

На правах рукописи



Носков Андрей Александрович

**Формирование изображений с расширенной
глубиной резкости для систем прикладного
телевидения**

Специальность 05.12.04 – Радиотехника, в том числе системы и устройства
телевидения

ДИССЕРТАЦИЯ

на соискание ученой степени
кандидата технических наук

Научный руководитель

д. т. н., доцент

Приоров Андрей Леонидович

Оглавление

Введение	4
Глава 1. Обзор методов формирования изображений с расширенной глубиной резкости	12
1.1. Причины размытия изображений	12
1.2. Обзор методов смешивания изображений с расширенной глубиной резкости	17
1.3. Общий алгоритм смешивания изображений с расширенной глубиной резкости	18
1.4. Пространственные методы смешивания	20
1.5. Трансформационные методы смешивания	22
1.6. Принципы работы клеточных автоматов	24
1.7. Краткие выводы	29
Глава 2. Алгоритм смешивания изображений с расширенной глубиной резкости на основе клеточного автомата	30
2.1. Тестовые серии исходных изображений	30
2.2. Предварительная обработка изображений исходной серии	31
2.3. Оценка сфокусированности областей изображений исходной серии	35
2.4. Работа клеточного автомата	49
2.5. Формирование итогового изображения	52
2.6. Краткие выводы	57
Глава 3. Исследование алгоритма формирования изображений с расширенной глубиной резкости на основе клеточных автоматов и пирамид изображений	59
3.1. Исследование работы алгоритма в условиях наличия шумов	59

3.2. Способы оценки качества изображений	68
3.3. Сравнение алгоритмов формирования изображений с расширенной глубиной резкости	78
3.4. Краткие выводы	87
Заключение	89
Список сокращений и условных обозначений	92
Список литературы	94
Приложение А. Примеры изображений тестовых серий	104
Приложение Б. Свидетельства о государственной регистрации программ для ЭВМ	105
Приложение В. Акты внедрения результатов работы	108

Введение

Актуальность темы. Цифровая обработка визуальной информации в настоящее время охватывает различные виды инфокоммуникационных и радиотехнических приложений и расширяет их спектр. Сюда относятся как традиционные приложения (вещательное, промышленное, охранное телевидение), так и относительно новые приложения (видеоконференцсвязь, техническое зрение, цифровое кино, телевидение высокой и сверхвысокой четкости, 3D-телевидение и вычислительная фотография).

Значительный вклад в данную область науки и техники внесли как отечественные ученые М.И. Кривошеев, В.А. Соيفер, М.К. Чобану, А.С. Крылов, Д.С. Ватолин, А.С. Конушин, Б.А. Алпатов, Ю.С. Бехтин, Б.В. Костров, В.Ю. Волков, Е.П. Петров, И.С. Трубин, Е.В. Медведева, М.Н. Фаворская, Ю.С. Радченко, так и зарубежные – S. Mitra, R. Gonzalez, R. Woods, Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, J. Astola, K. Egiazarian, R. Szeliski, R. Lukas и др.

Помимо систем телевидения, обработка визуальной информации активно используется во многих областях человеческой деятельности. В связи с массовым распространением мобильных устройств фото- и видеофиксации большое значение приобретают алгоритмы улучшения качества получаемых при помощи них изображений.

В последние годы как в России, так и за рубежом активно развивается отдельная ветвь обработки цифровых изображений – так называемая вычислительная фотография. Основной задачей этой области знаний является повышение качества получаемых цифровых изображений как аппаратными, так и программными средствами. Одним из основных направлений развития вычислительной фотографии является совмещение информации, содержащейся в нескольких изображениях одной сцены. Характер и цели совмещения могут быть различными: повышение динамического диапазона, повышение разрешения, формирование комбинированных изображений с целью выявления скры-

тых объектов в динамических сценах и другие.

Можно выделить достаточно большой класс специальных и бытовых устройств, а также условий съемки, при которых получаемые цифровые изображения имеют эффект ограниченной глубины резко изображаемого пространства. Наиболее наглядно данный эффект можно наблюдать при рассмотрении изображений, полученных при помощи микроскопа. Аналогичная ситуация имеет место при съемке в условиях низкой освещенности, когда для получения качественного изображения необходимо увеличивать размер диафрагмы снимающего устройства, уменьшая тем самым глубину резкости. Приведенные ограничения послужили основой для постановки отдельной задачи в рамках вычислительной фотографии – формирование изображений с расширенной глубиной резкости.

Основной задачей формирования изображений с расширенной глубиной резкости является комбинирование двух или более изображений в одно, которое является более информативным и пригодным для визуального распознавания или цифровой обработки, чем исходные. Данная область вычислительной фотографии начала активно развиваться с начала 2000-х годов. Техники смешивания на основе нескольких, снятых с различным фокусным расстоянием, изображений можно условно разделить на две группы: пространственные техники и техники с использованием трансформационных преобразований. В первых – итоговое изображение состоит из пикселей исходных изображений, находящихся в соответствующих позициях. Во вторых – изображение переводится в некоторое пространство признаков. Наиболее часто используемые методы решения данной задачи основаны на различного рода преобразованиях частотных характеристик исходных изображений. Примером могут служить методы на основе нейронных сетей, вейвлет-преобразования, дискретного косинусного преобразования и др. Основные результаты работ в этой области представлены в работах Ф. Сроубека (F. Sroubek), А. Малика (A. Malik), Д. Ванга (J. Wang), Х. Керке (H. Kekre) и др. Однако все эти техники вносят различного вида размытия и

искажения на границах объектов на итоговом изображении. Ключевой задачей всех алгоритмов смешивания является борьба с этими явлениями. Анализ существующих подходов к формированию изображений с расширенной глубиной резкости показал, что универсальных решений в данной области на настоящий момент не существует [1–6]. Алгоритмы, выполняющие рассматриваемую задачу, используются в большом числе практических приложений: системах технического зрения, робототехнике, медицине, криминалистике и т. д.

Все вышеперечисленное доказывает, что формирование изображений с расширенной глубиной резкости представляет собой актуальную научно-техническую задачу как для области телевидения, так и для ряда смежных научно-технических областей.

Целью работы является повышение информационной емкости изображений при помощи расширения глубины резко изображаемого пространства для систем прикладного телевидения.

Объектом исследования являются радиотехнические системы фильтрации и сжатия визуальной информации, системы технического зрения и прикладного телевидения.

Предметом исследования являются алгоритмы формирования изображений с расширенной глубиной резкости на основе клеточных автоматов и пирамид изображений.

Задачи диссертационной работы:

- проведение исследований в области существующих решений задачи формирования изображений с расширенной глубиной резкости;
- проведение исследований по выбору метрики сфокусированности пикселей изображений исходной серии;
- разработка и анализ алгоритма формирования изображений с расширенной глубиной резкости на основе аппарата клеточных автоматов;

- разработка и анализ алгоритма смешивания размеченных изображений исходной серии на основе пирамид гауссианов и лапласианов;
- проведение анализа работы разработанного алгоритма в условиях наличия аддитивного белого гауссовского шума.

Методы исследования. При решении поставленных задач использовались современные методы цифровой обработки изображений, технического зрения, распознавания образов, математического анализа, теории вероятностей и математической статистики. Для практической реализации алгоритмов применялись современные численные методы, методы программирования на языках Matlab и Python, а также методы объектно-ориентированного программирования на языке C++.

Научная новизна полученных результатов. В рамках работы получены следующие новые научные результаты:

- Предложено использование клеточных автоматов в качестве аппарата, формирующего правило смешивания в задаче формирования изображений с расширенной глубиной резкости.
- Предложено использование пирамид гауссианов и лапласианов как средства совмещения размеченных изображений исходных серий для повышения качества формируемых изображений.
- Предложены идеализированные значения метрики сфокусированности на основе физических аспектов рассматриваемой задачи.
- Разработан алгоритм формирования изображений с расширенной глубиной резкости на основе клеточного автомата.
- Разработан модифицированный алгоритм формирования изображений с расширенной глубиной резкости на основе пирамид изображений.

Практическая значимость полученных результатов

- Разработанный алгоритм формирования изображений с расширенной глубиной резкости и его модификация на основе пирамид изображений позволяют формировать достоверное представление о снимаемой сцене, что дает возможность использовать его как компонент систем технического зрения, либо анализа человеком на основе единственного изображения.
- Разработанный алгоритм является вычислительно эффективным, так как он работает в пространственной области и использует вычислительно не затратный аппарат клеточных автоматов, для которого применимы приемы SIMD-программирования и параллельных вычислений.
- Разработанная методика сравнения метрик сфокусированности может быть использована для оценки эффективности собственных метрик другими исследователями в данной области.
- Приведены рекомендации по выбору метрик сфокусированности. Установлено, что метрика на основе отношения коэффициентов вейвлет-преобразования (ОКВП) имеет наибольший коэффициент корреляции с идеализированными характеристиками ($r = 0,88$), а также наименьший доверительный интервал. Метрика на основе вариации уровня серого имеет наименьшую вычислительную сложность (в 22 раза быстрее ОКВП) при этом сохраняя высокий уровень точности ($r = 0,76$).
- Приведены рекомендации по выбору подхода к подавлению аддитивного белого гауссовского шума (АБГШ) при формировании изображений с расширенной глубиной резкости. Установлено, что применение фильтрации к изображениям исходной серии показывает наилучшие результаты с точки зрения средних значений эталонных метрик качества на основе пикового отношения сигнал/шум (улучшение на 5% и 24% по сравнению с постфильтрацией и без фильтрации соответственно) и коэффициента структурного подобия (4% и 47% соответственно).

- Предложенный алгоритм на основе пирамид изображений имеет наибольший средний балл субъективной оценки (4, 5) среди ряда известных алгоритмов.
- На основании проведенного анализа времени работы модифицированного алгоритма можно сделать вывод о том, что он применим для задач формирования изображений с расширенной глубиной резкости в офлайн-приложениях.

Результаты работы внедрены в соответствующие разработки ООО «Энергия-Инфо» г. Ярославль и ООО «А-Реал Консалтинг» г. Ярославль. Отдельные результаты диссертационной работы внедрены в учебный процесс ЯрГУ им. П.Г. Демидова в рамках дисциплин «Цифровая обработка изображений», «Системы технического зрения», а также в научно-исследовательские работы при выполнении исследований в рамках грантов РФФИ (№ 15-08-99639 и № 16-37-00301). Получены три свидетельства о государственной регистрации программ для ЭВМ (№ 2014615039, № 2015617434, № 2016613017).

Достоверность полученных научных результатов обусловлена применением адекватного математического аппарата и подтверждается их согласованностью с результатами проведенного компьютерного моделирования.

Апробация работы. Результаты работы докладывались и обсуждались на следующих научно-технических конференциях:

- Семнадцатая международная конференция «Цифровая обработка сигналов и ее применение», Москва, 2015.
- Семнадцатая международная конференция открытой инновационной ассоциации FRUCT, Ярославль, 2015.
- Шестая научно-техническая конференция «Техническое зрение в системах управления 2015», Москва, 2015.

- Восемнадцатая международная конференция «Цифровая обработка сигналов и ее применение», Москва, 2016.
- Восемнадцатая международная конференция открытой инновационной ассоциации FRUCT, Санкт-Петербург, 2016.
- Двадцатая международная конференция открытой инновационной ассоциации FRUCT, Санкт-Петербург, 2017.
- Международная конференция «Системы синхронизации, формирования и обработки сигналов в инфокоммуникациях «СИНХРОИНФО», Казань, 2017.
- Ярославские региональные конференции молодых ученых и аспирантов.

Публикации. По теме диссертации опубликовано 15 научных работ, из них 3 статьи в журналах, рекомендованных ВАК, 3 статьи, индексируемые в SCOPUS, и 9 докладов на научных конференциях. Получено три свидетельства о регистрации программ для ЭВМ.

Структура и объем диссертации. Диссертация состоит из введения, трех глав, заключения, списка использованных источников, содержащего 105 наименований, и 3 приложений. Она изложена на 111 страницах машинописного текста, содержит 28 рисунков и 5 таблиц.

Основные научные положения, выносимые на защиту:

- Алгоритм формирования изображений с расширенной глубиной резкости на основе клеточного автомата.
- Модифицированный алгоритм формирования изображений с расширенной глубиной резкости с использованием пирамид изображений.
- Методика выбора метрик сфокусированности для задачи синтеза изображений с расширенной глубиной резкости.

- Рекомендации по применению шумоподавления при использовании модифицированного алгоритма формирования изображений с расширенной глубиной резкости в условиях наличия аддитивного белого гауссовского шума.

Благодарности. Автор выражает слова глубокой благодарности своему научному руководителю д.т.н. Приорову Андрею Леонидовичу за помощь на всех этапах выполнения данной работы. Отдельная благодарность доценту В.А. Волохову за консультации по теме работы. Спасибо всем коллегам по лаборатории «Цифровые цепи и сигналы» кафедры инфокоммуникаций и радиофизики ЯрГУ, возглавляемой профессором Ю.А. Брюхановым, среди которых особенно хочется отметить И.Н. Трапезникова и Ю.А. Лукашевича, чьи рекомендации позволили существенно улучшить данную работу. Также автор выражает глубокую благодарность родственникам и друзьям, поддерживавшим его на протяжении написания работы.

Глава 1

Обзор методов формирования изображений с расширенной глубиной резкости

1.1. Причины размытия изображений

Решение задачи формирования изображений с расширенной глубиной резкости ставит основной целью повышение информационной емкости серии исходных изображений. Области изображений, которые несут наименьшее количество полезной информации, зачастую представляют из себя участки, характеризующиеся как размытые. Чтобы эффективно решать задачу формирования изображений с расширенной глубиной резкости, следует сначала рассмотреть вопрос о возможных причинах размытия.

В зависимости от природы происхождения, различные виды размытия изображения могут быть классифицированы на четыре основных группы [7]:

1. связанные с движением;
2. связанные с оптическими особенностями устройств съемки;
3. связанные с изменчивостью среды съемки;
4. вызванные применением обработки изображения.

Примеры изображений, подвергнутых различного рода размытиям, приведены на рис. 1.1.

Движение камеры или объекта съемки во время экспозиции приводит к образованию размытия. Особенно часто этот эффект встречается в бытовых камерах. Например, камеру мобильного телефона достаточно сложно удерживать в одном положении, даже если время экспозиции достаточно не большое,



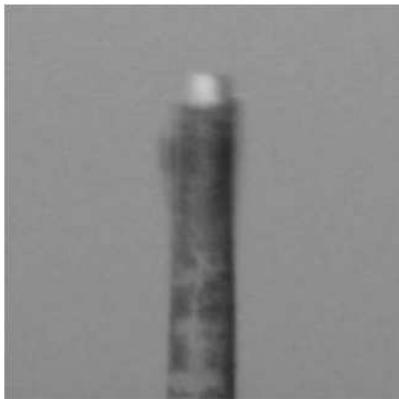
а)



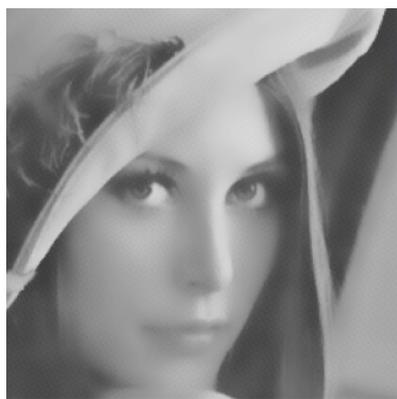
б)



в)



г)



д)



е)

Рис. 1.1. Различные источники образования артефактов размытия на изображениях: а) движение камеры относительно сцены; б) движение объекта в сцене; в) размытие в следствии расфокусировки; г) непостоянное размытие в следствии колебаний атмосферного воздуха; д) размытие, вызванное применением сглаживающих фильтров, например билатерального; е) размытие в следствии применения алгоритмов сжатия изображений

что, в конечном итоге, приводит к появлению данного вида искажений. Уменьшение времени экспозиции в некоторой степени может уменьшить этот эффект. Однако это приведет к искажениям другого рода, например, появлению шума в следствии повышения светочувствительности (*ISO*), так как уменьшится освещенность кадра.

В идеальном случае, если вся сцена неподвижна, а камера перемещается в одной плоскости, то размытие движением является пространственным инвариантом (одинаково на всем изображении, рис. 1.1 а), который может быть оценен

и устранен при помощи слепой обратной свертки [8–10]. Однако на практике движения, происходящие в сцене, являются более сложными. Может двигаться сама камера, объект в сцене, а также изменяться глубина сцены. Любой из этих факторов влечет за собой неравномерное пространственное искажение (например, рис. 1.1 б), задача оценки и устранения которого весьма сложная [11, 12].

Некорректная настройка оптической системы снимающего устройства, а также ограниченная глубина резкости приводят к размытию расфокусировкой (рис. 1.1 в). Даже в тех случаях, когда оптическая система настроена идеально, а снимаемая сцена находится в фокальной плоскости, большинство изображений имеют фундаментальную предельную границу резкости, обусловленную дифракцией [13]. С другой стороны, в большинстве используемых камер применяется антиалиазинговый фильтр, который удаляет высокочастотные компоненты изображения, превышающие порог Найквиста, определенный для данной модели оптической системы. Некоторые виды размытия могут помочь улучшить качество получаемых изображений. Размытия, обусловленные дифракцией или антиалиазингом, могут быть рассмотрены как пространственные инварианты, в то время как размытие расфокусировкой обусловлено двумя факторами: глубиной изображаемого пространства камеры и глубиной сцены [14, 15].

Для оптических систем, предназначенных для наблюдения на дальние дистанции (таких как астрономические), очень важным становится эффект атмосферной турбуленции, который порождает высокую изменчивостью коэффициента преломления на пути распространения светового потока и искажает его, что может существенным образом уменьшить качество изображения [16–20]. Турбуленция в других средах, таких как вода, также имеет схожий эффект, проявляющийся при подводной съемке [21]. Турбуленция в передающей среде, в общем случае, представляет собой комбинацию от двух искажений: геометрической деформации и размытия, каждое из которых меняется как в пространстве, так и во времени. Пример такого искажения приведен на рис. 1.1 г. Удаление

такого эффекта является очень важным для большого числа практических приложений и в настоящее время ему уделяется большое внимание [19, 21–24].

Применение различных методов обработки изображений, таких как шумоподавление, супер-разрешение и других, также может вносить искажения в виде размытия, в то время как большинство из них используется для повышения качества получаемых изображений (рис. 1.1 д). Одним из примеров может быть фильтр, который сглаживает изображение, чтобы уменьшить уровень шума, но имеет побочный эффект в виде размытия. Если фильтр работает одинаково на всем изображении, то получаемое размытие является пространственным инвариантом. Однако, на практике большинство фильтров являются локальными (билатеральный фильтр, NLM, BM3D), что позволяет с большей эффективностью обрабатывать наиболее информативные участки изображения [25–29]. Это приводит к тому, что размытие, вносимое ими, также является пространственно зависимым.

Процедуры сжатия, применяемые к изображениям, также вызывают эффекты размытия. Это, в основном, связано с удалением высокочастотных компонент сигнала. Пример такого изображения приведен на рис. 1.1 е [30, 31].

Для описания искажений, возникающих на изображении в следствии размытия расфокусировкой, воспользуемся моделью идеальной однолинзовой системы, к которой может быть сведено большинство реальных оптических систем. Схема модели изображена на рис. 1.2, здесь p — точечный объект, изображение которого строится в линзе с фокусным расстоянием f . Расстояние от объекта до линзы U , от линзы до изображения V . Формула тонкой линзы в таком случае определяется выражением 1.1

$$\frac{1}{f} = \frac{1}{U} + \frac{1}{V}. \quad (1.1)$$

Пусть камера имеет диафрагму размером D . Светочувствительная матрица камеры расположена на расстоянии S от линзы. Если $V \neq S$, изображение объекта на матрице будет представлять собой размытое пятно, радиус которого

исходит его смещение радиально от главной оптической оси системы. Однако, в большинстве случаев, данный эффект не оказывает существенного влияния [7].

1.2. Обзор методов смешивания изображений с расширенной глубиной резкости

Смешивание изображений — одно из приложений цифровой обработки сигналов. Основной задачей является комбинирование двух или более изображений в одно единственное, которое является более информативным и пригодным для визуального распознавания или цифровой обработки, чем исходные.

Русскоязычная терминология в данной области вычислительной фотографии еще не является устоявшейся. В ряде источников итоговое изображение называется «изображением с расширенной глубиной резкости», «полностью сфокусированным» или «полнофокусным», в других «мультифокусным» или «всюду сфокусированным», далее будем придерживаться первого варианта.

