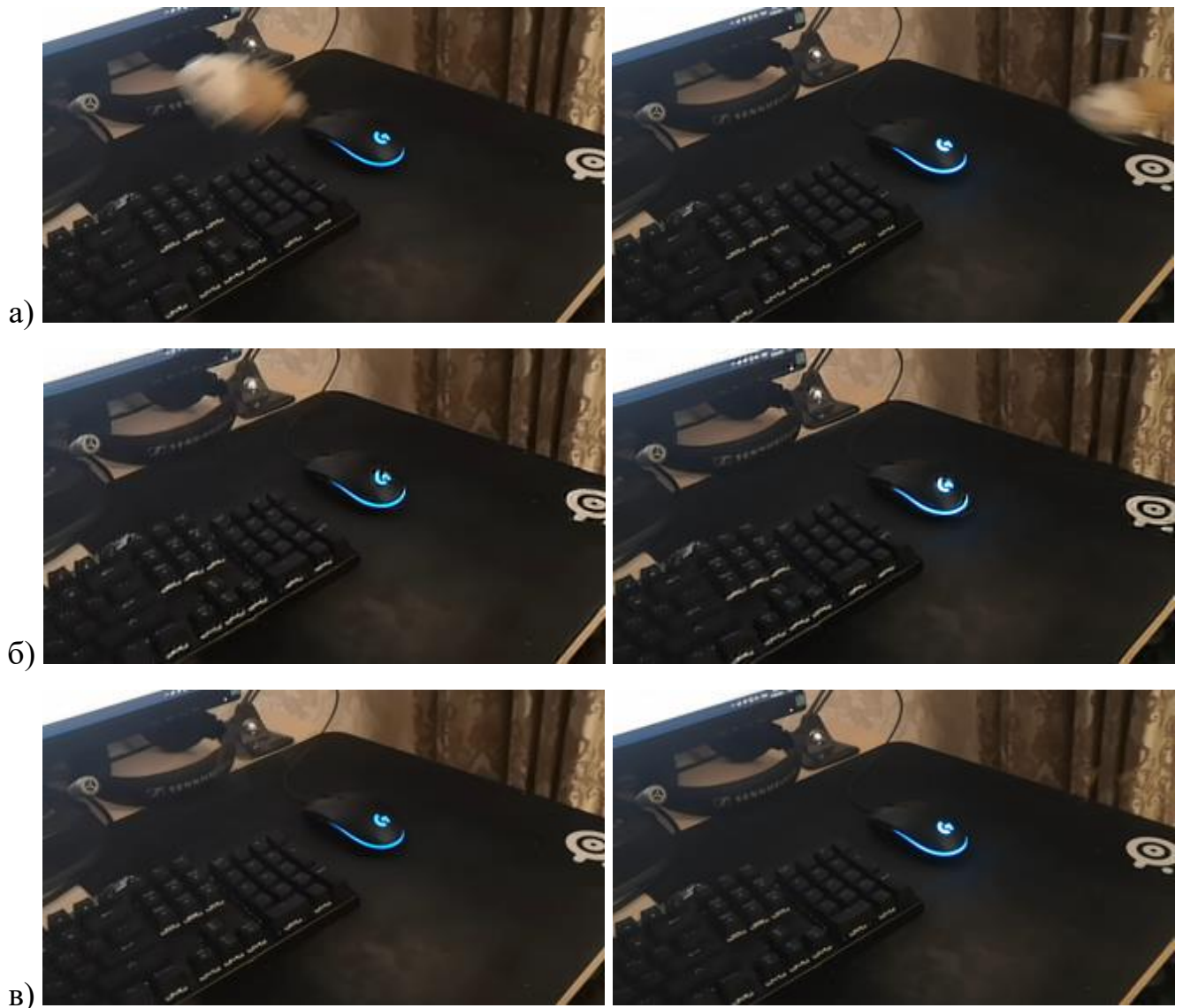


Рисунок 4.12 – Примеры кадров видеопоследовательностей, полученных после обработки разными алгоритмами: а) исходные кадры; б) результаты работы АСВ1 ( $m = 3$ ); в) результаты работы АСВ2 ( $m = 3$ )

В связи с отсутствием эталонных изображений для всех кадров используемых видеопоследовательностей, в качестве единственного образца для сравнения использовался первый кадр, не имеющий нежелательных объектов.

В таблице 4.7 представлены усреднённые значения используемых ранее показателей, полученные в ходе исследования.

Таблица 4.7 – Точность восстановления кадров ВР в случае наличия нежелательных объектов на входных кадрах НР

Критерий	АСВ1 ( $m = 3$ )	АСВ2 ( $m = 3$ )
<b>PSNR</b>	19,91	20,16
<b>SSIM</b>	0,81	0,82
<b>NIQE</b>	8,89	7,42

Их анализ показывает, что оба рассмотренных алгоритма справляются с поставленной задачей примерно на одном уровне, однако алгоритм АСВ2 имеет несколько большую точность.

### **Выводы по главе**

1. Рассмотрена общая схема построения СР видеоданных в условиях наличия АП на входных кадрах, основывающаяся на обработке нескольких кадров в окрестности от текущего и сегментации ЛОЗ при помощи скользящего во времени окна. Предложены варианты реализации такой схемы обработки видеоданных на основе ранее рассмотренной в главе 3 DAG-сети для улучшения качества изображений.

2. Предложен модифицированный алгоритм СР видеоданных в условиях АП, основывающийся на ранее рассмотренном комплексном алгоритме многокадрового СР, использующего искусственные нейронные сети на всех этапах обработки и, в частности, синтезированную ранее двухвходовую DAG-сеть для реализации итеративного алгоритма повышения качества изображений.

3. Предложен алгоритм СР видеоданных в условиях АП, также основанный на использовании глубоких нейронных сетей. Этот алгоритм отличается применением новой, специально модифицированной с учётом особенностей обработки видеоданных многовходовой DAG-сети, позволяющий принимать на вход все кадры в пределах скользящего окна. Предложенная архитектура нейронной сети учитывает особенности обработки видеопоследовательностей

методом скользящего по времени окна, помимо текущего кадра и оценки ВР принимая на вход несколько дополнительных кадров, которые предшествуют или следуют за текущим кадром. Проведено экспериментальное исследование синтезированного алгоритма, в рамках которого создана обучающая выборка и проведено обучение нейронной сети для повышения качества изображений. Проведено экспериментальное исследование, анализ результатов которого показывает, что синтезированный алгоритм успешно справляется с задачей построения СР для видеопоследовательностей и имеет большее быстродействие по сравнению с ранее рассмотренным алгоритмом, несколько уступая ему в точности.

4. Рассмотрена задача устранения на кадрах видеопоследовательности нежелательных объектов, закрывающих области полезных наблюдений, при помощи алгоритмов СР в условиях наличия АП. Предложен относительно простой алгоритм сегментации областей перекрытий на входных кадрах, основанный на использовании априорного изображения сцены в отсутствие АП и выполнении набора морфологических преобразования для вновь поступающих изображений. Проведено экспериментальное исследование на основе реальных видеофайлов с использованием ранее предложенных алгоритмов СР видеоданных, результаты которого демонстрируют, что данные алгоритмы способны эффективно справляться с данной задачей.

## Заключение

В ходе выполнения диссертационной работы были получены следующие результаты:

- проведён анализ существующих подходов к обработке последовательностей изображений с целью повышения разрешения (построения СР) как в традиционной постановке задачи, так и в условиях воздействия на них АП, проявляющихся в областях аномальных наблюдений;

- проведена декомпозиция итеративного процесса обработки информации при анализе последовательности изображений и обоснована общая структура комплексного алгоритма построения многокадрового СР в условиях АП;

- синтезированы и исследованы модели и комплексные алгоритмы многокадрового СР в условиях наличия АП на основе методов оптимальной условно-линейной и нелинейной (адаптивной) фильтрации и использования алгоритмов двухэтапной сегментации областей, подверженных воздействию АП;

- синтезированы и исследованы комплексные алгоритмы многокадрового СР в условиях наличия АП на основе использования свёрточных нейронных сетей, реализующих различные этапы обработки входных изображений;

- проведён сопоставительный анализ результатов построения многокадрового СР при помощи синтезированных алгоритмов на основе методов оптимальной фильтрации и машинного обучения;

- синтезированы и исследованы алгоритмы СР видеоданных в условиях наличия АП на основе метода скользящего окна и использования глубоких свёрточных нейронных сетей для повышения качества входных кадров.

В диссертационной работе получены следующие результаты и выводы, отличающиеся научной новизной.

1. По итогам системного анализа известных подходов к построению СР обоснована и описана общая схема проведения исследований в интересах построения и исследования алгоритмов многокадрового СР и СР видеоданных в условиях наличия АП на входных последовательностях изображениях.



2. Введена общая схема итеративной обработки последовательности изображений для комплексных алгоритмов построения многокадрового построения СР в условиях АП, включающая в себя этапы формирования начальной оценки выходного изображений ВР, оценки межкадровых сдвигов, сегментации локальных областей закрытия на входных изображения и уточнения текущей оценки ВР на основе обработки очередного входного кадра НР.

3. В рамках общей схемы итеративной обработки последовательности изображений при построении многокадрового СР в условиях АП предложены и исследованы модели и комплексные алгоритмы обработки информации, основанные на использовании методов оптимальной динамической фильтрации при неопределенности относительно параметров используемой математической модели и результатов независимой сегментации локальных областей закрытия на входных изображениях. Проведен синтез алгоритма при наличии параметрической неопределенности в условиях неизвестных межкадровых сдвигов изображений. Проведено экспериментальное сравнение результатов его работы с существующими альтернативными алгоритмами, основанными на моделях спиновых стёкол и случайных марковских полей. Анализ полученных результатов показывает, что предложенный алгоритм демонстрирует наибольшую точность восстановления изображения ВР по сравнению с альтернативными алгоритмами, обладая быстроедействием одного с ними порядка.