Разделяют два основных подхода к смешиванию изображений: первый — на основе трансформационных преобразований исходных изображений, во втором работа осуществляется непосредственно с пикселями изображений. По способу получения исходных изображений выделяют различные виды смешивания [7]:

1. объект снят с нескольких точек;
2. съемка панорамы с одной точки;
3. съемка с временными задержками;
4. съемка с разным фокусным расстоянием (мультифокус).

Техники смешивания на основе нескольких, снятых с различным фокусным расстоянием, изображений можно условно разделить на две группы: про-

странственные техники и техники с использованием трансформационных преобразований [32]. В первых — итоговое изображение состоит из пикселей исходных изображений, находящихся в соответствующих позициях. Во вторых — изображение переводится в некоторое пространство признаков. Смешивание изображений может быть разделено на четыре уровня по степени абстракции примитивов, с которыми осуществляется работа: уровень сигнала, уровень пикселя, уровень особенностей изображения, комбинированный уровень [33]. В смешивании на уровне сигнала, последовательности, приходящие с разных камер, смешиваются для получения нового сигнала, обладающего более высоким отношением сигнал/шум, чем исходные. На пиксельном уровне значение интенсивности каждого пикселя итогового изображения основано на значениях интенсивностей соответствующих пикселей исходных изображений. Уровень работы с особенностями подразумевает их извлечение или сегментацию на исходных изображениях, которые затем смешиваются в итоговое изображение. На комбинированном уровне применяются различные комбинации других уровней. После этого полученная информация обрабатывается в соответствии с некоторыми решающими правилами [1].

Однако, все эти техники вносят различного вида размытия и искажения на границах объектов на итоговом изображении. Ключевой задачей всех алгоритмов смешивания является борьба с этими явлениями.

1.3. Общий алгоритм смешивания изображений с расширенной глубиной резкости

Изображение с расширенной глубиной резкости представляет собой комбинацию нескольких изображений одного объекта, снятых с различным фокусным расстоянием.

Первым и самым важным шагом в подавляющем большинстве алгоритмов смешивания изображений с расширенной глубиной резкости является вычисле-

ние метрики сфокусированности исходных изображений или их отдельных областей. В работах зачастую используется один или несколько следующих методов оценки сфокусированности [2]:

1. гистограмма энтропии [34–36];
2. энергия градиента изображения [37, 38];
3. метрика Тененграда [39, 40];
4. пространственные частоты [41–43];
5. энергия Лапласиана [44];
6. M_2 -метрика [45–47];
7. вариация уровня серого [48–50];
8. отношение коэффициентов вейвлет-преобразования [51–53];
9. методы на основе дискретного косинусного преобразования [54, 55] (ДКП) и другие.

Пиксели с наибольшим значением метрики считаются находящимися в фокусе и выбираются для формирования итогового изображения. После выбора метрики сфокусированности необходимо определиться с правилом смешивания изображений. Например, выбор выделяющихся пикселей в некоторой области для выделения высокочастотной информации при помощи кратномасштабного анализа. Стандартная схема алгоритма смешивания изображений приведена на рис. 1.3.

В качестве иллюстрации различных существующих техник смешивания изображений с расширенной глубиной резкости рассмотрим ряд алгоритмов, позволяющих с различной эффективностью и качеством решить поставленную задачу.

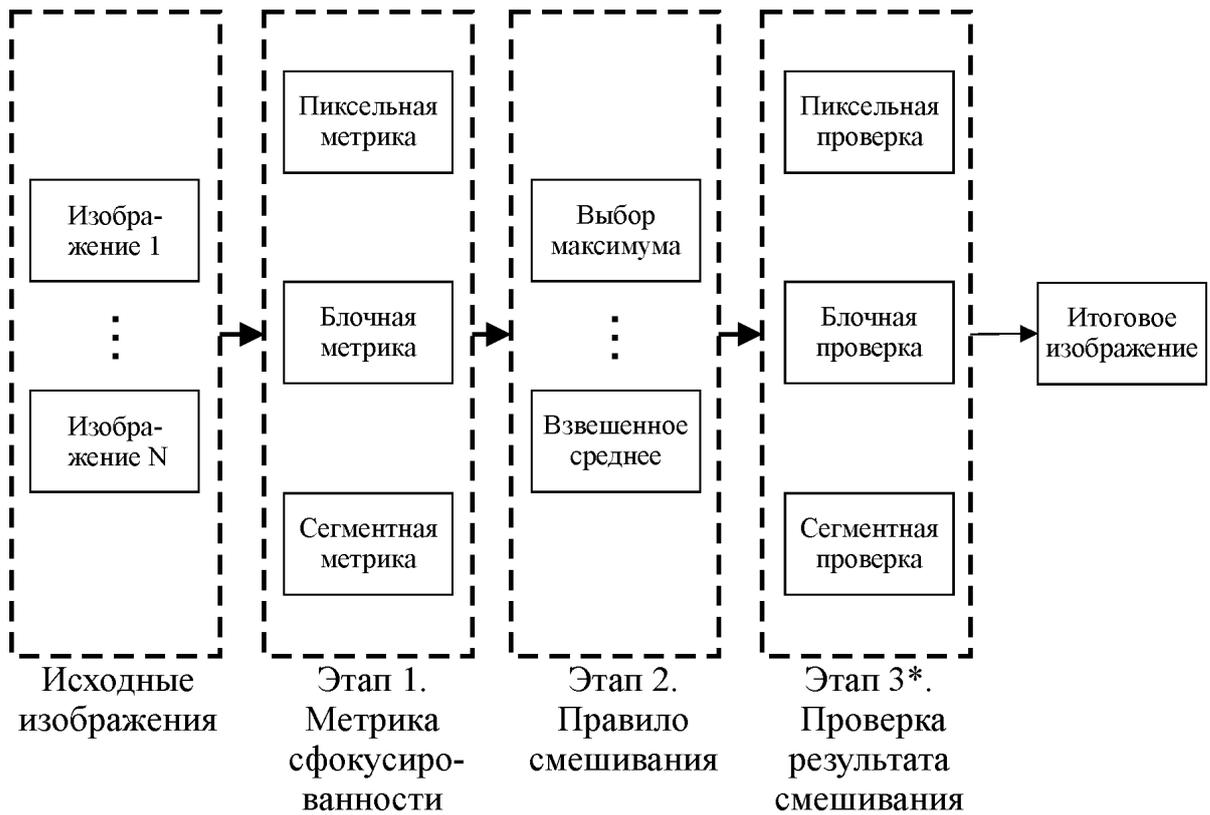


Рис. 1.3. Общая схема смешивания изображений с расширенной глубиной резкости

1.4. Пространственные методы смешивания

1.4.1. Смешивание на основе разбиения на блоки

Одним из самых простых и распространенных методов формирования изображений с расширенной глубиной резкости является смешивание на основе разбиения на блоки [7]. Метод предполагает, что все исходные изображения имеют одинаковый размер. Пусть исходные изображения имеют размер $M \times N$.

Алгоритм смешивания:

Алгоритм состоит из нескольких основных этапов, схема работы представлена на рис. 1.4.

Этап 1: Разбить каждое из исходных изображений на блоки размера $t \times n$.

Значения t и n являются параметрами алгоритма. Таким образом, каждое из исходных изображений будет разбито на $\lceil \frac{M}{t} \rceil * \lceil \frac{N}{n} \rceil$ блоков.

Этап 2: Рассчитать метрику сфокусированности каждого блока на каждом

исходном изображении.

Этап 3: Среди блоков с одинаковыми координатами, выбрать блок с наибольшей метрикой сфокусированности.

Этап 4: Сформировать конечное изображение, вставкой выбранного на Этапе 3 блока в соответствующее место на изображении.

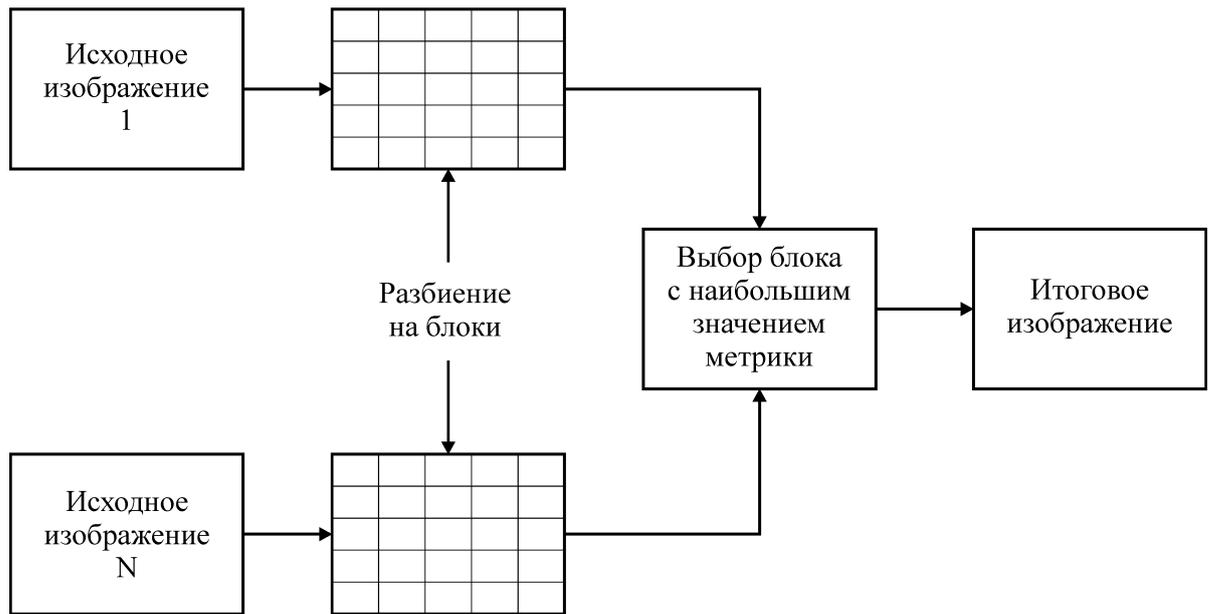


Рис. 1.4. Схема работы алгоритма смешивания изображений с расширенной глубиной резкости на основе разбиения на блоки

Данный алгоритм является простым и интуитивно понятным. Независимость обработки блоков позволяет использовать параллельные процессы для вычисления. Однако при использовании данного метода на границах блоков возникают существенные искажения. Также в некоторых сценах могут возникать ситуации, когда объект, который на одном из исходных изображений находится в фокусе, получается размытым, так как алгоритм не учитывает содержание изображения при разбиении на блоки. В практическом плане, алгоритм может быть использован как часть более сложной системы формирования изображений с расширенной глубиной резкости.

1.5. Трансформационные методы смешивания

1.5.1. Вейвлет-преобразование

На основе вейвлет-преобразования можно ввести новую статистическую меру резкости за счет использования распределения вейвлет-коэффициентов и оценить размытие изображения. Распределение вейвлет-коэффициентов оценивается при помощи локальной адаптивной модели смесей лапласианов [3]. Данная метрика позволяет выполнить смешивание в вейвлет-домене.

Алгоритм смешивания:

Вейвлет-разложение применяется к каждому исходному изображению. Для каждой подполосы детализирующих коэффициентов выполняется детализация при помощи локальной адаптивной модели смесей лапласианов, затем соответствующие коэффициенты объединяются с использованием некоторого правила. Далее применяется обратное вейвлет преобразование для получения итогового изображения. Если изображений больше чем два, то они складываются последовательно одно за другим по тому же алгоритму. Полученное таким образом изображение имеет высокое качество, по сравнению со многими другими методами. Однако данный подход имеет высокую вычислительную сложность.

1.5.2. Контурлет-преобразование без децимации

Контурлет-преобразование без децимации (КПБД) — полностью инвариантное к сдвигу, кратномасштабное расширение контурлет-преобразования для получения частотного разделения, комбинирует пирамиды и цифровые банки фильтров без прореживания [56]. В классическом контурлет-преобразовании пирамиды Лапласианов и банки фильтров используются для кратномасштабного анализа. В КПБД направленное разделение расположено таким образом, чтобы избежать наложения спектров, свойственного контурлет-преобразованию и достичь инвариантности к сдвигу.

Алгоритм смешивания:

На первом шаге происходит декомпозиция исходных изображений на низкочастотный и высокочастотные компоненты. Затем происходит вычисление метрик сфокусированности и определение наиболее четких регионов, таким образом, определяются коэффициенты смешивания компонент. Итоговое изображение получается путем выполнения обратного КПБД-преобразования.

Этот метод может успешно устранять артефакты на границах объектов.

1.5.3. Импульсная нейронная сеть

Импульсная нейронная сеть (ИНН), разработанная Экхорном (Eckhorn) в 1990 году, представляет собой нейронную сеть с обратными связями, каждый нейрон которой состоит из трех частей: рецепторного поля, модулирующего поля и генератора импульсов [57–59]. В ИНН каждый нейрон соответствует пикселю на входном изображении, входным значением является интенсивность пикселя. Выходные последовательности содержат информацию о входных изображениях, которая может быть использована для различных методов цифровой обработки изображений, таких как сегментация и извлечение отличительных особенностей.

По сравнению с традиционными подходами имеется ряд преимуществ, таких как устойчивость к шумам, независимость от геометрических преобразований входных данных, совместимость со слабосвязанными малыми перепадами интенсивности во входных данных. Этот метод основан на двухканальной ИНН. ИНН — это биологический аналог системы смешивания изображений [60]. Метод получен из стандартной модели ИНН путем упрощения процесса смешивания по сравнению с предыдущими методами. Обычно используется несколько ИНН или комбинация с другими алгоритмами, такими как дискретное вейвлет-преобразование (ДВП), в то время как данный метод использует только одну двухканальную ИНН для смешивания изображений. Схема работы данного метода представлена на рис. 1.5.

Данный метод выделяется среди других методов как с точки зрения визу-

альной эффективности, так и с точки зрения оценки работы. В практических приложениях метод более применим, так как быстро работает и может быть использован в системах реального времени. Однако существенной особенностью и сложностью является необходимость обучения нейронной сети, а также нестабильность работы на различных тестовых сценах.

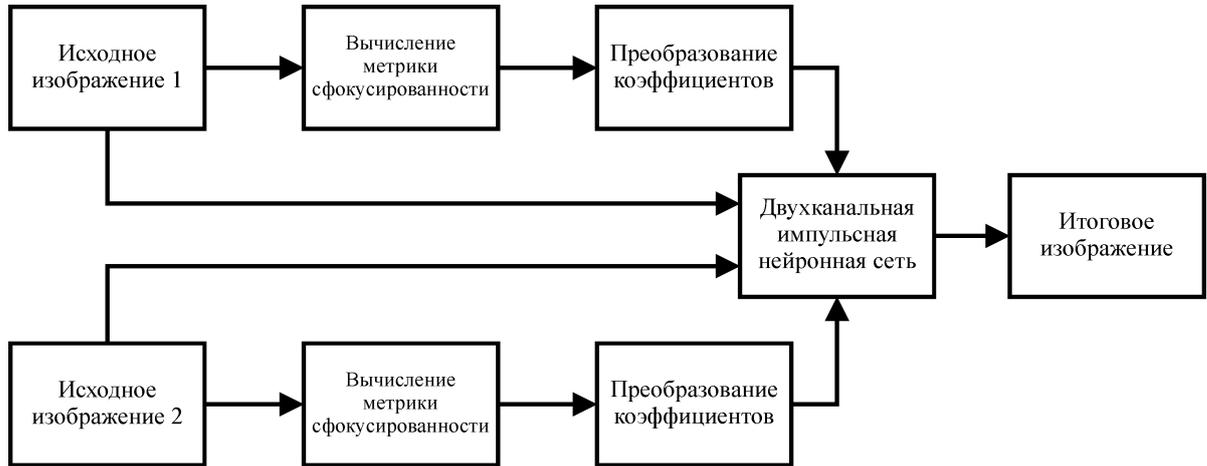


Рис. 1.5. Схема смешивания изображений с использованием двухканальной импульсной нейронной сети

1.6. Принципы работы клеточных автоматов

1.6.1. Понятие клеточного автомата

Клеточный автомат — дискретная модель, которая включает в себя несколько ключевых элементов и понятий [61]:

1. *Регулярная решетка ячеек.* Решетка может быть любой размерности. Формы решетки может быть различным, однако в большинстве приложений используется прямоугольная или гексагональная решетка. Каждая из ячеек может находиться в одном из конечного множества состояний $0, n - 1$. Во многих системах, описываемых клеточными автоматами, $n = 2$, таким образом каждая ячейка является бинарной, то есть может находиться в двух состояниях: 0 и 1.

2. *Окрестность.* Для каждой ячейки определено множество ячеек, называемых окрестностью. Две окрестности клетки имеют именные названия: это окрестность фон Неймана (рис. 1.6 а) — соседями являются только ближайшие клетки слева, справа, сверху и снизу, и окрестность Мура (рис. 1.6 б) — соседями являются все восемь окружающих клеток. В общем случае форма окрестности могут быть достаточно сложными, примеры таких окрестностей приведены на рис. 1.6 в и рис. 1.6 г.
3. *Исходное состояние.* Для начала работы необходимо определить начальное состояние всех ячеек решетки. Принципы задания состояний обычно определяются поведением моделируемой системы. Состояния могут быть распределены по решетке случайно, с заданным законом распределения, либо в соответствии с некой детерминированной закономерностью.
4. *Правило перехода.* На каждой итерации, используя правила перехода и состояния соседних ячеек, определяется новое состояние каждой ячейки. Правило перехода также может быть сколь угодно сложным и обычно определяется поведением моделируемой системы.

1.6.2. Математическое определение клеточного автомата

Клеточный автомат можно определить как множество конечных автоматов, каждый из которых может находиться в одном из $k + 1$ состояний:

$$\sigma \in \Sigma \equiv \{0, 1, 2, \dots, k - 1, k\}.$$

Изменение состояния клеточного автомата происходит согласно правилу перехода ϕ :

$$\sigma_{i,j}(t + 1) = \phi [\sigma_{k,l}(t) | \sigma_{k,l}(t) \in \Theta],$$

где Θ — множество автоматов, составляющих окрестность. Например, окрестность фон Неймана первого ранга определяется следующим образом:

$$\Theta_N^1(i, j) = \{\sigma_{k,l} | |i - k| + |j - l| \leq 1\}.$$

Число всех возможных правил перехода определяется числом состояний σ и количеством соседей n и составляет

$$N_r = \sigma^{\sigma^n}.$$

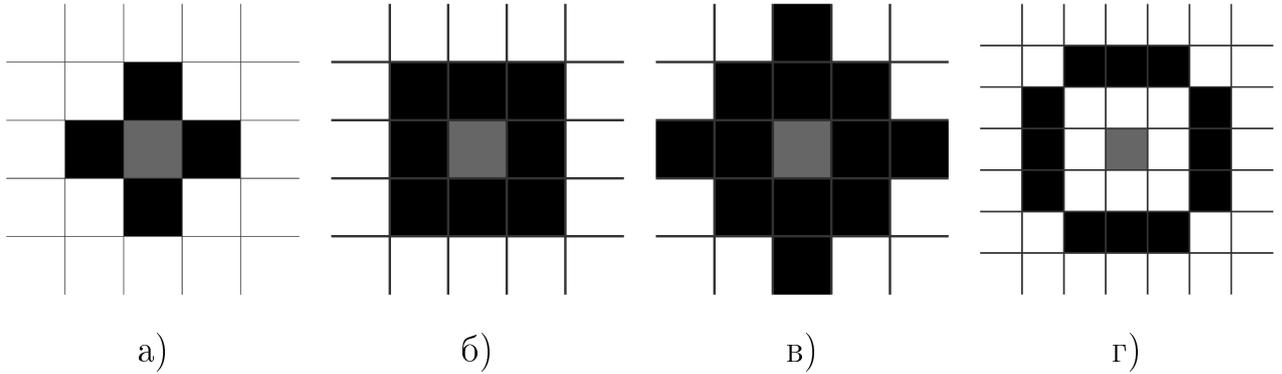


Рис. 1.6. Примеры окрестностей ячеек решетки клеточного автомата: а) окрестность фон Неймана; б) окрестность Мура; в,г) более сложные формы окрестностей

Основное направление исследования клеточных автоматов — алгоритмическая разрешимость тех или иных задач. Также рассматриваются вопросы построения начальных состояний, при которых клеточный автомат будет решать заданную задачу. Системы, описываемые клеточными автоматами, часто возникают в математике, теории вычислимости, физике, теоретической биологии, микромеханике, обработке изображений и других областях науки и инженерии [61, 62].