4. Обоснован и исследован двухэтапный алгоритм сегментации локальных областей закрытия на входных изображениях, основанный на использовании алгоритмов суперпиксельной сегментации и кластерного анализа (EM-алгоритма). Приведены соотношения для включения результатов сегментации в процесс условно-линейной оптимальной фильтрации последовательности изображений с целью повышения их разрешения в условиях наличия АП. В плане развития методов суперпиксельной сегментации предложен и исследован модифицированный алгоритм SLIC, реализующий, в отличие от известного, возможности сегментации областей с одинаковым средним уровнем яркости и различной текстурой.

5. В рамках общей схемы итеративной обработки последовательности изображений при построении многокадрового СР в условиях АП предложены и исследованы комплексные алгоритмы, основанные на использовании глубоких свёрточных сетей на различных этапах выполняемых преобразований и сети в виде направленного ациклического графа, обеспечивающей уточнение оценки изображения ВР на очередной итерации обработки. Рассмотрены два подхода к синтезу таких алгоритмов: без выделения этапа сегментации локальных областей закрытия и с реализацией этапа независимой сегментации каждого входного изображений НР. Проведено экспериментальное исследование предложенных алгоритмов, включавшее в себя как использование предварительно обученных существующих нейронных сетей, так и обучение предложенных сетей. Анализ результатов проведенного эксперимента показал, что алгоритм, использующий глубокие свёрточные сети и сегментацию локальных областей закрытия на входных изображениях, демонстрирует наилучшие результаты, как в плане достигаемой точности, так и в плане быстродействия.

6. Предложены и исследованы алгоритмы СР видеоданных в условиях АП на основе использования глубоких нейронных сетей для повышения качества исходных изображений и адаптированные в интересах обработки кадров видеопоследовательности в скользящем временном окне анализа. Рассмотрены два подхода к решению данной задачи. Первый подход основан на адаптации ранее рассмотренного алгоритма построения многокадрового СР, использующего глубокие нейронные сети на всех этапах обработки и, в частности, синтезированную двухвходовую DAG-сеть для реализации итеративного алгоритма повышения качества изображений. Второй подход основан на применении новой, специально модифицированной с учётом особенностей обработки видеоданных DAG-сети, позволяющий принимать на вход все кадры в пределах скользящего окна.

Проведено экспериментальное сравнение синтезированных алгоритмов, результаты которого показывают, что первый алгоритм имеет большую точность результатов, но уступает второму в скорости обработки

видеопоследовательности. Выполнено экспериментальное исследование предложенных алгоритмов на реальных видеоданных, включающих области закрытия полезной информации, которое показало, что оба предложенных алгоритма достаточно эффективно справляются с задачей повышения разрешения и восстановления искаженных участков изображения.

**Рекомендации по использованию.** Предложенные в диссертационной работе модели и алгоритмы построения многокадрового СР могут использоваться в различных информационных системах, ориентированных на автоматическую и автоматизированную обработку цифровых изображений (системы видеонаблюдения, системы аэрокосмического мониторинга, мобильные устройства). Их применение наиболее целесообразно и эффективно для обработки последовательностей изображений и видеоданных как в случае, отсутствия локальных участков входных изображений, искажённых аппликативными помехами, так и в случае, если исходные изображения могут быть подвержены воздействию подобных помех.

**Перспективы дальнейшего развития** связаны с дальнейшим совершенствованием комплексных алгоритмов на основе глубоких нейронных сетей и сегментации локальных областей закрытия на входных изображениях. Здесь, на наш взгляд, представляет интерес замена отдельных элементов процесса обработки изображений на более точных аналоги, лучше учитывающие специфику воздействия аппликативных помех, организация совместного обучения применяемых на различных этапах обработки нейронных сетей по одним и тем же данным. Кроме того, представляется важным более детальное рассмотрение вопросов обработки видеоданных с учетом специфики регистрации, например, синтез алгоритмов сегментации аппликативных помех, учитывающих особенности их проявления на видеопоследовательностях, которые исключают случайный и независимый характер воздействия на соседние кадры.

**Список использованных источников**

1. **Park, S.C.** Super-resolution image reconstruction: a technical overview / S.C. Park, M.K. Park, M.G. Kang // IEEE Signal Processing Magazine. – 2003. – Vol. 20(3). – P. 21-36. – DOI: 10.1109/MSP.2003.1203207.
2. **Nasrollahi, K.** Super-resolution: a comprehensive survey / K. Nasrollahi, T.B. Moeslund // Machine Vision and Applications. – 2014. – Vol. 25. – P. 1423-1468. – DOI: 10.1007/s00138-014-0623-4.
3. **Yang, C.Y.** Single-image super-resolution: A benchmark / C.Y. Yang, C. Ma, M.H. Yang // European Conference on Computer Vision. – Zurich, Switzerland, September 6-12, 2014. – P. 372-386. – DOI: 10.1007/978-3-319-10593-2\_25.
4. **Huang, D.** A Short Survey of Image Super Resolution Algorithms / D. Huang, H. Liu // Computer Science Technology Updates. – 2015. – Vol. 2(2). – P. 19-29. – DOI: 10.15379/2410-2938.2015.02.02.03.
5. Advanced Topics on Computer Vision / O.V. Villegas, M. Nandayapa, I. Soto (editors). – Springer, 2018. – 432 p.
6. **Журавель, И.М.** Краткий курс теории обработки изображений [Текст] / И.М. Журавель. – М., 1999.
7. **Савельев, М.И.** Эффективный алгоритм восстановления изображения со сверхвысоким разрешением / М.И. Савельев // Вестник СГТУ. – 2007. – Т. 1, №4. – С. 110-115.
8. **Ращупкин, А.В.** Методы обработки данных дистанционного зондирования для повышения качества выходных изображений [Текст] / А.В. Ращупкин // Вестник Самарского государственного аэрокосмического университета. – 2010. – №2. – С. 124-133.
9. **Михайлов, Б.А.** Сверхразрешение в обработке полихромных изображений [Текст] / Б.А. Михайлов // Т-Сomm – Телекоммуникации и Транспорт. – 2013. – №4. – С. 47-49.
10. **Ушакова, Н.Н.** Формирование космического изображения сверхвысокого разрешения с моделированием группировки из изображений реального

- аппарата / Н.Н. Ушакова // Прикладная математика и Физика. – Т. 50, №. 2. – 2018. – С. 229-238.
11. **Shultz, R.R.** A Bayesian approach to image expansion for improved definition / R.R. Schultz, R.L. Stevenson // IEEE Transactions on Image Processing. – 1994. – Vol. 3(3). – P. 233-242. – DOI:10.1109/83.287017.
12. **Hou, H.S.** Cubic splines for image interpolation and digital filtering / H.S. Hou, H.C. Andrews // IEEE Transactions on Acoustics, Speech and Signal Processing. – 1978. – Vol. 26(6). – P. 508-517. – DOI: 10.1109/TASSP.1978.1163154.
13. **Li, X.** New edge-directed interpolation / X. Li, M.T. Orchard // IEEE Transactions on Image Processing. – 2001. – Vol. 10(10). – P. 1521-1527. – DOI: 10.1109/83.951537.
14. **Lin, H.** Research on the Algorithms of Super resolution Image Reconstruction with Wavelet and Interpolation / H. Lin // Xidian University (Chinese). – 2005. – P. 7-14.
15. **Wang, Z.S.** Research on Super-resolution Image Processing Algorithm Based on Interpolation and Worked Examples / Z.S. Wang // Zhejiang Normal University (Chinese). – 2014. – P. 10-22.
16. **Cha, Y.** Edge forming methods for color image zooming / Y. Cha, S. Kim // IEEE Transactions on Image Processing. – 2006. – Vol. 15(8). – P. 2315-2323. – DOI: 10.1109/TIP.2006.875182.
17. **Timofte, R.** A+: Adjusted anchored neighborhood regression for fast super-resolution / R. Timofte, V. De Smet, L. Van Gool // 12th Asian Conference on Computer Vision. – Singapore, Singapore, November 1-5, 2014 – DOI: 10.1007/978-3-319-16817-3\_8.
18. **Schulter, S.** Fast and accurate image upscaling with super-resolution forests / S. Schulter, C. Leistner, H. Bischof. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Boston, USA, June 7-12, 2015. – DOI: 10.1109/CVPR.2015.7299003.
19. **Huang, J.-B.** Single image super-resolution using transformed self-exemplars / J.-B. Huang, A. Singh, N. Ahuja // IEEE Conference on Computer Vision and Pattern

- Recognition (CVPR). – Boston, USA, June 7-12, 2015. – DOI: 10.1109/CVPR.2015.7299156.
- 20.**Freedman, G.** Image and video upscaling from local self-examples / G. Freedman, R. Fattal // ACM Transactions on Graphics. – 2011. – Vol. 30(2). – DOI: 10.1145/1944846.1944852.
- 21.**Yang, J.** Fast image super-resolution based on in-place example regression / J. Yang, Z. Lin, S. Cohen // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Portland, USA, June 23-28, 2013. – P. 1059-1066. – DOI: 10.1109/CVPR.2013.141.
- 22.**Dong, C.** Image super-resolution using deep convolutional networks / C. Dong, C.C. Loy, K. He, X. Tang // IEEE Transactions on Pattern Analysis and Machine Intelligence – 2016. – Vol. 38(2). – P. 295-307. – DOI: 10.1109/TPAMI.2015.2439281.
- 23.**Dong, C.** Accelerating the super-resolution convolutional neural network / C. Dong, C.C. Loy, X. Tang // Proceeding of European Conference on Computer Vision. – 2016. – P. 391-407. – DOI: 10.1007/978-3-319-46475-6\_25.
- 24.**Kim, J.** Accurate Image Super-Resolution Using Very Deep Convolutional Networks / J. Kim, J.K. Lee, K.M. Lee // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Las Vegas, USA, 2016. – P. 1646-1654. – DOI: 10.1109/CVPR.2016.182.
- 25.**Ledig, C.** Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network / C. Ledig, L. Theis, F. Huszár et al. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Honolulu, USA, 2017. – P. 105-114. – DOI: 10.1109/CVPR.2017.19.
- 26.**Tsai, R.Y.** Multipleframe image restoration and registration / R.Y. Tsai, T.S. Huang // Advances in Computer Vision and Image Processing. – 1984. – P. 317-339.
- 27.**Kim, S.P.** Recursive reconstruction of high resolution image from noisy undersampled multiframe / S.P. Kim, N.K. Bose, H.M. Valenzuela // IEEE Transactions on Acoustics, Speech, and Signal Processing. – 1990. – Vol. 38(6). – P. 1013-1027. – DOI: 10.1109/29.56062.