1.6.3. Свойства клеточных автоматов

Отметим несколько ключевых свойств клеточных автоматов [61]:

1. *Параллельность.* Свойство означает, что обновление состояния всех ячеек происходит одновременно на всей решетке, независимо от изменения окрестности на текущей итерации.

2. *Локальность*. Свойство характеризует независимость изменения состояния ячейки от состояния других ячеек, за исключением ее самой и ячеек окрестности.
3. *Однородность*. Свойство означает, что ко всем ячейкам применяется одно и то же правило.

С точки зрения реализации клеточных, эти свойства позволяют успешно применять SIMD-оптимизацию, например, с использованием графических процессоров.

1.6.4. Классификация правил клеточных автоматов

Стивен Вольфрам в своей книге *A New Kind of Science* [63] предложил 4 класса, на которые все клеточные автоматы могут быть разделены в зависимости от типа их эволюции. Классификация Вольфрама являлась первой попыткой классифицировать сами правила, а не типы поведения правил по отдельности. В порядке возрастания сложности классы выглядят следующим образом:

1. *Класс 1*. Результатом эволюции почти всех начальных условий является быстрая стабилизация состояния и его гомогенность. Любые случайные конструкции в таких правилах быстро исчезают.
2. *Класс 2*. Результатом эволюции почти всех начальных условий является быстрая стабилизация состояния, либо возникновение колебаний. Большинство случайных структур в начальных условиях быстро исчезает, но некоторые остаются. Локальные изменения в начальных условиях оказывают локальный характер на дальнейший ход эволюции системы.
3. *Класс 3*. Результатом эволюции почти всех начальных условий являются псевдо-случайные, хаотические последовательности. Любые стабильные

структуры, которые возникают почти сразу же уничтожаются окружающим их шумом. Локальные изменения в начальных условиях оказывают широкое, неопределяемое влияние на ход всей эволюции системы.

4. *Класс 4.* Результатом эволюции почти всех правил являются структуры, которые взаимодействуют сложным и интересным образом с формированием локальных, устойчивых структур, которые способны выживать длительное время. В результате эволюции правил этого класса могут получаться некоторые последовательности Класса 2, описанного выше. Локальные изменения в начальных условиях оказывают широкое, неопределяемое влияние на ход всей эволюции системы. Некоторые клеточные автоматы этого класса обладают свойством универсальности по Тьюрингу. Последний факт доказан для Правила 110 и игры «Жизнь».

Такого рода определения носят по большей части качественный характер и их можно по разному интерпретировать. Однако, практически при всякой попытке классификации будут возникать ситуации, когда по одному свойству предмет можно отнести к одному классу, а какому-либо другому свойству — к другому классу. Такая же ситуация и с клеточными автоматами: встречаются правила, которые показывают свойства, присущие одновременно одному и другому классу.

Использование клеточных автоматов во многих дискретных конечных системах позволяет получать хорошие оценки тех или иных параметров этих систем. В виду того, что изображение представляет собой такую систему, а пиксели описываются как конечные автоматы, применение клеточных алгоритмов позволяет построить последовательность действий, которая решает поставленную задачу формирования изображений с расширенной глубиной резкости.

1.7. Краткие выводы

В первой главе рассмотрена классификация и причины возникновения эффектов размытия изображений, в том числе ограничение глубины резко изображаемого пространства. Приведен обзор существующих алгоритмов формирования изображений с расширенной глубиной резкости. Рассмотрен математический аппарат клеточных автоматов, использующийся в разработанном алгоритме. Клеточные автоматы представляют собой удобную математическую модель, позволяющую решить задачу формирования изображений с расширенной глубиной резкости с учетом контента изображений. Особенностью данного математического аппарата является возможность эффективного применения параллельного программирования и SIMD-оптимизации.

Глава 2

Алгоритм смешивания изображений с расширенной глубиной резкости на основе клеточного автомата

2.1. Тестовые серии исходных изображений

Для анализа и оценки качества работы алгоритмов формирования изображений с расширенной глубиной резкости в рамках диссертации используются несколько серий изображений. Каждая из этих серий содержит сцену, насыщенную областями с большим количеством деталей. В свою очередь, эти области располагаются на различном расстоянии от снимающего устройства.

Снимающее устройство представляет собой цифровую зеркальную фотокамеру Sony Alpha 37 с объективом Sony 50mm 1.8/ f (фиксированный SAL 50F18).

Характеристики серий изображений представлены в табл. 2.1

Таблица 2.1. Характеристики серий исходных изображений

Характеристика	Значение			
	Название серии	«Toys»	«Robot»	«Soldiers»
Формат файла	.jpg			
Цветовое пространство	RGB			
Глубина цвета	8 бит/канал			
Разрешение	786 × 523	786 × 523	786 × 523	892 × 592
Количество изображений	5	8	8	7

Примеры изображений данных исходных серий приведены в приложении А. Серии носят условные названия «toys», «robot», «soldiers» и «numbers», которые

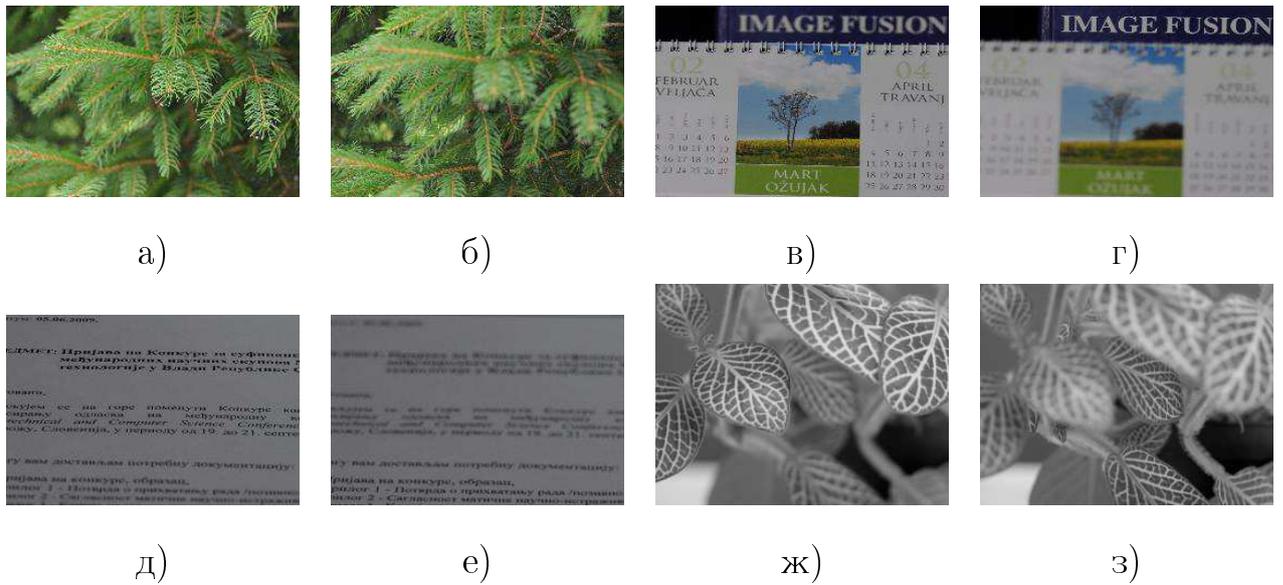


Рис. 2.1. Примеры изображений из заимствованных тестовых серий: а, б) «tree»; в, г) «calendar»; д, е) «text»; ж, з) «foliage»

отражают содержание сцены.

Помимо собственных серий изображений, описанных выше, также использовались несколько серий изображений, часто рассматриваемых в рамках англоязычных статей, касающихся вопросов смешивания изображений. Стоит отметить, что все заимствованные серии имеют небольшой размер, порядка 300×200 пикселей, и по два изображения в каждой серии. Примеры серий изображений приведены на рис. 2.1.

2.2. Предварительная обработка изображений исходной серии

Перед началом решения непосредственно задачи формирования изображений с расширенной глубиной резкости, зачастую полезно провести некоторую предварительную обработку, которая позволит в дальнейшем достичь более качественных результатов.

Одной из проблем, являются шумы, возникающие на изображениях по различным причинам. Также возможен линейный сдвиг изображений исходной

серии относительно друг друга на небольшие дистанции. Уменьшить влияние данных факторов можно применив методы обработки изображений.

2.2.1. Предфильтрация исходных изображений

Работа разработанного алгоритма формирования изображений с расширенной глубиной резкости в условиях наличия шумов может быть улучшена при помощи методов шумоподавления. В разделе 3.1 представлено исследование, подтверждающее, что предфильтрация изображений исходной серии способна повысить качество конечного изображения.

2.2.2. Линейный сдвиг кадров

При съемке изображений исходной серии могут возникнуть эффекты сдвига оптической оси. Наличие данного вида искажений может быть объяснено, например, дрожанием снимающего устройства в процессе съемки. Это приводит к наличию участков сцены, которые присутствуют не на всех снимках серии исходных изображений. Фазовая корреляция представляет собой способ оценки относительного сдвига пары изображений относительно друг друга, содержащих одну и ту же сцену, а также других множеств данных. Подход часто используется для размещения изображений в одной системе координат. Предполагается представление данных в частотной области, для вычисления которого обычно применяют дискретное преобразование Фурье.

Рассмотрим алгоритм вычисления относительного сдвига двух изображений A и B при помощи метода фазовой корреляции.

1. Применим к обоим изображениям оконное преобразование, позволяющее уменьшить влияние краевых эффектов:

$$g_a = W(A),$$

$$g_b = W(B).$$

Примером может служить окно Хэмминга, имеющее вид

$$w(n) = \alpha - \beta \cos\left(\frac{2\pi n}{N-1}\right),$$

здесь $\alpha = 0,54$, $\beta = 1 - \alpha = 0,46$ [64].

2. Вычислим двумерное дискретное преобразование Фурье от каждого изображения:

$$G_a = F(g_a),$$

$$G_b = F(g_b).$$

3. Вычислим совместную спектральную плотность мощности R :

$$R = \frac{G_a \circ G_b^*}{|G_a \circ G_b^*|},$$

здесь $X_1 \circ X_2$ — произведение Адамара [65], представляющее собой поэлементное перемножение матриц X_1 и X_2 , X^* — комплексно-сопряженная [66] к X матрица.

4. Вычислим кросс-корреляцию r путем применения обратного дискретного преобразования Фурье:

$$r = F^{-1}(R).$$

5. Найдем максимум r , который определяет сдвиги изображений относительно друг друга по осям:

$$(\Delta x, \Delta y) = \underset{x,y}{\operatorname{argmax}}(r).$$

Вычислив сдвиги всех изображений исходной серии $(\Delta x, \Delta y)_i$ относительно первого изображения и наложив их друг на друга, получим область которая перекрыта всеми изображениями исходной серии. Здесь можно рассмотреть два различных алгоритма действия. Выбрать для дальнейшей работы область, которая:

а) перекрыта всеми изображениями исходной серии;

б) перекрыта, как минимум, одним из изображений исходной серии;

с учетом особенностей большинства алгоритмов формирования изображений с расширенной глубиной резкости более предпочтительным является первый вариант.

На рис. 2.2 приведены примеры изображений, смещенные относительно друг друга на известные величины (в пикселях), а также рассчитанные приведенным способом сдвиги, с использованием окна Хэмминга размера 7.



а)



б)



в)



г)

Рис. 2.2. Примеры изображений, смещенных относительно друг друга: а) действительный сдвиг $(0, 0)$, рассчитанный сдвиг $(1, 1)$; б) действительный сдвиг $(33, 37)$, рассчитанный сдвиг $(32, 36)$; в) действительный сдвиг $(81, -1)$, рассчитанный сдвиг $(81, 1)$; г) действительный сдвиг $(-45, 64)$, рассчитанный сдвиг $(-47, 65)$

Таким образом корректировка взаимного положения изображений исходной серии относительно друг друга может быть осуществлена с точностью до единиц пикселей при помощи метода, основанного на фазовой корреляции. Выполнение данной операции приведет к потере части исходных данных, что в общем случае следует учитывать при построении системы формирования изображений с расширенной глубиной резкости.

2.3. Оценка сфокусированности областей изображений исходной серии

2.3.1. Постановка задачи оценки сфокусированности изображения

Задача оценки глубины резкости изображения является одной из ключевых в вычислительной фотографии вообще и в задаче формирования изображений с расширенной глубиной резкости в частности. Данная проблема возникает в момент перехода от трехмерного восприятия к двумерной проекции на изображении.

Если невозможно никакое физическое взаимодействие с отснятой сценой, глубина изображения может быть восстановлена при помощи бинокулярных (тринокулярных) систем, а также при помощи нескольких кадров, сделанных при различных настройках монокулярной системы.

Оценка глубины изображения лежит в основе многих важных приложений вычислительной фотографии, например, в задаче восстановления формы объекта. Также, свое применение данная задача находит в областях, связанных с управлением роботами, неразрушающем контроле изделий, реконструкции моделей и прочих сферах промышленности.

В работе оценка сфокусированности некоторой области изображения производится для определения наиболее информативных пикселей изображения, которые в дальнейшем будут играть более важную роль при формировании

изображений с расширенной глубиной резкости.

Все способы оценки сфокусированности изображений или их различных областей можно разделить на несколько больших групп. Схема деления приведена на рис. 2.3.

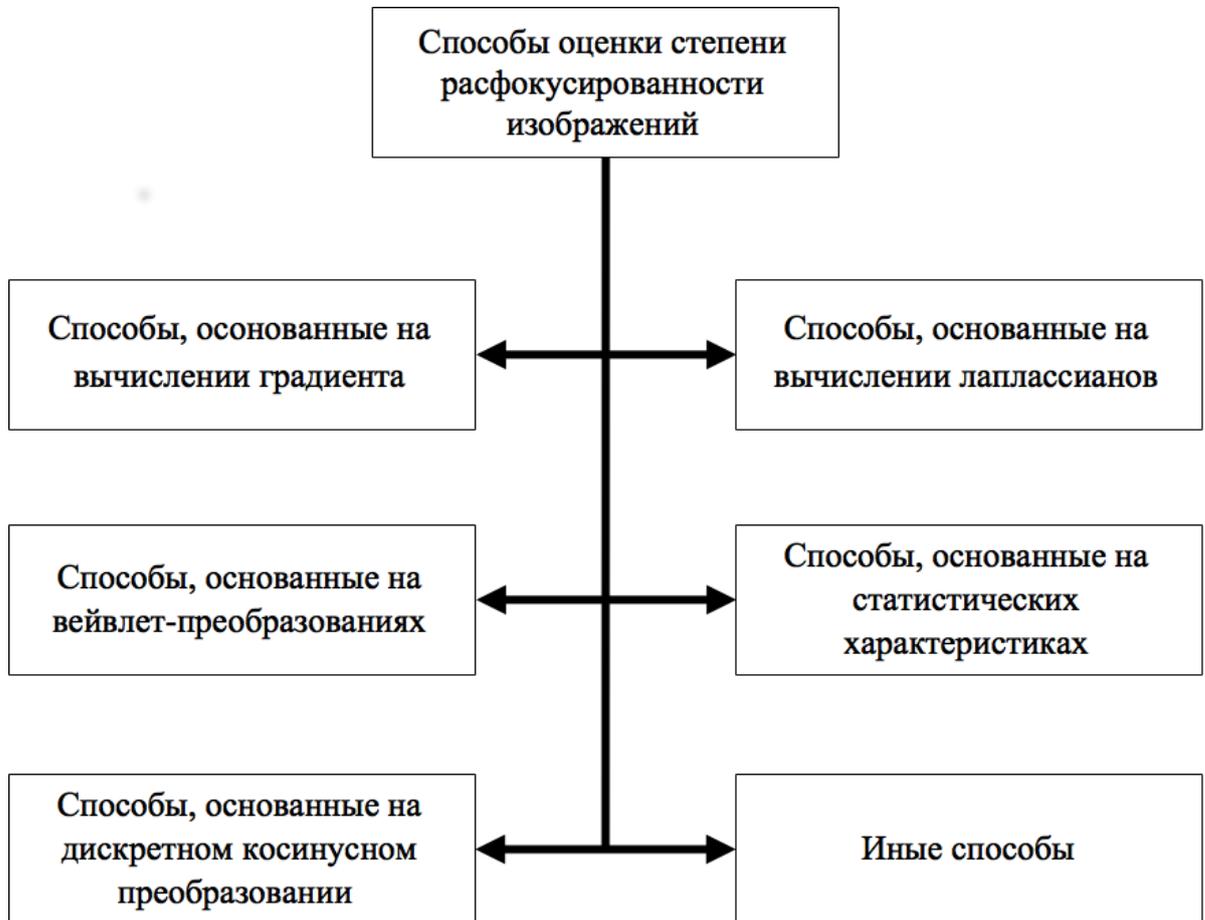


Рис. 2.3. Классификация способов оценки сфокусированности изображений и их областей

Рассмотрим каждую группу более подробно.

1. *Способы, основанные на вычислении градиента.* Данная группа методов базируется на вычислении градиента изображения или первой производной от изображения. Основная идея в методах этой группы, заключается в утверждении, что чем больше выражены границы в данной области, тем она более сфокусирована. Таким образом, значение метрики больше на резких областях чем на размытых.

2. *Способы, основанные на вычислении лапласианов.* Также как и предыдущая группа основана на идее выделения границ, однако, с той разницей, что вычисляются они при помощи второй производной.
3. *Способы, основанные на вейвлет-преобразованиях.* Данная группа способов оценки сфокусированности основана на способности коэффициентов дискретного вейвлет-преобразования описывать пространственные и частотные характеристики изображения. Таким образом, коэффициенты преобразования и их различные соотношения могут быть использованы для получения оценок сфокусированности.
4. *Способы, основанные на статистических характеристиках.* Методы этой группы основаны на вычислении статистических характеристиках структуры изображения.
5. *Способы, основанные на дискретном косинусном преобразовании.* Также как и методы основанные на вейвлет-преобразовании, используют коэффициенты соответствующего преобразования для вычисления пространственных и частотных характеристик изображения в целях оценки сфокусированности.
6. *Иные способы.* В данную группу отнесены методы, которые не вошли ни в одну из предыдущих. Как правило, они основаны на каких-либо априорных данных о снимаемой сцене, и применяются в случае изображений со специализированным содержанием.

2.3.2. Способы оценки сфокусированности изображения

Для выбора наилучшего метода оценки сфокусированности в рамках работы проведено исследование, целью которого являлось выявление наиболее подходящего алгоритма оценки сфокусированности.

Так как предлагаемый алгоритм, основанный на клеточных автоматах, предполагает наличие оценки сфокусированности каждого пикселя изображения, то наиболее предпочтительными выглядят методы, основанные на свертках исходных изображений с различного рода масками. Однако, для общности, в исследование включены несколько метрик, которые используют статистические и другие алгоритмы.

1. *Абсолютный центральный момент.* Метрика предложена в [67]. Основана на статистических оценках и гистограмме изображения H .

$$\text{АЦМ} = \sum_{k=1}^L |k - \mu| P_k,$$

где μ — среднее значение H , L — количество уровней серого на изображении, P_k — относительная частота k -го уровня серого.

2. *Метрика Тененграда (маска Собеля).* Распространенная метрика оценки сфокусированности, основанная на измерении амплитуды градиента яркости изображения

$$\text{Тененград (Собель)} = \sum_{(i,j) \in \Omega(x,y)} (G_x(i,j)^2 + G_y(i,j)^2),$$

где Ω — область оценки сфокусированности, G_x и G_y — вертикальный и горизонтальный градиенты, вычисленные путем свертки исходных изображений с соответствующими масками Собеля, имеющими следующий вид:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

3. *Метрика Тененграда (маска Робертса).* Метрика аналогична вышеописанной, однако основанная на масках Робертса, имеющих вид:

$$S_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, S_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

4. *Метрика Тененграда (маска Превитт)*. Метрика аналогична вышеописанной, однако основанная на масках Превитт, имеющих вид:

$$S_x = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, S_y = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

5. *Вариация уровня серого (ВУС)*. Одна из самых широко распространенных метрик оценки сфокусированности изображений, применяемая в задачах автофокусировки и восстановления формы объектов, определяется следующим образом:

$$\text{ВУС} = \sum_{(i,j) \in \Omega(x,y)} (I(i,j) - \mu)^2,$$

где μ — среднее значение уровня серого по всей области Ω оценки сфокусированности.