28. **Kim, S.P.** Recursive high-resolution reconstruction of blurred multiframe images / S.P. Kim, W.Y. Su // IEEE Transactions on Image Processing – 1993. – Vol. 2(4). – P. 534–539. – DOI: 10.1109/83.242363.
29. **Nguyen, N.** An efficient wavelet-based algorithm for image superresolution / N. Nguyen, P. Milanfar // IEEE International Conference on Image Processing. – Vancouver, Canada, September 10-13, 2000. – P. 351-354. – DOI: 10.1109/ICIP.2000.899387.
30. **Ji, H.** Robust wavelet-based super-resolution reconstruction: theory and algorithm / H. Ji, C. Fermüller // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2009. – Vol. 31(4). – P. 649-660. – DOI: 10.1109/TPAMI.2008.103.
31. **Ur., H.** Improved Resolution from Subpixel Shifted Pictures / H. Ur, D. Gross // CVGIP: Graphical Models and Image Processing. – 1992. – Vol. 54(2). – P. 181-186. – DOI: 10.1016/1049-9652(92)90065-6.
32. **Bose, N.K.** Superresolution and Noise Filtering Using Moving Least Squares / N.K. Bose, N.A. Ahuja // IEEE Transactions on Image Processing. – 2006. – Vol. 15(8). – P. 2239-2248. – DOI: 10.1109/TIP.2006.877406.
33. **Chantas, G.K.** Super-resolution based on fast registration and maximum a posteriori reconstruction / G.K. Chantas, G.P. Galatsanos, N.A. Woods // IEEE Transactions on Image Processing. – 2007. – Vol. 16(7). – P. 1821-1830. – DOI: 10.1109/TIP.2007.896664.
34. **Tian, J.** Stochastic super-resolution image reconstruction / J. Tian, K.-K. Ma // Journal of Visual Communication and Image Representation. – 2010. – Vol. 21(3). – P. 232-244. – DOI: 10.1016/j.jvcir.2010.01.001.
35. **Belekos, S.P.** Maximum a Posteriori Video Super-Resolution Using a New Multichannel Image Prior / S.P. Belekos, N.P. Galatsanos, A.K. Katsaggelos // IEEE Transactions on Image Processing. – 2010. – Vol. 19(6). – P. 1451-1464. – DOI: 10.1109/TIP.2010.2042115.
36. **Irani, M.** Super resolution from image sequences / M. Irani, S. Peleg // Proceeding of 10th International Conference on Pattern Recognition. – 1990. – Vol. 2. – P. 115-120.

- 37.**Ji, H.** Wavelet-Based Super-Resolution Reconstruction: Theory and Algorithm / H. Ji, C. Fermüller // European Conference on Computer Vision. – Graz, Austria, May 7-13, 2006. – Vol. 3954. – P. 295-307. – DOI: 10.1007/11744085\_23.
- 38.**Patti, A.J.** Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time / A.J. Patti, M.I. Sezan, A.M. Tekalp // IEEE Transactions on Image Processing. – 1997. – Vol. 6(8). – P. 1064-1076. – DOI: 10.1109/83.605404.
- 39.**Elad, M.** Super-resolution reconstruction of image sequences / M. Elad, A. Feuer // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1999. – Vol. 21(9). – P. 817-834. – DOI: 10.1109/34.790425.
- 40.**Newland, C.** Time invariant steady-state Kalman filter for image super-resolution / C. Newland, D. Gray // Image and Vision Computing New Zealand 2005 (IVCNZ05). – 2006. – P. 381-387.
- 41.**Newland, C.** Modified Kalman filtering for image super-resolution / C. Newland, D. Gray, D. Gibbins // Image and Vision Computing New Zealand 2006 (IVCNZ06). – 2006. – P. 79-84.
- 42.**Wei, Z.** Kalman filter-based method for image superresolution using a sequence of low-resolution images / Z. Wei, F. Tao, W. Jun // Journal of Electronic Imaging. – 2014. – Vol. 23(1). – DOI: 10.1117/1.JEI.23.1.013008.
- 43.**Sajjadi, M.** Frame-Recurrent Video Super-Resolution. / M. Sajjadi, R. Vemulapalli, M. Brown // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2018. – P. 6626-6634. – DOI: 10.1109/CVPR.2018.00693.
- 44.**Kawulok, M.** Deep Learning for Multiple-Image Super-Resolution / M. Kawulok, P. Benecki, S. Piechaczek, K. Hrynczenko, D. Kostrzewa, J. Nalepa // IEEE Geoscience and Remote Sensing Letters. – 2019. – Vol. 17. – P. 1062-1066. – DOI: 10.1109/LGRS.2019.2940483.
- 45.**Molini, A.B.** DeepSUM: Deep Neural Network for Super-Resolution of Unregistered Multitemporal Images / A.B. Molini, D. Valsesia, G. Fracastoro, E. Magli // IEEE Transactions on Geoscience and Remote Sensing. – 2019. – P. 1-13. – DOI: 10.1109/TGRS.2019.2959248.



46. **Kim, J.** Deeply-Recursive Convolutional Network for Image Super-Resolution / J. Kim, J.K. Lee, K.M. Lee // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA. – 2016. – P. 1637-1645. – ISSN: 1063-6919. – DOI: 10.1109/CVPR.2016.181.
47. **Liu, C.** On Bayesian Adaptive Video Super Resolution / C. Liu, D. Sun // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2014. – Vol. 36(2). – P. 346-360. – DOI: 10.1109/TPAMI.2013.127.
48. **Kappeler, A.** Video Super-Resolution with Convolutional Neural Networks / A. Kappeler, S. Yoo, Q. Dai, A.K. Katsaggelos // IEEE Transactions on Computational Imaging. – 2016. – Vol. 2(2). – P. 109-122. – DOI: 10.1109/TCI.2016.2532323.
49. **Makansi, O.** End-to-End Learning of Video Super-Resolution with Motion Compensation / O. Makansi, E. Ilg, T. Brox. // Pattern Recognition. GCPR 2017. Lecture Notes in Computer Science. – P. 203-214. – DOI: 10.1007/978-3-319-66709-6\_17.
50. **Caballero, J.** Real-Time Video Super-Resolution with Spatio-Temporal Networks and Motion Compensation / J. Caballero, C. Ledig, A. Aitken et al. // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Honolulu, USA, July 21-26, 2017. – P. 2848-2857. – DOI: 10.1109/CVPR.2017.304.
51. **Ma, Z.** Handling motion blur in multi-frame super-resolution / Z. Ma, R. Liao, X. Tao, L. Xu, J. Jia, E. Wu // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Boston, USA, June 7-12, 2015. – P. 5224-5232. – DOI: 10.1109/CVPR.2015.7299159.
52. **Kanemura, A.** Image superresolution under spatially structured noise / A. Kanemura, S. Maeda, S. Ishii // IEEE Int. Symp. Signal Processing and Information Technology. – Cairo, Egypt, December 15-18, 2007. – P. 275-280. – DOI: 10.1109/ISSPIT.2007.4458156.
53. **Fukuda, W.** Superresolution from occluded scenes / W. Fukuda, A. Kanemura, S. Maeda, S. Ishii // International Conference on Neural Information Processing. –

- Bangkok, Thailand, December 1-5, 2009. – P. 19-27. – DOI: 10.1007/978-3-642-10684-2\_3.
54. **Negi, C.S.** Super-resolution de-fencing: Simultaneous fence removal and high-resolution image recovery using videos / C.S. Negi, K. Mandal, R.R. Sahay, M.S. Kankanhalli // Proc. Of IEEE International Conference on Multimedia and Expo Workshops. – 2014. – P. 1-6. – DOI: 10.1109/ICMEW.2014.6890641.
55. **Ivankov, A.Y.** Image sequence filtering algorithms for resolution improvement in applicative noise conditions / A.Y. Ivankov, A.A. Sirota // Optoelectronics, Instrumentation and Data Processing. – 2015. – Vol. 51. – P. 622-631. – DOI: 10.3103/S8756699015060151.
56. **Иванков, А.Ю.** Блочные алгоритмы обработки изображений на основе фильтра Калмана в задаче построения сверхразрешения / А.Ю. Иванков, А.А. Сирота // Компьютерная оптика. – 2014. – Т. 38, № 1. – С. 118-125.
57. **Иванков, А.Ю.** Адаптивные алгоритмы построения сверхразрешения на основе обработки последовательности изображений / Иванков А.Ю., Сирота А.А. // Оптический журнал. – 2017. – Т. 84, № 5. – С. 38-45.
58. **Иванков, А.Ю.** Адаптивный алгоритм оптимальной фильтрации для повышения разрешения на основе обработки последовательности цифровых изображений // Вестн. Воронежского гос. ун-та. Сер. Системный анализ и информационные технологии. – 2016. – № 4. – С. 142-147.
59. **Саввин, С.В.** Алгоритм построения сверхразрешения на основе калмановской фильтрации и суперпиксельной сегментации изображений / С.В. Саввин, А.А. Сирота, А.Ю. Иванков // XIX Международная конференция «Информатика: проблемы, методология, технологии». – Воронеж, Россия, 14-15 февраля, 2019. – С. 862-866.
60. **Savvin, S.V.** Superresolution algorithms in applicative noise conditions using superpixel segmentation / S.V. Savvin, A.A. Sirota, A.Yu. Ivankov // International Conference on Digital Image & Signal Processing. – Oxford, United Kingdom, April, 29-30, 2019.