6. *Отношение коэффициентов вейвлет-преобразования (ОКВП)*. Метрика предложена в [68], предполагает использование отношения высокочастотных M_H и низкочастотных M_L коэффициентов вейвлет-преобразования. Метрика определена следующим образом

$$\text{ОКВП} = \frac{M_H^2}{M_L^2},$$

где

$$M_H^2 = \sum_k \sum_{(i,j) \in \Omega} W_{LHk}(i,j)^2 + W_{HLk}(i,j)^2 + W_{HHk}(i,j)^2,$$

$$M_L^2 = \sum_k \sum_{(i,j) \in \Omega} W_{LLk}(i,j)^2,$$

здесь $W_{LHk}(i,j)$, $W_{HLk}(i,j)$ и $W_{HHk}(i,j)$ — детализирующие коэффициенты вейвлет-преобразования, $W_{LLk}(i,j)$ — аппроксимирующие коэффициенты вейвлет-преобразования, k — количество уровней вейвлет-преобразования.

2.3.3. Экспериментальное сравнение алгоритмов оценки сфокусированности изображений

Для численной оценки избирательных свойств различных метрик сфокусированности разработана программа экспериментальных исследований и организован экспериментальный стенд.

Экспериментальный стенд

Экспериментальный стенд состоит из устройства фотофиксации (1), набора опорных меток, содержащих высоконтрастные изображения (2) и экспериментального изображения (3). Общая схема представлена на рис. 2.4.

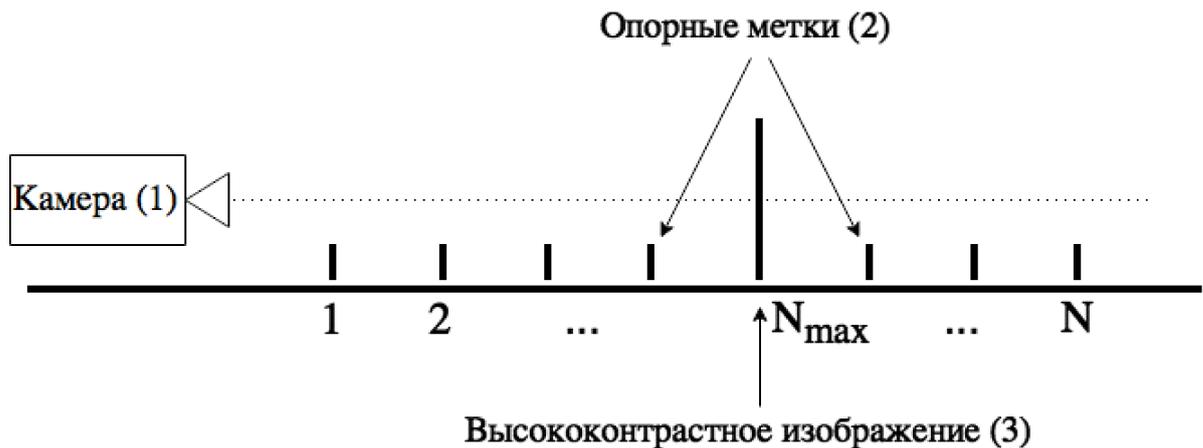


Рис. 2.4. Области тестовых последовательностей

В качестве устройства фотофиксации используется фотокамера Sony Alpha 37 с объективом Sony 50mm 1.8/ f . Опорные точки необходимы для установки фокальной плоскости устройства фотофиксации в определенной позиции, таким образом, что расстояние между двумя ее последовательными положениями оставалось неизменным. Экспериментальное изображение представляет собой телевизионную испытательную таблицу ТИТ-0249. Экспериментальное изображение представлено на рис. 2.5.

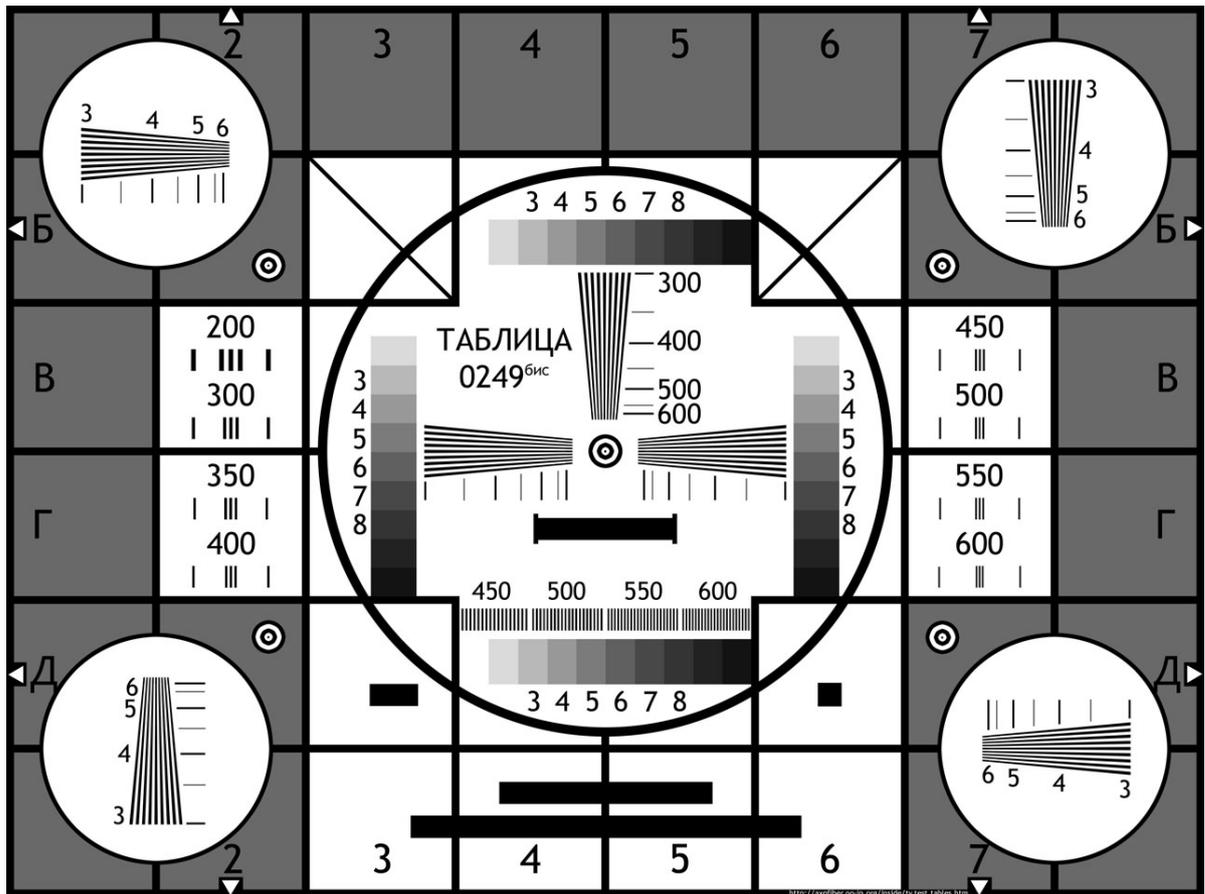


Рис. 2.5. Экспериментальное изображение: телевизионная испытательная таблица ТИТ-0249

Порядок проведения эксперимента

- Этап 1.** Формирование четырех тестовых последовательностей, состоящих из нескольких изображений каждая, в соответствии с выбранными областями.
- Этап 2.** Оценка, полученных на этапе 1, изображений с использованием каждой из рассматриваемых метрик.
- Этап 3.** Формирование и вычисление метрики сравнения качества работы метрик сфокусированности.

Для проведения исследования по первому этапу, создана расширенная тестовая последовательность, состоящая из 16 изображений, с различным положением фокальной плоскости. Наиболее четкое изображение расположено на

расстоянии 70 сантиметров и имеет номер 9. Шаг изменения положения фокальной плоскости снимающего устройства выбран равным пяти сантиметрам, начиная с 30 сантиметрового отступа от снимающего устройства.

В целях набора различных тестовых последовательностей выбраны области снимаемой сцены с различным содержанием. Выбранные области и соответствующие им тестовые последовательности изображены на рис. 2.6.

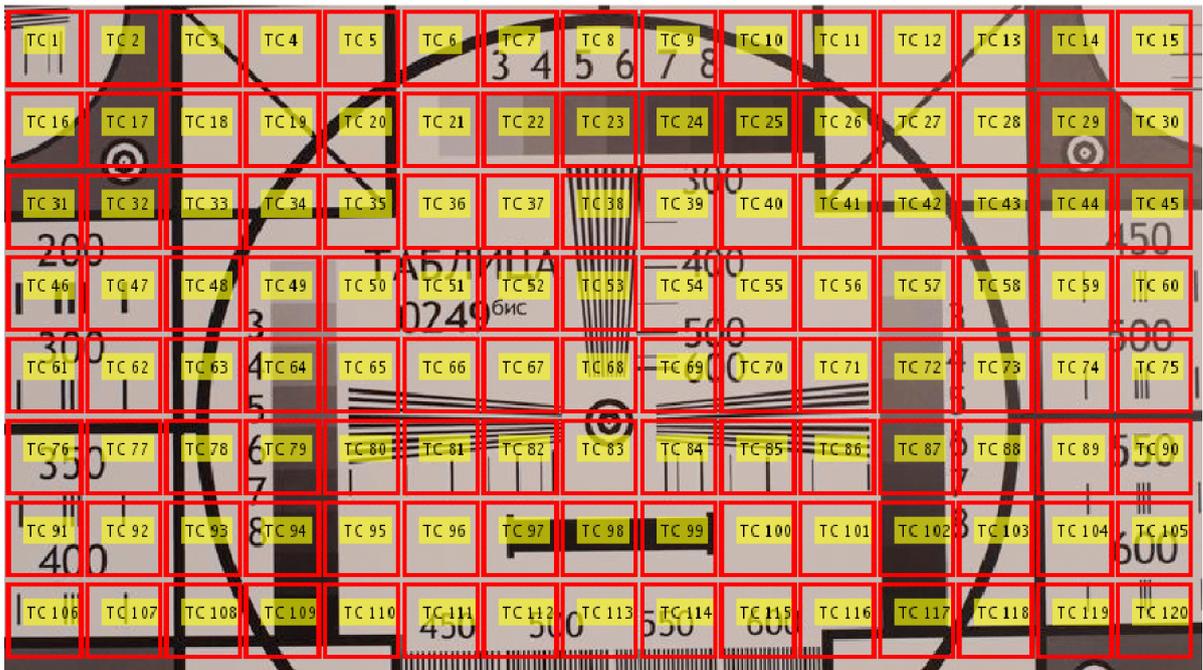


Рис. 2.6. Разделение снимаемой сцены на тестовые серии

Во втором этапе эксперимента проведены оценки сфокусированности выбранных областей при помощи анализируемых алгоритмов. Вычисленные значения метрик нормированы на максимальное значение метрики в текущей серии изображений для приведения к диапазону $[0; 1]$. Визуализация некоторых полученных значений метрик сфокусированности приведены на рис. 2.7.

Для численного сравнения различных алгоритмов оценки сфокусированности изображения в диссертации предлагается использовать значение коэффициента корреляции Пирсона между снятой характеристикой X и некоторой идеальной характеристикой Y . Коэффициент корреляции вычисляется соглас-

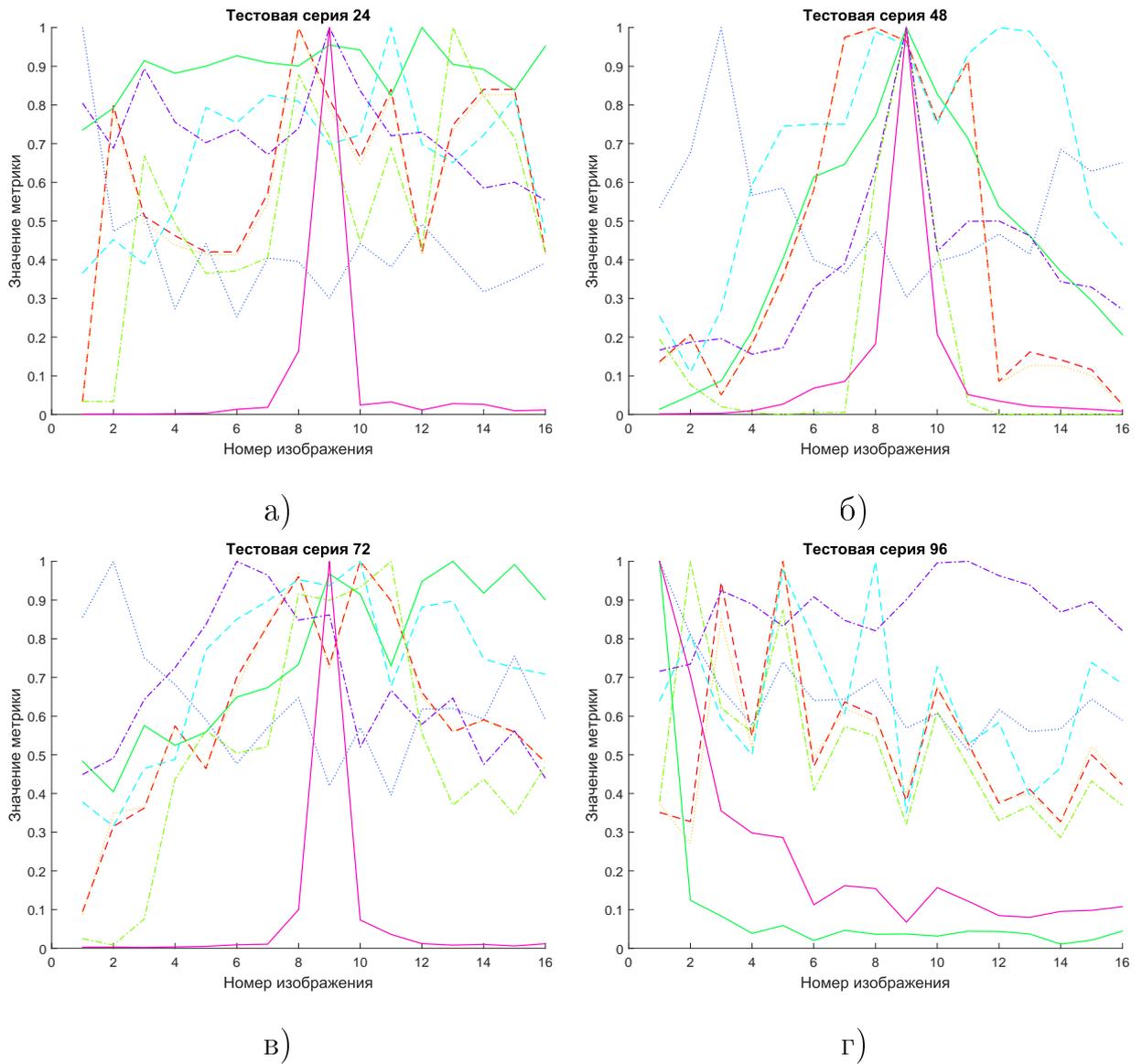


Рис. 2.7. Нормированные значения метрик сфокусированности для тестовых последовательностей: а) тестовая последовательность 24; б) тестовая последовательность 48; в) тестовая последовательность 72; г) тестовая последовательность 96

но выражению 2.1.

$$r = \frac{\sum_n [(x_n - \bar{x})(y_n - \bar{y})]}{\sqrt{\sum_n (x_n - \bar{x})^2 \sum_n (y_n - \bar{y})^2}}, \quad (2.1)$$

где $x_n = X(n)$, $y_n = Y(n)$, \bar{x} и \bar{y} — средние значения X и Y соответственно.

В рамках диссертации предложены следующие формы идеальной характеристики:

1. **На основе единичного импульса.** Идея идеальной характеристики на

основе единичного импульса базируется на предположении о том, что наилучшая метрика должна иметь значение 1 при оценке наиболее сфокусированного изображения N_{max} и 0 при оценке остальных изображений.

График характеристики приведен на рис. 2.8.

$$FM(n) = \begin{cases} 1, & \text{если } n = N_{max} \\ 0, & \text{иначе} \end{cases}.$$

2. **На основе формулы тонкой линзы. Обратная линейная зависимость.** Идеальная характеристика на основе тонкой линзы основана на физически наблюдаемых аспектах динамики изменения сфокусированности изображений. Это подтверждается формулой тонкой линзы 1.2, из которой следует, что радиус пятна нечеткого изображения точки прямо пропорционален удалению фокальной плоскости от плоскости матрицы камеры. Из этого следует, что если значение метрики на идеально сфокусированном изображении равно 1, то оно будет уменьшаться при удалении от плоскости объекта по закону $\frac{1}{x}$.

График характеристики приведен на рис. 2.8.

$$FM(n) = \begin{cases} \frac{1}{-n+N_{max}+1}, & \text{если } n < N_{max} \\ \frac{1}{n-N_{max}+1}, & \text{иначе} \end{cases}.$$

3. **На основе формулы тонкой линзы. Обратная квадратичная зависимость.** Данная форма идеальной характеристики аналогична предыдущей, за исключением того, что в качестве критерия расфокусированности используется не радиус пятна размытия, а его площадь, что дает обратный квадратичный закон.

График характеристики приведен на рис. 2.8.

$$FM(n) = \begin{cases} \frac{1}{(-n+N_{max}+1)^2}, & \text{если } n < N_{max} \\ \frac{1}{(n-N_{max}+1)^2}, & \text{иначе} \end{cases}.$$

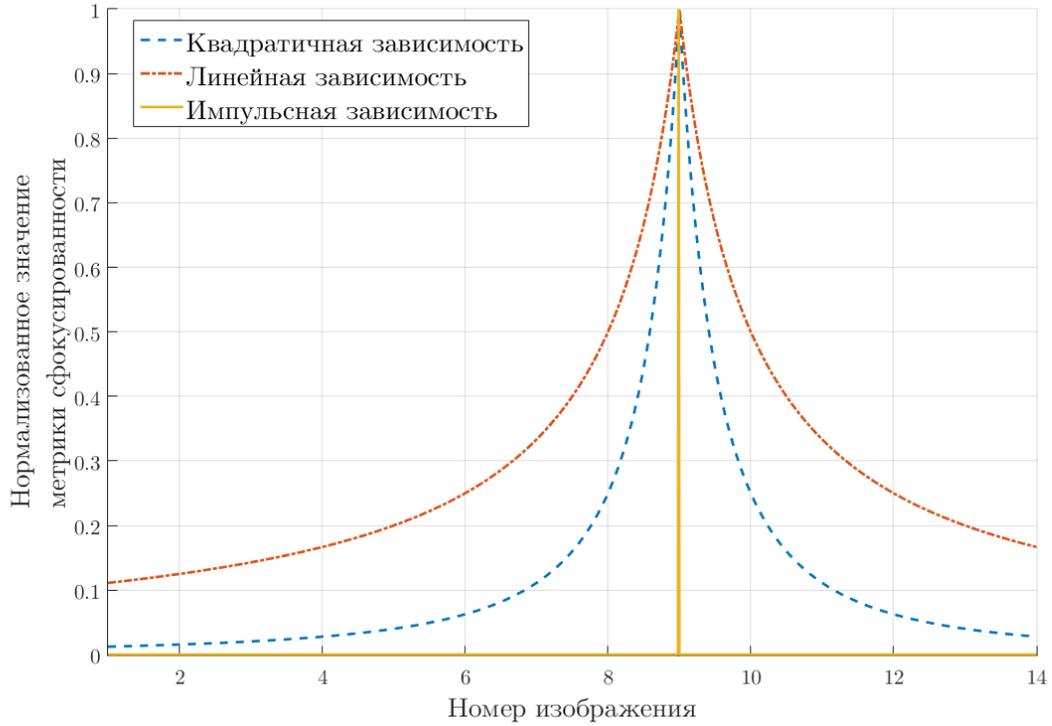


Рис. 2.8. Идеальные характеристики метрик сфокусированности

Результаты моделирования

В таблице 2.2 приведены численные значения среднего взвешенного коэффициента корреляции r для предложенных идеальных характеристик, вычисленные по следующей формуле 2.2:

$$r = \sum_{k=1}^N w_k r_k, \quad (2.2)$$

где N — общее число тестовых последовательностей r_k — коэффициент корреляции в k -ой тестовой последовательности, w_k — весовой коэффициент k -ой тестовой последовательности, вычисляемый по формуле $w_k = \frac{n_k}{NI}$, где n_k — количество изображений в k -ой последовательности, NI — общее количество изображений во всех последовательностях.