61. **Саввин, С.В.** Синтез и анализ алгоритмов построения многокадрового сверхразрешения в условиях аппликативных помех / С.В. Саввин, А.А. Сирота // Международная научно-техническая конференция «Актуальные проблемы прикладной математики, информатики и механики». – Воронеж, Россия, 11-13 ноября, 2019. – С. 376-383.
62. **Саввин, С.В.** Анализ алгоритмов построения сверхразрешения изображений в условиях аппликативных помех / С.В. Саввин, А.А. Сирота, А.Ю. Иванков // XX Международная конференция «Информатика: проблемы, методология, технологии». – Воронеж, Россия, 13-14 февраля, 2020. – С. 1155-1161.
63. **Savvin, S.V.** Performance of super-resolution algorithms under applicative noise / S.V. Savvin, A.A. Sirota // Journal of Physics: Conference Series. Applied Mathematics, Computational Science and Mechanics: Current Problems. – 2019. – Vol. 1479. – P. 11-13. – DOI: 10.1088/1742-6596/1479/1/012080.
64. **Savvin, S.V.** An Algorithm for Multi-Frame Image Super-Resolution Under Applicative Noise Based on a Convolutional Neural Network / S.V. Savvin, A.A. Sirota // 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA). – Lipetsk, Russia, November 11-13, 2020. – P. 422-424. – DOI: 10.1109/SUMMA50634.2020.9280698.
65. **Саввин, С.В.** Методы и алгоритмы построения сверхразрешения для последовательности изображений в условиях аппликативных помех / С.В. Саввин, А.А. Сирота, А.Ю. Иванков // Известия РАН. Теория и системы управления. – 2021. – № 3. – С. 136-148. – DOI: 10.31857/S0002338821030070.
66. **Саввин, С.В.** Алгоритмы построения многокадрового сверхразрешения изображений в условиях аппликативных помех на основе глубоких нейронных сетей / С.В. Саввин, А.А. Сирота // Компьютерная оптика. – 2022. – Т. 46, № 1. – С. 130-138. – DOI: 10.18287/2412-6179-CO-904.
67. **Саввин, С.В.** Алгоритмические способы построения сверхразрешения видеоданных в условиях аппликативных помех с использованием глубоких нейронных сетей / С.В. Саввин // Вестн. Воронежского гос. ун-та, Сер.

- Системный анализ и информационные технологии. – Воронеж: ИПЦ ВГУ: 2021. – № 4. – С. 107-120. – DOI: 10.17308/sait.2021.4/3801.
68. Обработка информации в пространственно-распределенных системах радиомониторинга: статистический и нейросетевой подходы / Э.А. Кирсанов, А.А. Сирота. – М.: Физматлит, 2012. – 343 с.
69. **Калинин, П.В.** Статистические, нейросетевые и комбинированные алгоритмы фильтрации аппликативных помех на изображениях / П.В. Калинин, А.А. Сирота // Автометрия. – 2012. – Т. 48, №6. – С. 18-28.
70. **Boykov, Y.** Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images / Y. Boykov, M.P. Jolly // Proceeding of the 8th Intern. Conf. on Computer Vision, Vancouver, Canada, July 7–14, 2001. – Vol. 1. – P. 105-112. – DOI: 10.1109/ICCV.2001.937505.
71. **Lloyd, S.** Least Squares Quantization in PCM / S. Lloyd // IEEE Transactions on Information Theory. 1982. – Vol. 28(2). – P. 129-137. – DOI: 10.1109/TIT.1982.1056489.
72. Finite Mixture Models / G. McLachlan, D. Peel. – Hoboken, NJ: John Wiley & Sons, Inc., 2000. – 427 p.
73. Методы и алгоритмы анализа данных и их моделирование в MATLAB: учеб. пособие. / А.А. Сирота – СПб.: БХВ-Петербург, 2016. – 384 с.
74. **Bertalmío, M.** Image inpainting / M. Bertalmío, G. Sapiro, V. Caselles, C. Ballester // Proceedings of the ACM SIGGRAPH Conference on Computer Graphics. – July, 2000. – P. 417-424. – DOI: 10.1145/344779.344972.
75. **Jiahui, Y.** Free-Form Image Inpainting with Gated Convolution / Y. Jiahui, Z. Lin, J. Yang, X. Shen, X. Lu, T. Huang // IEEE/CVF International Conference on Computer Vision (ICCV). – Seoul, Korea, October 27-November 2, 2019. – P. 4470-4479. – DOI: 10.1109/ICCV.2019.00457.
76. **Stutz, D.** Superpixels: An evaluation of the state-of-the-art / D. Stutz, A. Hermans, B. Leibe // Computer Vision and Image Understanding. – 2018. – Vol. 166. – P. 1-27. – DOI: 10.1016/j.cviu.2017.03.007.