На рис. 2.9 приведены значения среднего взвешенного коэффициента корреляции в виде столбчатой диаграммы и доверительных интервалов при уровне значимости 0,95.

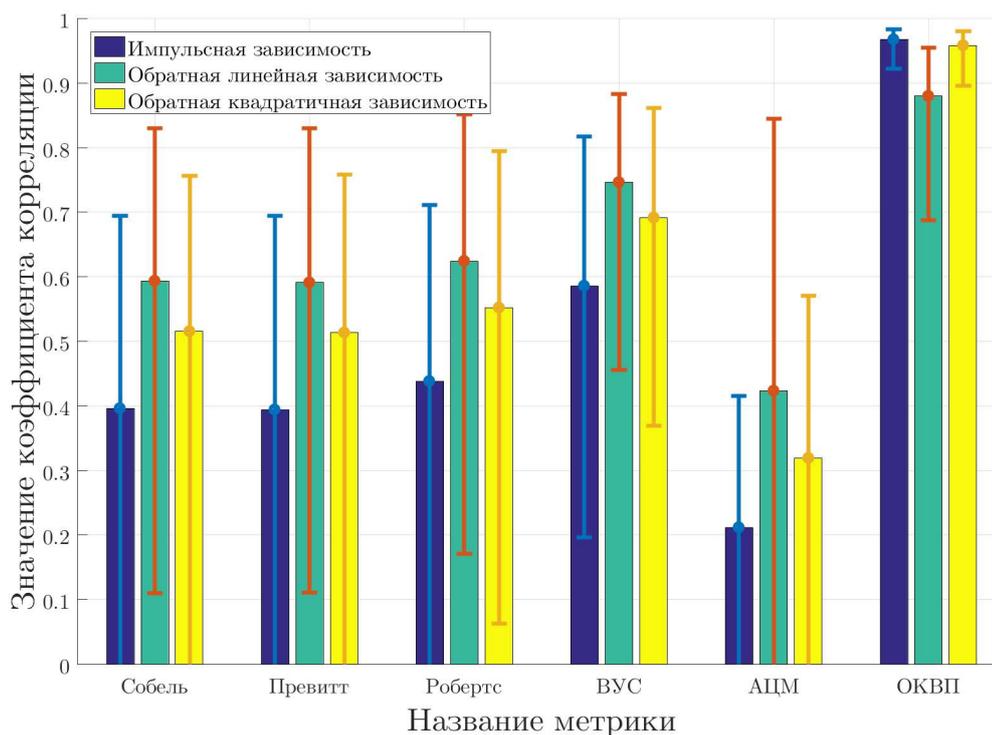


Рис. 2.9. Значения среднего взвешенного коэффициента корреляции

Таблица 2.2. Среднее взвешенное значение коэффициента корреляции r для предложенных идеальных характеристик

Название метрики	На основе единичного импульса	На основе формулы тонкой линзы (радиус)	На основе формулы тонкой линзы (площадь)
Тененград (Робертс)	0,39	0,59	0,51
Тененград (Собель)	0,39	0,59	0,51
Тененград (Превитт)	0,43	0,62	0,55
ВУС	0,58	0,74	0,69
АЦМ	0,21	0,42	0,31
ОКВП	0,96	0,87	0,95

На рис. 2.10 приведена столбцовая диаграмма, отражающая среднее время вычисления метрик сфокусированности в блоке конкретного размера.

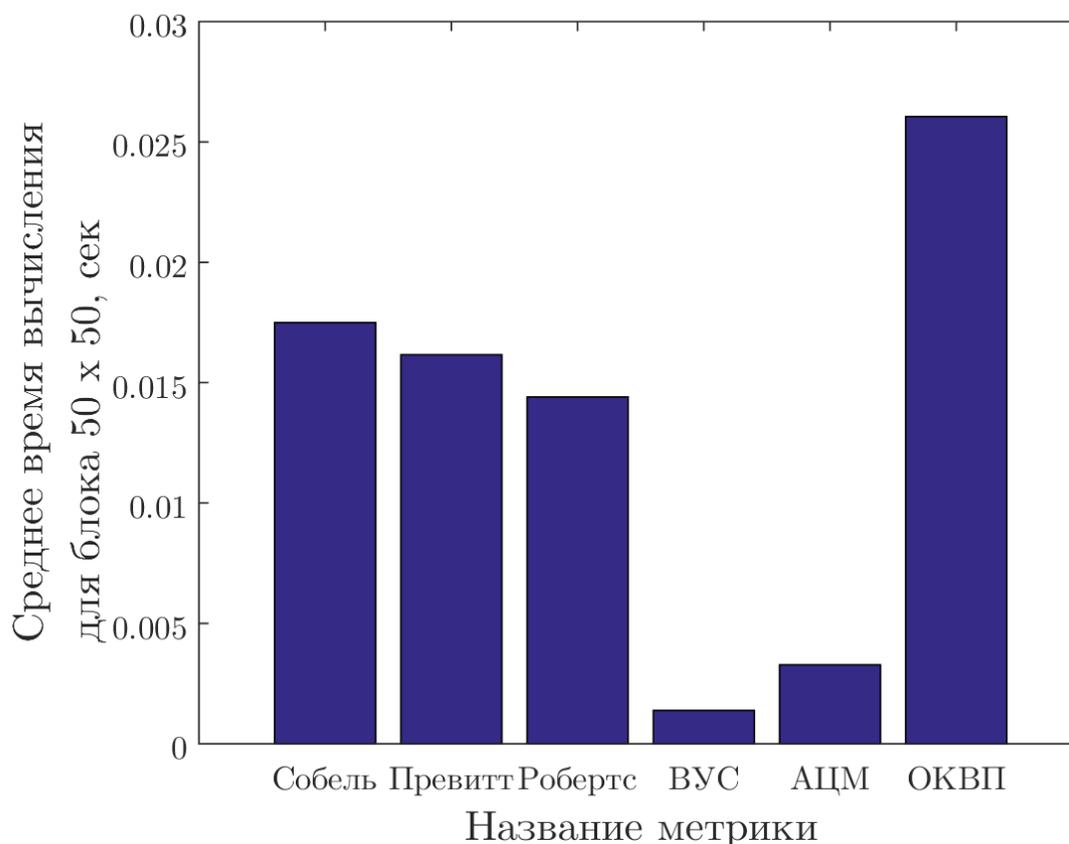


Рис. 2.10. Средние значения времени вычисления метрик сфокусированности для области определенного размера

Основываясь на результатах проведенного эксперимента, можно сделать ряд выводов, позволяющих выдать следующие рекомендации по использованию методов оценки сфокусированности:

1. Из анализа значений коэффициентов корреляции для различных способов оценки сфокусированности и способов задания идеальной характеристики можно заключить, что выбор той или иной идеальной характеристики не играет существенной роли, так как они ведут себя схожим образом в рамках данного эксперимента. Действительно, если рассчитать матрицу корреляции для полученных значений метрик, то она будет выглядеть

следующим образом:

$$\begin{bmatrix} 1 & 0,92 & 0,98 \\ 0,92 & 1 & 0,98 \\ 0,98 & 0,98 & 1 \end{bmatrix},$$

что свидетельствует о наличии сильной положительной связи. Однако при использовании предложенного подхода для исследования иных метрик сфокусированности, может быть не истинным.

2. Из рассмотренных алгоритмов, наиболее точно сфокусированные области выделяет алгоритм на основе вычисления отношения коэффициентов вейвлет-преобразования ($r = 0,88$). Однако использовать его в задачах, требующих работы в реальном времени, не рекомендуется, в виду высокой вычислительной сложности. Также следует отметить, данная метрика имеет наименьший доверительный интервал, а следовательно более точна в оценке.
3. Среди алгоритмов, основанных на сверточных операциях, наилучшие результаты по выделению сфокусированных областей показывает алгоритм вычисления метрики Тененграда с использованием маски Робертса ($r = 0,62$). Этот алгоритм примерно равен по вычислительной сложности остальным сверточным алгоритмам и может использоваться в задачах реального времени.
4. Наилучшие показатели с точки зрения быстродействия показал алгоритм на основе вычисления вариации уровня серого (ВУС). Также, среди представленных, данный алгоритм занимает второе место по качеству выделения сфокусированных областей.

Применив один из рассмотренных выше алгоритмов, возможно получить набор оценок сфокусированности $\{FM_i\}$ каждого пикселя на каждом из изображений исходной серии.

2.4. Работа клеточного автомата

Следующий этап работы системы заключается в формировании матрицы меток M , с использованием информации, полученной на этапе оценки сфокусированности областей изображения $\{FM_i\}$. В рамках диссертации, рассматривается подход в основе которого лежит математический аппарат клеточных автоматов. Этот подход находит свое применение в большом количестве практических приложений, в которых система может быть описана как совокупность элементарных объектов, каждый из которых может находиться в конечном числе состояний. Изображение является одним из наиболее наглядных примеров таких систем. Элементарными объектами здесь являются пиксели, а состояния — значения их интенсивности. Искомая матрица меток представляет собой матрицу, размер которой равен размеру изображений исходной серии, а в ячейках расположены номера изображений исходной серии, из которых пиксель с соответствующими координатами, необходимо вынести на итоговое изображение.

2.4.1. Формирование исходных значений для клеточного автомата

Предварительным этапом в работе клеточного автомата является инициализация исходных значений в решетке ячеек. В общем случае, расстановка начальных значений может происходить различными способами: случайно, в соответствии с некоторым распределением, заранее детерминированным образом и каким-либо иным способом.

В работе предлагается подход, основанный на использовании информации об оценке сфокусированности некоторых областей изображения. Данные области предлагается выделить на каждом из изображений исходной серии и использовать их как начальные семена для работы клеточного автомата. Подход основывается на предположении о том, что наиболее четкие пиксели сгруппированы вокруг областей, пиксели которых имеют наибольшие значения метрики сфокусированности. Таким образом, клеточный автомат разметит эти области

одинаковыми маркерами, что соответствует корректному выбору пикселей для извлечения на итоговое изображение. Результатом вышеописанных действий является создание множества матриц меток $\{S_i\}$, в каждой из ячеек (m, n) которых содержится два числа:

- первое число $S_i(m, n)_{img}$ — номер изображения из которого взят данный пиксель, либо 0, в случае, если он был отсеян;
- второе число $S_i(m, n)_{val}$ — значение метрики сфокусированности, либо 0, в случае, если он был отсеян.

На данном этапе возникает задача определения порогового значения, отделяющего пиксели, которые будут использованы как семена, от тех, что не играют существенной роли для построения изображений с расширенной глубиной резкости. Фактически, данная задача сводится к бинаризации изображения. В работе рассмотрено два подхода к выбору данного значения:

- *Эмпирический подход.* Подход основан на эмпирическом подборе порога. Обычно порог представляет собой константу значение которой находится в диапазоне $[0; 1]$. Правило бинаризации в таком случае может выглядеть следующим образом:

$$S_i(m, n)_{img} = \begin{cases} i, & \text{если } FM_i(m, n) > \alpha FM_i^{max} \\ 0, & \text{иначе} \end{cases},$$

$$S_i(m, n)_{val} = \begin{cases} FM_i(m, n), & \text{если } FM_i(m, n) > \alpha FM_i^{max} \\ 0, & \text{иначе} \end{cases},$$

где α — значение порога, FM_i^{max} — максимальное значение метрики сфокусированности на изображении FM_i .

- *Непараметрический подход.* Подход основан на использовании одного из непараметрических методов бинаризации. В обработке изображений часто

применяется непараметрический метод Оцу [69], целью которого является подбор такого порога t , что взвешенная сумма дисперсий

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t),$$

двух классов пикселей, на которые разделяется изображение, была минимальной. В соответствии с выбранным порогом t , в матрицы помещаются следующие значения:

$$S_i(m, n)_{img} = \begin{cases} i, & \text{если } FM_i(m, n) > t \\ 0, & \text{иначе} \end{cases},$$

$$S_i(m, n)_{val} = \begin{cases} FM_i(m, n), & \text{если } FM_i(m, n) > t \\ 0, & \text{иначе} \end{cases}.$$

Завершающим этапом формирования исходных значений является объединение всего множества матриц $\{S_i\}$ в единственную матрицу S , содержащую в каждой ячейке пару значений, по следующему правилу:

$$S(m, n)_{val} = \max_i S_i(m, n)_{val},$$

$$S(m, n)_{img} = i,$$

таким образом матрица S содержит всю информацию для работы клеточного автомата.

2.4.2. Формирование матрицы меток итогового изображения

Формирование матрицы меток, используемой для формирования итогового изображения, происходит при помощи алгоритма, описываемого математическим аппаратом клеточных автоматов. Клеточным автоматом называется дискретная модель, состоящая из следующих ключевых элементов:

- *Регулярная решетка ячеек.* Роль этого элемента в работе принимает на себя двумерная матрица, количество строк и столбцов в которой равно высоте и ширине изображений исходной серии, соответственно.
- *Начальное состояние всех ячеек.* Данный элемент определяется значениями матрицы S , полученной в 2.4.1, и представляют собой упорядоченную пару значений для каждой ячейки.
- *Правило перехода и окрестность.* Данная пара элементов является сильно связанной, так как вид окрестности часто определяется правилом перехода.

В рамках диссертации, работа клеточного автомата является компонентом, вносящим в конечный результат решающий вклад. Алгоритм функционирования может быть представлен в виде блок-схемы, изображенной на рис. 2.11. Здесь M — итоговая матрица меток, $M_{m,n}$ — значение метки, содержащееся в элементе с индексами (m, n) , Θ — множество ячеек, составляющих окрестность элемента $M_{m,n}$. Функция $\phi(M_{m,n})$ определяет правило перехода между двумя последовательными состояниями одной ячейки клеточного автомата и имеет следующий вид:

$$\phi(M_{m,n}) = \begin{cases} 0, & \forall (i, j) \in \Theta : S(i, j)_{val} = 0 \\ S(i, j)_{img}, & \exists (i, j) \in \Theta : \max_{i,j} [S(i, j)_{val} \neq 0] \\ M_{m,n}, & M_{m,n} \neq 0 \end{cases}$$

Классическим вариантом формы окрестности Θ является 8-связная область, приведенная на рис. 1.6 б), также называемая окрестностью Мура.

2.5. Формирование итогового изображения

Сформированная в 2.4.2 матрица меток M в сочетании с серией исходных изображений $\{I_n\}$ могут быть использованы для получения изображения

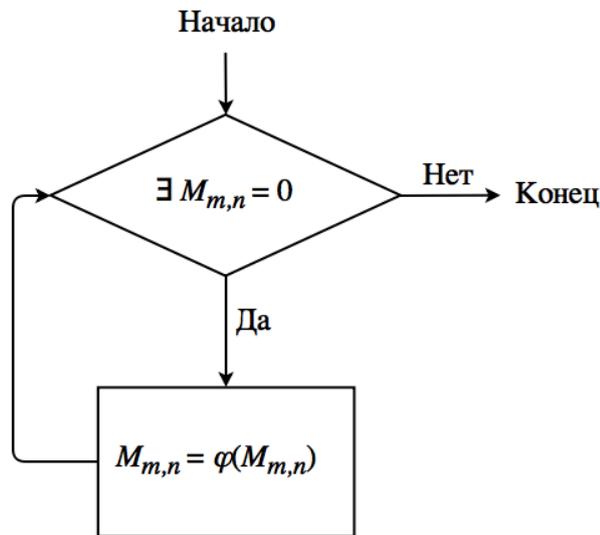


Рис. 2.11. Блок-схема работы клеточного автомата

с расширенной глубиной резкости F . В работе рассмотрено несколько методов, позволяющих сформировать итоговое изображение.

2.5.1. Прямой перенос значений интенсивности пикселей

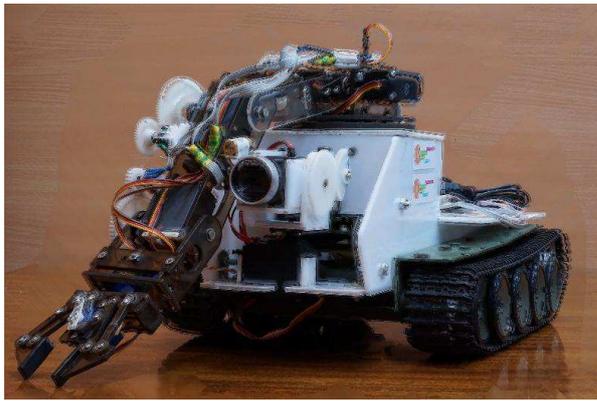
Простейший подход заключается в прямом переносе значения интенсивности пикселя с координатами (x, y) с изображения исходной серии с номером $M(x, y)$ на итоговое изображение F . Таким образом для каждого пикселя изображения F верно выражение 2.3.

$$F(x, y) = I_{M(x,y)}(x, y). \quad (2.3)$$

Перенос выполняется для каждого цветового канала.

Примеры изображений, синтезированных при помощи данного способа представлены на рис. 2.12. Основным преимуществом данного подхода является скорость выполнения, так как все действия сводятся к копированию областей памяти с возможностью параллельного выполнения, что крайне эффективно при реализации на графических процессорах. При использовании данного подхода может быть получено быстрое общее представление о снимаемой сцене.

К недостаткам подхода можно отнести наличие ярко выраженных перехо-



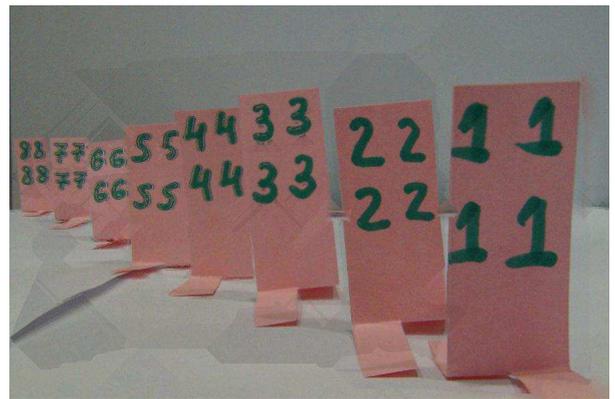
а)



б)



в)



г)

Рис. 2.12. Примеры изображений с использованием прямого переноса пикселей для различных тестовых серий: а) «robot»; б) «toys»; в) «soldiers»; г) «numbers»

дов между областями, которые промаркированы различными метками. С точки зрения человека, такие переходы выглядят не естественно и уменьшают общее качество изображения. Также возможна потеря некоторых деталей. Кроме того, если полученное изображение передать на систему компьютерной обработки, которая использует частотные методы анализа, резкие границы, возникшие при формировании изображения, внесут сильные искажения в результат анализа.

2.5.2. Построение пирамид гауссианов и лапласианов

Частично устранить недостатки прямого переноса призваны методы, основанные на разложении изображений в пирамиды гауссианов и лапласианов. Данный подход позволяет существенно сгладить переходы между областями,

извлеченными из разных изображений, при этом сохранив большую часть деталей.

При описании работы алгоритма приняты следующие обозначения:

M — матрица меток, полученная в результате работы клеточного автомата.

$\{M_n\}, M_n$ — множество и n -ая матрица-маска меток

$\{I_n\}, I_n$ — множество и n -ое исходное изображение исходной серии

$\{LP_n\}, LP_n$ — множество и n -ая пирамида лапласианов исходных изображений

GP_i — пирамида гауссианов текущей интегральной маски

IM_i — текущее значение интегральной маски

Алгоритм формирования итогового изображения с использованием данного метода, в качестве входных параметров использует следующие данные:

1. набор изображений из исходной серии;
2. набор матриц-масок $\{M_n\}$ с метками соответствия исходным изображениям, полученными из матрицы M следующим образом:

$$M_n(i, j) = \begin{cases} 1, & \text{если } M(i, j) = n \\ 0, & \text{иначе} \end{cases},$$

где n — номер матрицы-маски;

3. число уровней разложения изображений;
4. параметр α , характеризующий использующийся для интерполяции фильтр.

В оригинальной статье [70] предложен алгоритм смешивания двух изображений. В работе предлагается модификация, позволяющая объединять произвольное количество изображений. Процесс происходит следующим образом:

1. Дополнить изображения исходной серии $\{I_n\}$ и матрицы-маски $\{M_n\}$ до размера 2^x — ближайшей, в большую сторону, степени двойки. Это необходимо для целочисленного деления на 2 при децимации изображений при построении пирамиды лапласианов.
2. Сделать матрицы-маски трехканальными, т.е. продублировать трижды для обработки цветных изображений. Для изображений, представленных в градациях серого, данный шаг пропускается.
3. Создать множество пирамид лапласианов LP_n для каждого изображения $\{I_n\}$ исходной серии.
4. Инициализировать начальные значения интегральной маски IM_i и итоговой пирамиды LP_i^B значениями M_1 и LP_1 соответственно.
5. Создать итоговую пирамиду изображений, по следующему алгоритму (представлена i -ая итерация алгоритма, работа алгоритма начинается с $i = 2$):

Шаг 1. Создать пирамиду гауссианов GP_{i-1} для текущего значения интегральной маски IM_{i-1}

Шаг 2. Создать пирамиду лапласианов смешанного изображения LP_i^B по следующей формуле:

$$LP_i^B = LP_{i-1}^B * GP_{i-1} + LP_i * (1 - GP_{i-1}).$$

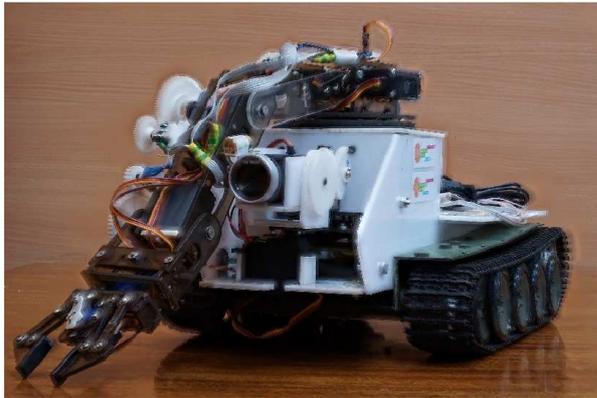
Шаг 3. Обновить текущую интегральную маску:

$$IM_i = IM_{i-1} + M_i.$$

В ходе работы i принимает значения от 2 до N — количества различных меток равное количеству матриц-масок $\{M_n\}$.

6. Создать итоговое изображение F из итоговой пирамиды лапласианов LP_N^B .

Примеры изображений, синтезированных при помощи данного способа, представлены на рис. 2.13. Преимуществом данного подхода является качественное, с точки зрения визуального восприятия, сглаживание границ между участками, извлеченными из различных изображений. Однако подход достаточно затратен по вычислительной стоимости, что делает его сложно применимым в системах реального времени.



а)



б)



в)



г)

Рис. 2.13. Примеры изображений, синтезированных с использованием пирамид гауссианов и лапласианов для различных тестовых серий: а) «robot»; б) «toys»; в) «soldiers»; г) «numbers»

2.6. Краткие выводы

В главе описан разработанный алгоритм формирования изображений с расширенной глубиной резкости на основе клеточного автомата, а также его

модификация на основе пирамид изображений. Проведено сравнение метрик сфокусированности применительно к рассматриваемой задаче. Проведен эксперимент по сравнению различных метрик сфокусированности, в рамках которого предложены идеализированные характеристики метрик сфокусированности для рассматриваемой задачи. Установлено, что наиболее точной метрикой оценки сфокусированности является метрика на основе соотношения коэффициентов вейвлет-преобразования, для которой коэффициент корреляции с идеализированными метриками $r = 0,88$. Также данная метрика имеет наименьший динамический диапазон при уровне значимости $0,95$. Установлено, что наиболее эффективной с точки зрения скорости, а также весьма эффективной с точки зрения точности $r = 0,76$ является метрика сфокусированности на основе расчета вариации уровня серого. Среди алгоритмов, основанных на сверточных операциях, наилучшие результаты по выделению сфокусированных областей показывает алгоритм вычисления метрики Тененграда с использованием маски Робертса $r = 0,62$. Этот алгоритм примерно сопоставим по вычислительной сложности с остальными сверточными алгоритмами и может использоваться в задачах реального времени. Установлено, что диапазон $[0,03; 0,12]$ значения пороговой бинаризации в алгоритме смешивания на основе клеточного автомата соответствует наиболее качественному смешиванию изображений исходной серии. Разработана и протестирована модификация алгоритма смешивания изображений с расширенной глубиной резкости на основе пирамид гауссианов и лапласианов.

Глава 3

Исследование алгоритма формирования изображений с расширенной глубиной резкости на основе клеточных автоматов и пирамид изображений

3.1. Исследование работы алгоритма в условиях наличия шумов

В научно-технической литературе задача формирования изображений с расширенной глубиной резкости рассматривается в отрыве от проблемы возникновения шумов, появляющихся в процессе формирования цифрового изображения и оказывающих негативное влияние на процедуру построения изображений. В диссертации рассматривается подход, позволяющий выполнить построение изображений с расширенной глубиной резкости при наличии шумового воздействия. Суть подхода заключается в использовании алгоритмов предфильтрации (фильтрация до объединения зашумленных изображений) или постфильтрации (фильтрация после объединения зашумленных изображений) с целью формирования качественного изображения с расширенной глубиной резкости.

На сегодняшний день алгоритмы фильтрации цифровых изображений [71], широко применяются в области современной науки и техники и имеют множество практических приложений. Достаточно интересным подходом к задаче фильтрации изображений является использование методов машинного обучения [72, 73].

3.1.1. Алгоритм фильтрации цифровых изображений

В работе при решении задачи шумоподавления предлагается использовать двухэтапная схема фильтрации изображений на основе анализа главных компонент (АГК) [74, 75], краткое описание которой представлено ниже, что анализируемое цифровое изображение x искажено аддитивным белым гауссовским шумом (АБГШ) n с нулевым математическим ожиданием и среднеквадратическим отклонением σ .

Первый этап обработки

1. Считаем среднеквадратическое отклонение шума σ на входном зашумленном изображении $y = x + n$ известным.
2. Разбиваем входное зашумленное изображение на совокупность перекрывающихся блоков. Внутри каждого из них можно выделить: *область обучения*, *область фильтрации* и *область наложения блоков*. Размеры рассматриваемых областей могут варьироваться.
3. Внутри области обучения выбираем всевозможные квадратные блоки (*векторы обучения*). Последние, будучи представленными в виде векторов-столбцов, позволяют сформировать некоторую выборочную матрицу, столбцами которой являются рассматриваемые векторы-столбцы.
4. На основе выборочной матрицы, предварительно центрировав ее, составляем ковариационную матрицу, для которой находим собственные числа и соответствующие им собственные векторы (*главные компоненты* данных, заключенных в центрированной выборочной матрице). Итогом данного этапа является построение матрицы ортогонального преобразования.
5. Находим проекции (коэффициенты преобразования) множества векторов, заключенных в центрированной выборочной матрице, на множество собственных векторов, найденных шаге 4.

6. Осуществляем обработку полученного множества проекций с использованием линейной среднеквадратической оценки [73].
7. На основе множества обработанных данных восстанавливаем оценку выборочной матрицы незашумленных данных, а на основе последней – отдельную обработанную область на изображении. При этом, во-первых, восстанавливается область обучения путем вложения в нее, с учетом наложения, векторов обучения в соответствующие им пространственные позиции. Векторы обучения, находясь внутри оценки выборочной матрицы незашумленных данных в виде векторов-столбцов, снова были преобразованы в квадратные блоки перед вложением в область обучения. Необходимо отметить, что область наложения векторов обучения усредняется с использованием арифметического усреднения. Во-вторых, после восстановления области обучения из нее выделяется область фильтрации меньших размеров. Повторяя аналогичную операцию для остальных областей фильтрации, с учетом их наложения, можно обработать зашумленное изображение целиком и получить первичную «грубую» оценку \hat{x}^I неискаженного изображения x . При этом обработанные области фильтрации вкладываются в соответствующие пространственные позиции изображения \hat{x}^I , а их область наложения усредняется арифметически.

Второй этап обработки

1. Используя зашумленное изображение y , повторяем шаги 2 – 5, рассмотренные на первом этапе обработки. При этом устанавливаются другие размеры областей обучения, областей фильтрации и областей наложения, а также векторов обучения.
2. Осуществляем обработку полученного множества проекций с использованием эмпирического фильтра Винера, представленного в области главных компонент и вычисленного на основе первичной оценки \hat{x}^I .

3. Повторяя рассуждения, изложенные на шаге 7 первого этапа обработки, получаем вторую, «точную» оценку \hat{x}^{II} неискаженного изображения x .

На рис. 3.1 представлена блок-схема описанного алгоритма фильтрации цифровых изображений. Необходимо отметить, что при обработке цветных RGB-изображений данная схема фильтрации применяется поканально.

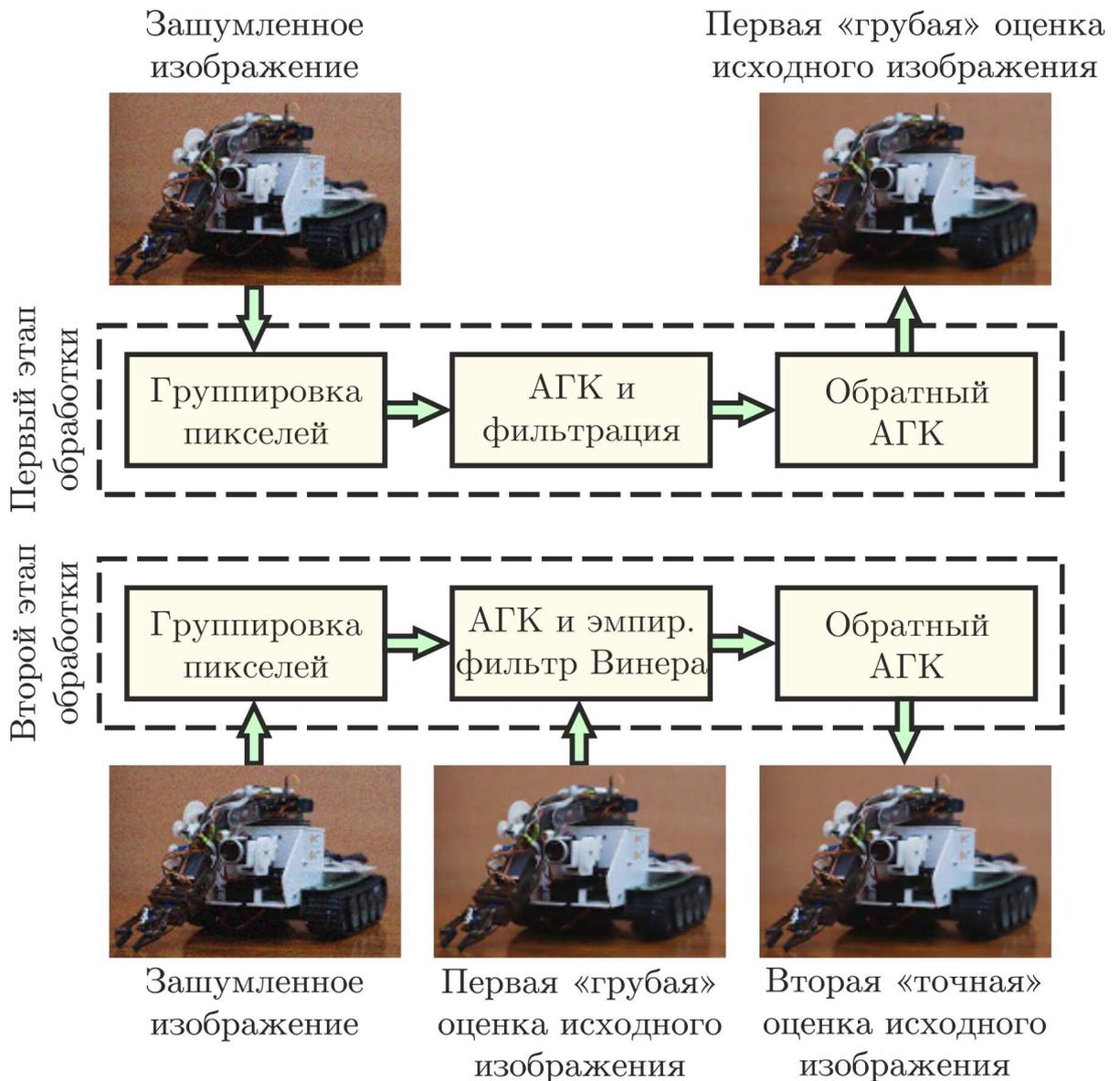
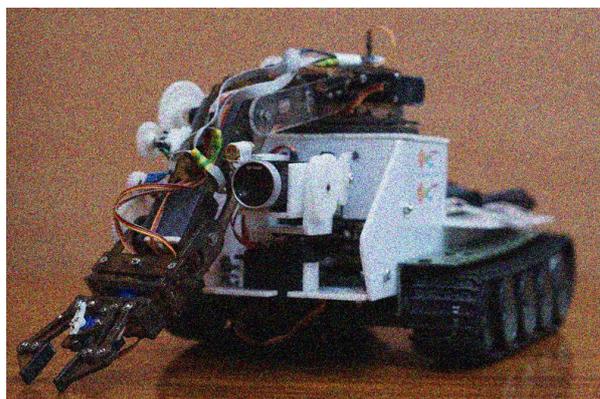


Рис. 3.1. Блок-схема обработки цифрового изображения с использованием двухэтапной схемы фильтрации изображений на основе анализа главных компонент

На рис. 3.2 представлен пример фильтрации на примере изображений из тестовой серии «robot» (рис. A.1 г)). Каждое из изображений сцены «robot»,

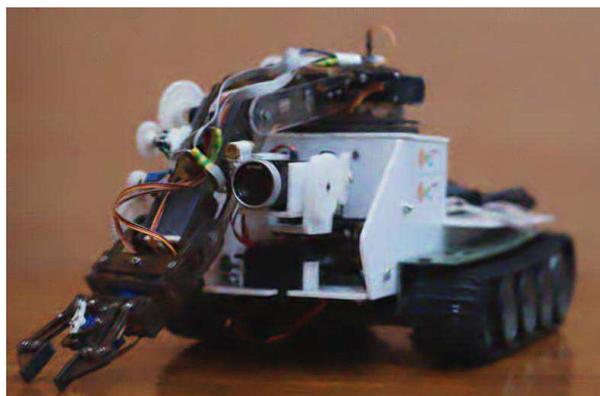
сформированное с разным фокусным расстоянием, зашумлено АБГШ со среднеквадратическим отклонением $\sigma = 35$. Дополнительно на рис. 3.2 указаны численные оценки ПОСШ, дБ (пиковое отношение сигнала к шуму) [76] и КСП (коэффициент структурного подобия) [77] для зашумленных и восстановленных изображений.



а) ПОСШ = 17,86 / КСП = 0,252



б) ПОСШ = 17,77 / КСП = 0,197



в) ПОСШ = 29,34 / КСП = 0,845



г) ПОСШ = 31,92 / КСП = 0,871

Рис. 3.2. Примеры изображений для сцены «robot»: а) и б) зашумленные изображения ($\sigma = 35$); в) и г) соответствующие восстановленные изображения

3.1.2. Сравнение подходов к фильтрации изображений с расширенной глубиной резкости

Для проведения исследования были рассмотрены следующие подходы к объединению изображений:

- *Подход 1 — Без фильтрации.* Исходные изображения сцен зашумлялись АБГШ с фиксированным σ и затем объединялись, используя алгоритм формирования изображений с расширенной глубиной резкости.
- *Подход 2 — Предфильтрация.* Исходные изображения сцен зашумлялись АБГШ с фиксированным σ , подвергались фильтрации, используя двухэтапную схему обработки изображений на основе анализа главных компонент и затем объединялись, используя алгоритм формирования изображений с расширенной глубиной резкости.
- *Подход 3 — Постфильтрация.* Исходные изображения сцен зашумлялись АБГШ с фиксированным σ , объединялись, используя алгоритм формирования изображений с расширенной глубиной резкости и затем итоговое объединенное изображение подвергалось фильтрации, используя двухэтапную схему обработки изображений на основе анализа главных компонент. В данном случае предполагалось, что модель шума и его параметры на объединенном изображении являются идентичными тем, что были у исходных изображений до слияния.

При выполнении процедуры моделирования использовалось четыре группы цветных RGB-изображений фиксированных сцен: «numbers», «robot», «soldiers» и «toys». Полученные изображения с расширенной глубиной резкости были взяты за эталонные при проведении исследования качества работы алгоритма формирования изображений с расширенной глубиной резкости в присутствии шумов.

В табл. 3.1 представлено численное сравнение различных подходов к объединению нескольких изображениях фиксированной сцены для различных степеней зашумления σ . Наилучшие результаты выделены жирным шрифтом. Дополнительно на рис. 3.3 представлены визуальные результаты построения изображений с расширенной глубиной резкости. Для проведения исследования ис-

пользовались стандартные метрики эталонной оценки качества – пиковое отношение сигнал/шум (ПОСШ) и коэффициент структурного подобия (КСП).

Метрика ПОСШ определяется как

$$\text{ПОСШ} = 10 \log_{10} \left(\frac{\text{max}^2}{\text{СКО}} \right) = 20 \log_{10} \left(\frac{\text{max}}{\text{СКО}} \right), \quad (3.1)$$

здесь под пиковым значением (max) подразумевается максимальное возможное значение, которое может принять пиксель на изображении. Для 8-битового изображения оно равно 255.

Метрика КСП характеризует похожесть двух изображений A и B и определена следующим образом:

$$\text{КСП}(A, B) = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\sigma_A^2 + \sigma_B^2 + C_2)},$$

где μ_A и μ_B – среднее значение интенсивности на изображениях A и B соответственно, σ_A и σ_B – среднее квадратичное отклонение интенсивностей на изображениях A и B соответственно, σ_{AB} – ковариация A и B , C_1 и C_2 – небольшие константы для A и B соответственно. Как и в предыдущем случае, большее значение метрики соответствует большей похожести изображений.

Данные метрики применялись к двум изображениям. Первое изображение, полученное с использованием алгоритма формирования изображений с расширенной глубиной резкости в отсутствие шумов, бралось за эталонное. Второе изображение было получено с использованием одного из трех вышеуказанных подходов.

Анализ результатов показывает, что качественное объединение зашумленных изображений фиксированной сцены с использованием алгоритма формирования изображений с расширенной глубиной резкости при наличии шумов возможно (рис. 3.3 в)), если исходные зашумленные данные будут предварительно обработаны с использованием определенной схемы фильтрации, например, приведенной в 3.1.1. Объединение изображений при отсутствии этой схемы не позволяет качественно идентифицировать пиксели, которые должны внести существенный вклад в построение изображения с расширенной глубиной резкости.

Таблица 3.1. ПОСШ, дБ / КСП изображений с расширенной глубиной резкости при наличии АБГШ

σ	Подход 1	Подход 2	Подход 3
Сцена «numbers»			
5	31,28 / 0,749	32,51 / 0,876	34,83 / 0,879
15	24,21 / 0,416	31,99 / 0,862	31,83 / 0,845
20	22,03 / 0,308	31,82 / 0,857	31,25 / 0,838
25	20,26 / 0,230	31,80 / 0,852	30,84 / 0,833
35	17,52 / 0,140	31,46 / 0,843	30,17 / 0,825
Сцена «robot»			
5	32,21 / 0,881	35,60 / 0,938	34,63 / 0,933
15	24,28 / 0,510	31,04 / 0,847	29,78 / 0,805
20	22,23 / 0,415	29,84 / 0,816	28,73 / 0,776
25	20,55 / 0,340	28,94 / 0,791	27,86 / 0,753
35	17,99 / 0,243	27,39 / 0,748	26,49 / 0,717
Сцена «soldiers»			
5	31,87 / 0,907	33,85 / 0,954	33,63 / 0,953
15	23,32 / 0,584	28,39 / 0,869	26,42 / 0,809
20	21,34 / 0,488	27,06 / 0,834	24,94 / 0,752
25	19,79 / 0,412	26,04 / 0,801	23,92 / 0,709
35	17,40 / 0,298	24,53 / 0,742	22,55 / 0,646
Сцена «toys»			
5	32,67 / 0,887	33,92 / 0,942	35,96 / 0,951
15	22,66 / 0,450	29,80 / 0,885	26,19 / 0,835
20	20,72 / 0,342	29,06 / 0,870	24,95 / 0,809
25	19,24 / 0,266	28,49 / 0,858	24,32 / 0,792
35	16,84 / 0,167	27,58 / 0,835	23,38 / 0,766

а) ПОСШ = ∞ / КСП = 1

б) ПОСШ = 19,24 / КСП = 0,266



в) ПОСШ = 28,49 / КСП = 0,858



г) ПОСШ = 24,32 / КСП = 0,792

Рис. 3.3. Примеры построения изображений с расширенной глубиной резкости для тестовой сцены «toys», с указанием ПОСШ, дБ / КСП: а) слияние без добавления АБГШ в объединяемые изображения; б) слияние с добавлением АБГШ ($\sigma = 25$) в объединяемые изображения; в) слияние с добавлением АБГШ ($\sigma = 25$) в объединяемые изображения и их фильтрацией до объединения; г) слияние с добавлением АБГШ ($\sigma = 25$) в объединяемые изображения и фильтрацией объединенного изображения

Как следствие, объединенное изображение в данном случае будет содержать шум, а также значительное размытие объектов интереса, присутствующих в сцене (рис. 3.3 б)). Поэтому фильтрация после объединения не дает качественного результата обработки (рис. 3.3 г)). Однако необходимо отметить то, что фильтрация до объединения изображений приводит к многократному применению схемы шумоподавления. В этом случае каждое зашумленное изображение фиксированной сцены подвергается фильтрации, что значительно увеличивает вычислительную стоимость подхода 2. В случае подхода 3 фильтрация выпол-

няется однократно применительно к объединенному изображению.