77. **Achanta, R.** SLIC Superpixels Compared to State-of-the-Art Superpixel Methods / R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Ssstrunk // IEEE transactions on pattern analysis and machine intelligence. – 2012. – V. 34(11). – P. 2247-2281. – DOI: 10.1109/TPAMI.2012.120.
78. **Саввин, С.В.** Алгоритмы суперпиксельной сегментации изображений и их применение для анализа изображения с разнородной текстурой / С.В. Саввин, А.А. Сирота // Вестн. Воронежского гос. ун-та, Сер. Системный анализ и информационные технологии. – 2016. – № 4. – С. 165-173.
79. **Саввин, С.В.** Модели и алгоритмы суперпиксельной сегментации изображений / С.В. Саввин, А.А. Сирота // Сборник студенческих научных работ факультета компьютерных наук ВГУ. – Воронеж: ИПЦ ВГУ, 2016. – Вып. 10. – С. 160–165.
80. **Malik, J.** Normalized Cuts and Image Segmentation / J. Malik, J. Shi // IEEE transactions on pattern analysis and machine intelligence. – 2000. – Vol 22(8). – P. 888-905. – DOI. 10.1109/34.868688.
81. **Liu, M.** Entropy Rate Superpixel Segmentation / M. Liu, O. Tuzel, S. Ramalingam, R. Chellappa // Conference on Computer Vision and Pattern Recognition – Colorado Springs, USA, June 20-25, 2011. – DOI: 10.1109/CVPR.2011.5995323.
82. **Veksler, O.** Superpixels and Supervoxels in an Energy Optimization Framework / O. Veksler, Y. Boykov, P. Mehrani // 11th European Conference on Computer Vision. – Heraklion, Greece, September 5-11, 2010. – P. 211-224. – DOI. 10.1007/978-3-642-15555-0\_16.
83. **Long, J.** Fully convolutional networks for semantic segmentation / J. Long, E. Shelhamer, T. Darrel // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2014. – Vol. 39(4). – P. 640-651. – DOI: 10.1109/TPAMI.2016.2572683.
84. Handbook of Deep Learning Applications / V.E. Balas, S.S. Roy, D. Sharma, P. Samui (editors). – Springer, 2019. – 383 p.
85. **Chen, L.C.** Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation / L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam //

- European Conference on Computer Vision. – Munich, Germany, September 8-14, 2018. – DOI: 10.1007/978-3-030-01234-2\_49.
- 86.**Ronneberger, O.** U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // 18th International Conference on Medical Image Computing and Computer Assisted Intervention. – Munich, Germany, October 5-9, 2015. – DOI: 10.1007/978-3-319-24574-4\_28.
- 87.**Zhou, W.** Image Quality Assessment: From Error Visibility to Structural Similarity / W. Zhou, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli // IEEE Transactions on Image Processing. – 2004. – Vol. 13(4). – P. 600-612. – DOI: 10.1109/TIP.2003.819861.
- 88.**Mittal, A.** Making a Completely Blind Image Quality Analyzer / A. Mittal, R. Soundararajan, A.C. Bovik // IEEE Signal Processing Letters. – 2013. – V. 22(3). – P. 209-212. – DOI: 10.1109/LSP.2012.2227726.
- 89.**Sun, D.** PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume / D. Sun, X. Yang, M. Liu, J. Kautz // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. – Salt Lake City, USA, June 18-23, 2018. – P. 8934-8943. DOI: 10.1109/CVPR.2018.00931.
- 90.**Hui, T.-W.** A Lightweight Optical Flow CNN – Revisiting Data Fidelity and Regularization / T.-W. Hui, X. Tang, C.C. Loy // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2020. – DOI: 10.1109/TPAMI.2020.2976928.
- 91.**Krizhevsky, A.** ImageNet Classification with Deep Convolutional Neural Networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in neural information processing systems. – 2012. – Vol. 25(2). – P. 1-9. – DOI: 10.1145/3065386.
- 92.**Nickolls, J.** Scalable Parallel Programming with CUDA: Is CUDA the parallel programming model that application developers have been waiting for? / J. Nickolls, I. Buck, M. Garland, K. Skadron // Queue. – 2008. – Vol 6(2). – P. 40-53. – DOI: 10.1145/1365490.1365500.
- 93.**Pattern Recognition and Machine Learning / C.M. Bishop.** – Springer, New York, NY, 2006. – 758 p.

94. **Kingma, D.** Adam: A Method for Stochastic Optimization / D. Kingma, J.L. Ba // 3rd International Conference for Learning Representations. San Diego, USA, 2015. – <https://arxiv.org/abs/1412.6980>.
95. Введение в статистическую теорию распознавания образов / К. Фукунага. – М.: Наука, 1979. – 368 с.
96. Анализ и компьютерное моделирование информационных процессов и систем / Э.К. Алгаинов, А.А. Сирота. – М.: Диалог-МИФИ, 2009. – 416 с.
97. Statistical Physics of Spin Glasses and Information Processing / H. Nishimori. – Oxford: Oxford University Press, 2001. – 256 p.
98. Matrix Computations, 3rd ed / G.H. Golub, F.V. Loan. – Baltimore: Johns Hopkins University Press, 1996. – 694 p.
99. Markov Random Field Modeling in Image Analysis / S.Z. Li. – Berlin: Springer-Verlag, 2001. – 384 p.
100. **Xiao, J.** SUN Database: Large-scale Scene Recognition from Abbey to Zoo / J. Xiao, J. Hays, K. Ehinger, A. Oliva, A. Torral // IEEE Conference on Computer Vision and Pattern Recognition. – San Francisco, USA, June 13-18, 2010. – P. 3485-3492. – DOI: 10.1109/CVPR.2010.5539970.