3.2. Способы оценки качества изображений

Качество изображения описывает то, насколько ухудшилось изображение по сравнению с некоторым эталоном. Системы обработки, подобные рассматриваемой, могут вносить некоторые искажения и артефакты в получаемые изображения, следовательно — оценка их качества является важной задачей. Все множество методов и оценок качества изображений можно разделить на две большие группы: методы, основанные на эталонном сравнении и методы без использования эталона. В первом случае, изображение сравнивается с эталоном, качество которого считается идеальным. Во втором — подобного сравнения не происходит [4].

3.2.1. Эталонные метрики оценки качества изображений

Пусть изображение A является эталонным, а изображение B получено при помощи системы смешивания изображений с расширенной глубиной резкости. Также пусть (i, j) — номера строки и столбца соответственно на изображении, являющиеся координатами некоторого пикселя. Изображение имеет размер $M \times N$. Таким образом, имеют место выражения:

$$1 \leq i \leq M,$$

$$1 \leq j \leq N.$$

Используя приведенные выше обозначения и допущения, можно ввести ряд эталонных метрик оценки качества изображений.

Средняя квадратичная ошибка (СКО)

Метрика определяет средний квадрат разности изображений относительно друг друга

$$\text{СКО} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [A(i, j) - B(i, j)]^2.$$

Корень из СКО (КСКО)

Часто, вместо СКО используется корень из этого значения, представляющий собой среднюю разницу между соответствующими пикселями изображения

$$\text{СКО} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [A(i, j) - B(i, j)]^2}.$$

Взаимная информация (ВИ)

Взаимная информация используется для оценки схожести распределения интенсивности пикселей на двух изображениях. Гистограмма изображения может быть использована для получения распределения вероятностей. Чем больше значение данной метрики, тем лучше отработал алгоритм смешивания изображений. Взаимная информация определяется следующей формулой:

$$\text{ВИ}(A, B) = \sum_{a,b} \left(P_{AB}(a, b) \frac{\log_2(P_{AB}(a, b))}{P_A(a)P_B(b)} \right),$$

где $P_{AB}(a, b)$ — взаимное распределение вероятностей, $P_A(a)$ и $P_B(b)$ — распределения вероятностей на изображениях A и B соответственно.

Также в качестве эталонных рассматриваются введенные в разделе 3.1.2 метрики ПОСШ и КСП.

3.2.2. Безэталонные метрики оценки качества изображений

Когда эталонное изображение не доступно — необходимо использовать методы, не опирающиеся на знание истинных значений соответствующих пикселей. Для этого могут быть использованы безэталонные метрики [5, 7].

Энтропия

Энтропия — мера информативности сообщения. Увеличение значения энтропии изображения, полученного после смешивания, по сравнению со значением энтропии исходных изображений, свидетельствует о том, что полученное изображение несет большее количество информации. Энтропия определяется следующим образом:

$$E = - \sum_{i=0}^G p(i) \log_2 p(i),$$

где G — количество уровней серого в гистограмме изображения (255 для 8-битного изображения), $p(i)$ — нормализованная частота встречи i -го уровня серого. Информативность изображения выражается в битах на пиксель. Следует отметить, что данная метрика чувствительна к шумам и другим резким флуктуациям интенсивности пикселей.

Дисперсия

Эта метрика наиболее эффективна в отсутствии шумов. Она позволяет оценить контраст смешанного изображения. Изображения с большим контрастом, имеют более высокое значение метрики.

$$D = \frac{1}{MN} \sum_i^M \sum_j^N [I(i, j) - \mu]^2,$$

здесь μ — среднее значение интенсивности изображения.

Пространственные частоты

Пространственные частоты характеризуют изменения интенсивности происходящие на смешанном изображении. Метрика вводится следующим образом:

$$SF = \sqrt{RF^2 + CF^2},$$

где RF и CF — вертикальные и горизонтальные частоты соответственно, определяемые следующим образом:

$$RF = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=2}^N [I(i, j) - I(i, j - 1)]^2},$$

$$CF = \sqrt{\frac{1}{MN} \sum_{i=2}^M \sum_{j=1}^N [I(i, j) - I(i - 1, j)]^2}.$$

Взаимная смешанная информация (ВСИ)

Данная метрика идейно близка к смешанной информации в эталонных метриках, однако в связи с отсутствием эталонного изображения, внесены некоторые изменения. Пусть A и B — исходные изображения, а F — изображение после смешивания. Тогда определим взаимную смешанную информацию как

$$\text{ВСИ}(A, B) = \text{ВИ}(F, A) + \text{ВИ}(F, B).$$

Метрика определяет схожесть смешанного и исходных изображений. Здесь также высокое значение метрики соответствует более высокому качеству работы алгоритма. Существуют более точные способы вычисления похожести двух изображений, основанные на идее вычисления взаимной энтропии двух изображений, однако, они, как правило, вычислительно более затратны.

Дивергенция и расстояние между совместным распределением и произведением маргинальных распределений пары изображений могут быть использованы в качестве мер их подобия. Классом мер дивергенции, использующих взаимную информацию, является класс F информации или F дивергенции. Мерами F информации являются следующие величины [78]:

$$I_\alpha = \frac{1}{\alpha(\alpha - 1)} \left(\sum_{i=0}^{255} \sum_{j=0}^{255} \frac{p_{i,j}^\alpha}{(p_i p_j)^{\alpha-1}} - 1 \right),$$

$$M_\alpha = \sum_{i=0}^{255} \sum_{j=0}^{255} |p_{i,j}^\alpha - (p_i p_j)^\alpha|^{\frac{1}{\alpha}},$$

$$\chi_\alpha = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{|p_{i,j} - p_i p_j|^\alpha}{(p_i p_j)^{\alpha-1}}.$$

Мера I_α определена при $\alpha \neq 0$, $\alpha \neq 1$ и сводится к взаимной информации Шеннона при $\alpha = 1$. Мера M_α определена при $0 \leq \alpha \leq 1$, а χ_α — при $\alpha > 1$.

Для вычисления некоторых из приведенных выше метрик, необходимо вычислить $p_{i,j}$ — элементы совместной плотности распределения вероятностей (СПРВ) яркостей изображений, которая может быть оценена, например, с использованием гистограммы.

Для двух 8-битовых одноканальных изображений X и Y , каждое из которых имеет размер $M \times N$, значения $p_{i,j}$ можно представить как значения карманов двумерной гистограммы размера 256×256 .

$$p_{i,j} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, & \text{если } I(X_{m,n}) = i \text{ и } I(Y_{m,n}) = j \\ 0, & \text{иначе} \end{cases},$$

где $I(X_{m,n})$ — интенсивность пикселя с координатами (m, n) на изображении X , а $I(Y_{m,n})$ — на изображении Y соответственно. Примеры попарной совместной плотности распределения вероятностей для трех изображений приведены на рис. 3.4. Для наглядности гистограммы приведены в логарифмическом масштабе по значению.

$Q^{AB/F}$ -метрика

Метрика предложена Ксидеасом (Xydeas) и Петровичем (Petrovic) в работах [6, 79]. Основной идеей метрики является предположение о том, что большая часть границ должна переноситься с исходных изображений на сформированное. Этот метод использует оператор Собеля для вычисления информации о модуле $g(n, m)$ и направлении $\alpha(n, m)$ градиента в каждом пикселе. Для двух исходных изображений A и B и сформированного изображения F , оператор Собеля применяется следующим образом:

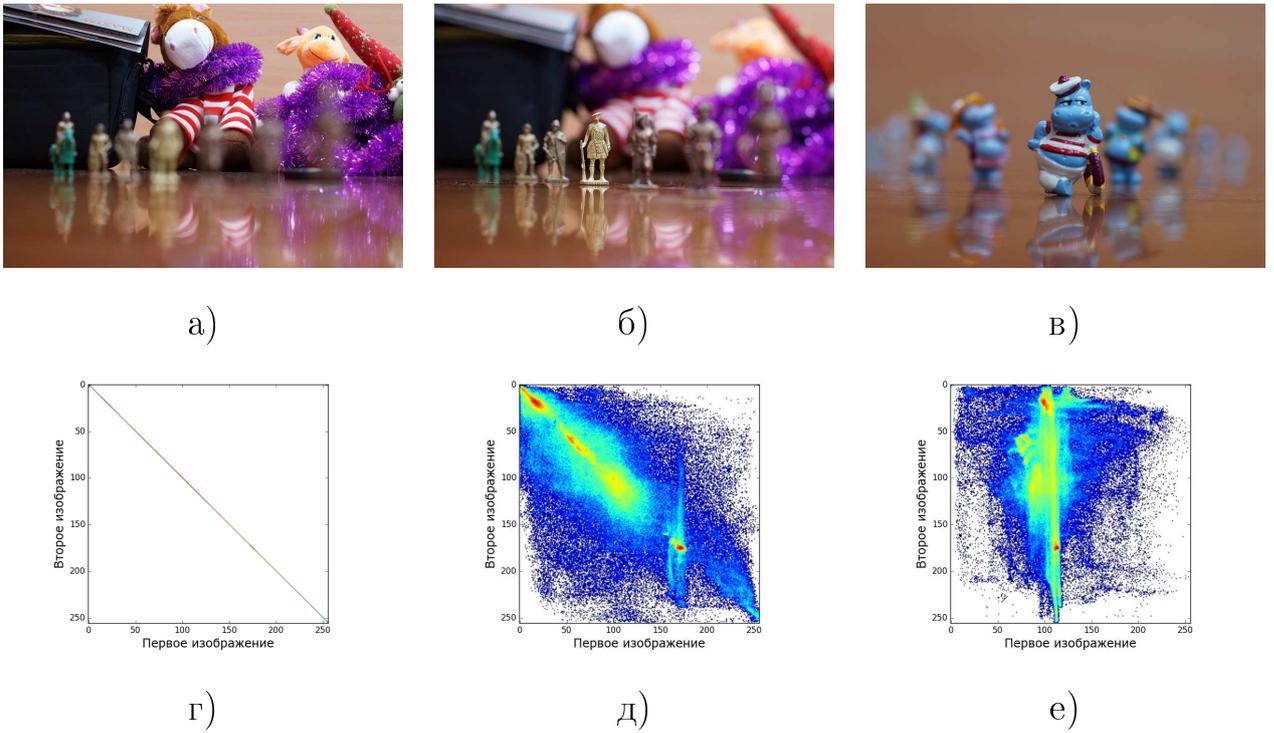


Рис. 3.4. Совместная плотность распределения вероятностей (СПРВ) для различных пар изображений: а-в) изображения, выбранные для иллюстрации СПРВ; г) СПРВ одинаковых изображений а); д) СПРВ для изображений а) и б); е) СПРВ для изображений а) и в)

$$g_A(n, m) = \sqrt{s_A^x(n, m)^2 + s_A^y(n, m)^2},$$

$$\alpha_A(n, m) = \tan^{-1} \left(\frac{s_A^y(n, m)^2}{s_A^x(n, m)^2} \right),$$

где $s_A^y(n, m)$ и $s_A^x(n, m)$ — значения вертикального и горизонтального шаблонов Собеля, свернутые с областью с центром в точке с координатами (n, m) на изображении A . Относительные значения силы и ориентации градиента между изображениями A и F формируются следующим образом:

$$(G_{n,m}^{AF}, A_{n,m}^{AF}) = \left[\left(\frac{g_{n,m}^F}{g_{n,m}^A} \right)^M, 1 - \frac{|\alpha_A(n, m) - \alpha_F(n, m)|}{\pi/2} \right],$$

где

$$M = \begin{cases} 1, & \text{если } g_A(n, m) > g_F(n, m) \\ -1, & \text{иначе} \end{cases}.$$

Информация о границах вычисляется следующим образом:

$$Q_{n,m}^{AF} = \frac{\Gamma_\alpha}{\left(1 + e^{k_\alpha(A_{m,n}^{AF} - \sigma_\alpha)}\right)} \frac{\Gamma_g}{\left(1 + e^{k_g(G_{m,n}^{AF} - \sigma_g)}\right)},$$

здесь $\Gamma_\alpha, \Gamma_g, k_\alpha, k_g$ и σ_α, σ_g — константы определяющие кривизну и амплитуду сигмоидальной функции. Авторы в [80] предлагают использовать следующие значения этих величин: $\Gamma_\alpha * \Gamma_g = 1$, $k_\alpha = 0,8$, $k_g = 0,7$ и $\sigma_\alpha = 24$, $\sigma_g = 11$. Таким образом значение метрики $Q^{AB/F}$ вычисляется как

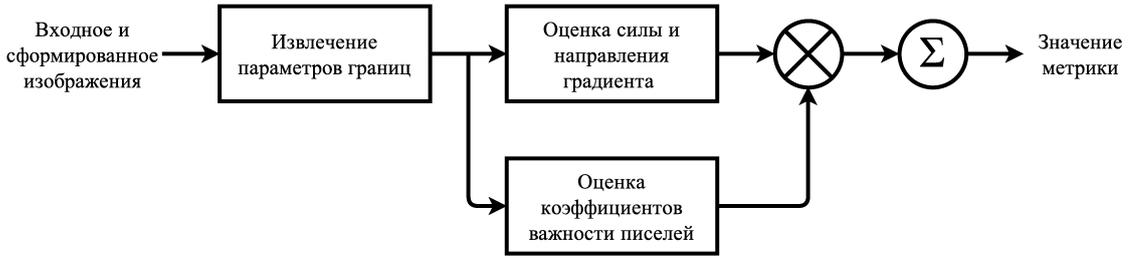
$$Q^{AB/F} = \frac{\sum_{n,m} (Q_{n,m}^{AF} w_{n,m}^A + Q_{n,m}^{BF} w_{n,m}^B)}{\sum_{n,m} (w_{n,m}^A + w_{n,m}^B)},$$

и представляет собой взвешенную сумму значений $Q_{n,m}^{AF}$ и $Q_{n,m}^{BF}$ и множителей w^A и w^B , которые определяют важность пикселя (m, n) . Веса определяются как

$$w_{n,m}^A = d_A(m, n) C_A(m, n) P_A(m, n) g_A(n, m)^L,$$

здесь $d(m, n)$ — флаг детектирования, равен 1, если пиксель с координатами (m, n) принадлежит границе, 0 — иначе. $C(m, n)$ — коэффициент, определяющий корреляцию по направлению с соседними границами, стремится к 1, если соседние границы сонаправлены с данной границей, и стремится к 0 в обратном случае. $P_A(m, n)$ — коэффициент, зависящий от положения рассматриваемого пикселя, в общем случае определяется исходя из метода смешивания, например, линейно уменьшается от центра к краям изображения. Последний множитель зависит от модуля градиента в данной точке, так как чем больше модуль, тем сильнее сфокусированность, L здесь — некоторая константа [81], которую авторы предлагают принять равной $L = 1,5$.

Для идеального смешивания $Q^{AB/F} = 1$. Для данной метрики верно правило: чем больше значение — тем лучше результат. Общая схема вычисления метрики приведена на рис. 3.5.

Рис. 3.5. Схема вычисления $Q^{AB/F}$ -метрики

Модифицированный универсальный индекс качества изображения (МУИК)

Данная метрика предложена Вангом (Wang) и Бовиком (Bowik) в [82]. Пусть есть два изображения A и B размера $M \times N$ каждое. Пусть \bar{A} и \bar{B} — средние значения интенсивности изображений A и B соответственно, а σ_A^2 и σ_B^2 — дисперсии. Также обозначим ковариацию изображений как σ_{AB} . Таким образом можно записать:

$$\sigma_A^2 = \frac{1}{1 - MN} \sum_{m=1}^M \sum_{n=1}^N (A(m, n) - \bar{A})^2,$$

$$\sigma_B^2 = \frac{1}{1 - MN} \sum_{m=1}^M \sum_{n=1}^N (B(m, n) - \bar{B})^2,$$

$$\sigma_{AB} = \frac{1}{1 - MN} \sum_{m=1}^M \sum_{n=1}^N (A(m, n) - \bar{A})(B(m, n) - \bar{B}).$$

Определим

$$Q_0(A, B) = \frac{4\sigma_{AB}\bar{A}\bar{B}}{(\bar{A}^2 + \bar{B}^2)(\sigma_A^2 + \sigma_B^2)}, \quad (3.2)$$

которое может быть представлено в виде:

$$Q_0(A, B) = \frac{\sigma_{AB}}{\sigma_A\sigma_B} \cdot \frac{2\bar{A}\bar{B}}{\bar{A}^2 + \bar{B}^2} \cdot \frac{2\sigma_A\sigma_B}{\sigma_A^2 + \sigma_B^2}. \quad (3.3)$$

В оригинальной статье авторы описывают Q_0 как индекс качества изображений и используют его для количественной оценки структурных искажений между изображениями A и B . Можно рассматривать Q_0 как меру схожести

изображений A и B , которая принимает значения в диапазоне $[-1; 1]$. Следует отметить, что компоненты выражения 3.3 обладают определенным смыслом. Первый компонент представляет собой коэффициент корреляции между изображениями. Второй — характеризует среднее искажение яркости и находится в диапазоне $[0; 1]$. Третий компонент определяет искажение контраста между изображениями и также принимает значения от 0 до 1. Таким образом, при $Q_0 = 1$ изображения A и B являются идентичными.

Так как изображения является нестационарным сигналом необходимо ограничивать области в которых производится оценка значения Q_0 , а затем объединить полученные данные для оценки изображений в целом. В оригинальной статье авторы предлагают использовать для этого метод скользящего окна. Скользящее окно w размера $m \times n$ проходит по изображению начиная в левого верхнего угла вправо и вниз с шагом в один пиксель по каждому направлению. Таким образом оценивается локальное значение Q_0 внутри текущего положения окна. Итоговое значение Q_0 определяется как усредненное значение всех полученных локальных значений, согласно выражению 3.4, где W — множество всех окон, а $|W|$ — общее количество окон.

$$Q_0(A, B) = \frac{1}{|W|} \sum_{w \in W} Q_0(A, B|w). \quad (3.4)$$

Индекс качества смешивания (ИКС)

Основываясь на результатах [82], в частности на выражении 3.4, в [83], предложено расширение индекса качества для случая смешивания изображений. Пусть A и B — исходные изображения, а F — изображение, полученное в результате работы алгоритма смешивания. Таким образом, значение метрики $Q(A, B, F)$ должно описывать качество работы алгоритма при смешивании изображений A и B .

Обозначим $s(A|w)$ некоторую количественную характеристику, определяющую изображение A внутри окна w . Значение $s(A|w)$, может, например, за-

висеть от контраста, дисперсии, энтропии или других характеристик изображения или их комбинаций. Определив $s(A|w)$ и $s(B|w)$ для конкретного окна w , введем величину $\lambda_A(w)$, изменяющуюся в диапазоне $[0; 1]$ и отражающую относительную важность изображения A по сравнению с изображением B для описания содержания w . В статье для этого предлагается выражение 3.5.

$$\lambda_A w = \frac{s(A|w)}{s(A|w) + s(B|w)}. \quad (3.5)$$

Аналогичное выражение можно записать для $\lambda_B(w)$. Таким образом можно записать выражение определяющее индекс качества смешивания $Q(A, B, F)$ как

$$Q(A, B, F) = \frac{1}{|W|} \sum_{w \in W} (\lambda_A(w)Q_0(A, F|w) + \lambda_B(w)Q_0(B, F|w)). \quad (3.6)$$

Таким образом, в тех областях изображения, где наиболее важно сохранение информации от изображения A , значение метрики по большей части определяется значением $Q_0(A, F|w)$ и соответственно в областях, где B более полно описывает содержание сцены значение $Q_0(B, F|w)$ включено с большим весом.

Также автор предлагает ряд модификаций предложенной метрики качества смешивания изображений. Во-первых, предлагается учитывать более сильную значимость некоторых областей изображений по сравнению с другим. Для этого предлагается использовать весовые коэффициенты $C(w)$, характеризующие значимость окна w , таким образом выражение 3.6 преобразуется к следующему виду:

$$Q_W(A, B, F) = \frac{1}{|W|} \sum_{w \in W} c(w) (\lambda_A(w)Q_0(A, F|w) + \lambda_B(w)Q_0(B, F|w)), \quad (3.7)$$

где $c(w) = \frac{C(w)}{\sum_{w' \in W} C(w')}$ — относительный вес окна w .

Следующим уточнением метрики является дополнение, связанное с особенностью человеческого зрения. Особенность состоит в том, что при анализе изображения, человеческий мозг извлекает большую часть информации из границ

объектов, присутствующих в сцене. Если в выражении 3.4 заменить исходные выражения на изображения границ A', B', F' , которые могут быть вычислены, например, с применением оператора Собеля, то комбинируя полученные выражения, можно записать:

$$Q_E(A, B, F) = Q_W(A, B, F)^{1-\alpha} \cdot Q_W(A', B', F')^\alpha,$$

где α — параметр, лежащий в диапазоне $[0; 1]$, и характеризующий важность изображений с границами в итоговом значении метрики.

Метрики $Q(A, B, F)$, $Q_W(A, B, F)$ и $Q_E(A, B, F)$ имеют диапазон значений $[-1; 1]$, чем ближе значение любой из метрик к 1, тем более качественным является результат смешивания.

3.3. Сравнение алгоритмов формирования изображений с расширенной глубиной резкости

3.3.1. Формирование тестовой базы

Количественная оценка качества работы алгоритмов формирования изображений с расширенной глубиной резкости сфокусированных изображений представляет собой важную, с практической точки зрения, задачу. При ее решении необходимо учитывать ряд свойств, которым обладают рассматриваемые системы, а также изображения, формируемые ими:

- отсутствие эталонного изображения;
- широкий спектр возможных сцен и параметров съемки;
- потенциальное наличие искажений схожего вида, но иной природы происхождения.

Выбор тестовых сцен

Перечисленные выше особенности показывают, что формируемая тестовая база требует наличия разнообразных изображений для обеспечения репрезентативности. Тестовые серии, описанные в пункте 2.1, а также проиллюстрированные в приложении А, покрывают достаточно большой спектр возможных композиций объектов в сценах:

- один или несколько объектов;
- однородный или сложно текстурированный фон;
- сложные или простые объекты интереса.

Таким образом, серии изображений «numbers», «robot», «toys» и «soldiers» отобраны как тестовые для оценки качества работы алгоритмов. К данным применены различные алгоритмы и наборы параметров, описанные далее.

Выбор алгоритмов и диапазонов параметров

Определившись с наборами исходных изображений, необходимо составить список алгоритмов и наборы их параметров, при помощи которых будет осуществляться формирование изображений с расширенной глубиной резкости. В работе предлагается следующий набор алгоритмов для сравнения:

- BF — алгоритм на основе блочного смешивания (п. 1.4.1);
- WT — алгоритм на основе вейвлет-преобразования (п. 1.5.1);
- CF — алгоритм на основе клеточного автомата (п. 2.3 — 2.4 и 2.5.1);
- CF+P — алгоритм на основе клеточного автомата с пирамидальным смешиванием (п. 2.3 — 2.4 и 2.5.2).

Для каждого из алгоритмов следует определить диапазон значений параметров. При выборе граничных значений этих диапазонов следует учитывать

Таблица 3.2. Диапазоны параметров алгоритмов смешивания для проведения эксперимента по сравнению алгоритмов оценки качества

Алгоритм	Параметр	Диапазон значений	Шаг
BF	Размер квадратного блока	[5 : 75]	10
WT	Порядок вейвлета Добеши	$D4 - D18$	два порядка
CF	Порог бинаризации α	[0, 01 : 0, 15]	0, 02
CF+P	Порог бинаризации α	[0, 01 : 0, 15]	0, 02

особенности работы конкретного алгоритма. Так, например, не следует выбирать значения размеров блока более 20% длины стороны в блочном алгоритме смешивания, или количество уровней пирамиды более 4 в алгоритмах на основе пирамидального смешивания. Шаг, с которым будет происходить изменение значений параметров, следует выбирать из соображений визуальной различимости последовательно сформированных изображений с точки зрения человеческого восприятия. Выбранные значения диапазонов параметров приведены в таблице 3.2.

Согласно данным таблицы 3.2 на каждый алгоритм приходится 8 значений параметров. Таким образом общее число тестовых изображений составляет 4 (сцены) \times 4 (алгоритмы) \times 8 (параметры) = 128 изображений.

3.3.2. Постановка эксперимента

Проведение эксперимента по сбору субъективных оценок сформированных изображений с расширенной глубиной резкости является важнейшим этапом сравнения качества работы алгоритмов смешивания. В ходе эксперимента каждому изображению из тестовой базы, созданной в 3.3.1, необходимо сопоставить набор субъективных оценок экспертов, отражающих общее качество оцениваемого изображения.



Рис. 3.6. Графический интерфейс пользователя для субъективной оценки качества изображений с расширенной глубиной резкости

Графический интерфейс пользователя

Для ускорения процесса сбора субъективных оценок разработан графический интерфейс пользователя (GUI). Интерфейс содержит элементы позволяющие в доступной и удобной форме оценить предлагаемое изображение и поставить ему соответствующую оценку. Внешний вид интерфейса приведен на рис. 3.6.

На рис. 3.6 обозначены следующие основные элементы управления:

1. *Область отображения оцениваемого изображения.* В данной области происходит отображение оцениваемого в данный момент изображения. Область отображения меняет размер в зависимости от размера отображаемого изображения, но составляет не более чем 800×600 пикселей. В случае, если изображение имеет больший размер, размер отображаемого изображения меняется, чтобы помещаться в указанное разрешение, с

сохранением соотношения сторон.

2. *Кнопки оценки текущего изображения.* Здесь расположен набор из десяти кнопок, на каждой из которых отображена соответствующая оценка. При нажатии на кнопку, в соответствие изображению ставится данная оценка и сохраняется во временную базу данных, хранящуюся в оперативной памяти. Также при нажатии на кнопку в области отображения появляется следующее изображение для оценки.
3. *Кнопка сохранения оценок.* Кнопка предназначена для перенесения сделанных оценок из временной базы данных в постоянную, хранящуюся на жестком диске. При нажатии происходит обращение к постоянной базе данных, оценки полученные в ходе последнего сеанса дополняются к уже имеющимся в базе.

Графический интерфейс пользователя реализован на базе модуля GUIDE среды MATLAB [84].

База данных эксперимента

Для удобства обработки полученных оценок, а также оптимизации времени расчета необходимых характеристик создана специализированная база данных. Технически база представляет собой файл `stat_data.mat`, содержащий набор полей и данных, совместимый со средой разработки MATLAB. На рис. 3.7 изображена часть сформированной базы данных с указанием названий полей и некоторых возможных значений.

База данных содержит следующие поля:

1. *index* — индекс тестового изображения;
2. *votes* — массив всех экспертных оценок, выставленных данному изображению;

index	votes	method	params	set	qaf	mi	wab
1	[2 2]	'cell'	0.0100	'toys'	0.5591	5.4199	0.6819
2	[4 6 7]	'cellpyr'	[0.0100 4]	'toys'	0.4195	4.9210	0.7516
3	[3 2 3 3]	'cell'	0.0300	'toys'	0.4678	5.7780	0.6804
4	[7 6 6 6 6 7 ...]	'cellpyr'	[0.0300 4]	'toys'	0.4221	4.8581	0.7389
5	[5 4 4 4 4]	'cell'	0.0500	'toys'	0.4500	5.7962	0.6785
6	[6 4 4]	'cellpyr'	[0.0500 4]	'toys'	0.4213	4.8929	0.7242
7	[6 5 6 5 5]	'cell'	0.0700	'toys'	0.4378	5.7702	0.6753
8	[6 5 7 7 4 5]	'cellpyr'	[0.0700 4]	'toys'	0.4201	4.9137	0.7071
9	[2 1 2 2 2]	'cell'	0.0900	'toys'	0.4282	5.6426	0.6717
10	[4 4 4 3 3 3 ...]	'cellpyr'	[0.0900 4]	'toys'	0.4162	4.8281	0.6943
11	[2 4 4 4 3]	'cell'	0.1100	'toys'	0.4129	5.3906	0.6672
12	[6 7 5 4 4]	'cellpyr'	[0.1100 4]	'toys'	0.4061	4.5804	0.6825
13	[5 4 5 4 3 3]	'cell'	0.1300	'toys'	0.4010	5.2317	0.6639
14	[4 4 3 3 4 4 ...]	'cellpyr'	[0.1300 4]	'toys'	0.3964	4.4065	0.6766
15	[4 3 3 3]	'cell'	0.1500	'toys'	0.3960	5.2162	0.6611
16	[3 3 3 3 4 3 ...]	'cellpyr'	[0.1500 4]	'toys'	0.3927	4.4312	0.6703

Рис. 3.7. Пример данных, содержащихся в базе `stat_data.mat`

3. *method* — алгоритм смешивания, при помощи которого получено данное изображение;
4. *params* — список параметров, с которыми был запущен алгоритм смешивания;
5. *set* — название сцены;
6. *qaf* — значение метрики $Q^{AB/F}$ для данного изображения;
7. *mi* — значение метрики взаимной смешанной информации для данного изображения;
8. *wab* — значение метрики модифицированного универсального индекса качества для данного изображения.

Значения оценки качества изображений, полученные при помощи автоматических алгоритмов, также внесены в базу данных для ускорения последующей обработки без повторного расчета.

3.3.3. Обработка данных эксперимента

Сформированная база данных содержит набор оценок, выставленных экспертами и автоматическими алгоритмами. Эти данные имеют различный характер и не могут сравнены непосредственно. На данном этапе необходимо выбрать характеристику, отражающую схожесть множеств оценок, а также способ приведения данных к единому типу и диапазону значений.

Коэффициент корреляции Спирмена

В качестве меры схожести множеств оценок, собранных из различных источников, хорошо подходят различные корреляционные методы. Так как оценки экспертов, носят ранговый характер, то коэффициент ранговой корреляции Спирмена выглядит подходящим инструментом для выполнения задачи сравнения.

Коэффициент корреляции Спирмена [85] — мера линейной связи между случайными величинами. Корреляция Спирмена является ранговой, то есть для оценки силы связи используются не численные значения, а соответствующие им ранги. Коэффициент инвариантен по отношению к любому монотонному преобразованию шкалы измерения. Пусть имеются две выборки $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$, тогда коэффициент определяется следующим образом:

$$\rho = 1 - \frac{6}{n(n-1)(n+1)} \sum_{i=1}^n (R_i - S_i)^2,$$

где ρ — коэффициент корреляции Спирмена, принимает значения в диапазоне $[-1; 1]$, R_i — ранг наблюдения x_i в ряду x , S_i — ранг наблюдения y_i в ряду y . Равенство $\rho = 1$ указывает на строгую прямую линейную зависимость, $\rho = -1$ на обратную.

Приведение оценок

Оценки качества изображений, полученные при помощи различных автоматических алгоритмов и экспертов, имеют различный характер распределения, диапазоны значений и тип. Так, например, экспертные оценки представляют собой дискретный ряд целых чисел в диапазоне $[1; 10]$, в то время как оценки, полученные при расчете $Q^{AB/F}$ -метрики, имеют непрерывный ряд значений, в общем случае, имеющий диапазон $(0; \infty)$.

Приведение оценок к общему ранговому виду, в рамках работы, включает следующие этапы:

1. Для экспертных оценок вычисляется среднее значение для каждого изображения. При этом учитывается возможность наличия статистических выбросов по правилу 3σ . Так например, при средней оценке 7 и $\sigma = 1$, экспертные оценки в 1 – 3 балла исключаются из рассмотрения.
2. Для всех оценок (экспертных и автоматических) строятся гистограммы распределения с десятью ячейками по всему диапазону имеющихся значений.
3. Каждой оценке в соответствие ставится номер ячейки гистограммы, в который попало ее значение, начиная с единицы.

После проведения приведенной выше операции, каждой оценке поставлен в соответствие некоторый ранг в диапазоне $[1; 10]$, который в дальнейшем будет использован при расчете коэффициента корреляции Спирмена.

Расчет коэффициентов корреляции

Расчет коэффициентов корреляции Спирмена может быть произведен для различных рядов данных. В рамках работы рассмотрены корреляции для всей тестовой базы в целом, так и для отдельных сцен. На рис. 3.8 приведено графическое представление коэффициентов корреляции для различных автоматиче-

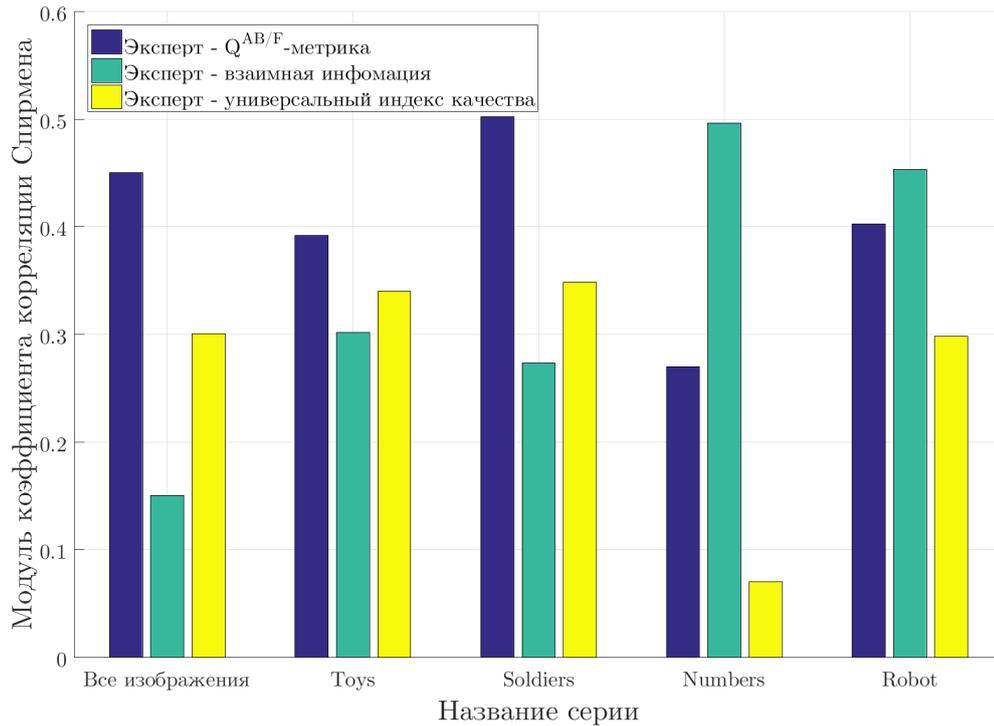


Рис. 3.8. Коэффициент корреляции Спирмена между автоматическими и экспертными оценками для различных тестовых сцен

ских алгоритмов оценки качества и экспертных оценок для всей тестовой базы и отдельно для каждой из рассмотренных сцен.

Приведенные на рис. 3.8 данные показывают, что рассчитанный коэффициент корреляции имеет среднее значение $0,4 - 0,6$, что можно характеризовать как наличие слабой или средней корреляции, таким образом представленные метрики обладают невысокой точностью неэталонной оценки качества изображений с расширенной глубиной резкости. В связи с этим наиболее корректным будет использование среднего балла субъективной оценки экспертов, полученного в ходе проведения эксперимента. Средние значения оценок приведены в таблице 3.3. Также распределение оценок можно оценить при помощи накопительной гистограммы количества оценок, приведенной на рис. 3.9. Как видно из графика, предлагаемый алгоритм имеет наибольшее смещение в правый нижний угол, что говорит о более высоких экспертных оценках.

Таблица 3.3. Средние значения субъективных оценок качества сформированных изображений с расширенной глубиной резкости

Алгоритм	Средняя оценка
Блочный алгоритм	3,9
Вейвлет-преобразование	4,0
Клеточный автомат	3,6
Пирамиды изображений	4,5

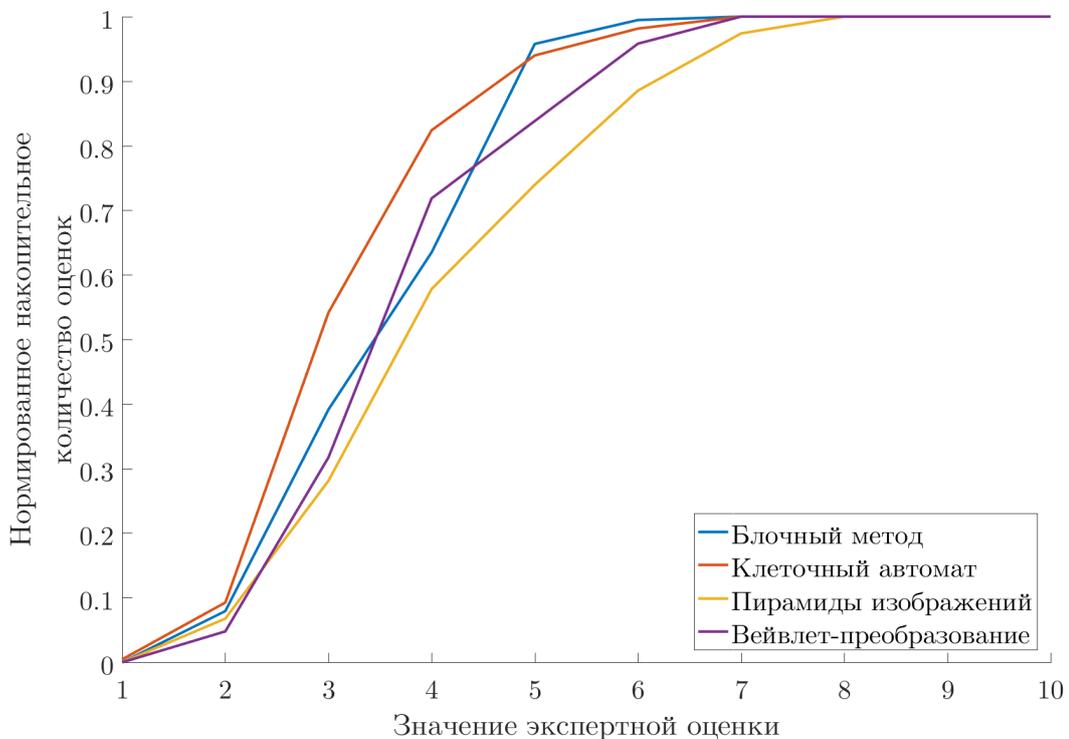


Рис. 3.9. Распределение нормированное накопительного количества субъективных оценок

3.4. Краткие выводы

В третьей главе приведено исследование разработанного алгоритма формирования изображений с расширенной глубиной резкости и его модификации на основе пирамид изображений. Проведено исследование влияния аддитивного белого гауссовского шума на работу алгоритма, а также сравнение разра-

ботанного алгоритма с известными на основе экспертных оценок и ранговой корреляции Спирмена.

Приведены рекомендации по выбору подхода к подавлению АБГШ при формировании изображений с расширенной глубиной резкости. Установлено, что применение фильтрации к изображениям исходной серии показывает наилучшие результаты с точки зрения средних значений эталонных метрик качества на основе пикового отношения сигнал/шум (улучшение на 5% и 24% по сравнению с постфильтрацией и без фильтрации соответственно) и коэффициента структурного подобия (4% и 47% соответственно). Проведен эксперимент по сравнению неэталонных метрик оценки качества изображений с расширенной глубиной резкости на основе ранговой корреляции Спирмена. Установлено, что существующие автоматические неэталонные метрики оценки качества имеют средний уровень корреляции $\rho = 0,4 - 0,6$ с экспертными оценками. Можно выделить характерные особенности сцен, при наличии которых та или иная метрика демонстрирует существенно лучшие результаты, чем иные. Взаимная информация на 73% лучше справляется со сценой, имеющей малое количество высоких значений градиента, чем $Q^{AB/F}$ -метрика. Установлено, что разработанный модифицированный алгоритм на основе пирамид изображений имеет в среднем наибольший балл согласно экспертным оценкам (4,5 балла).

Заключение

Основные выводы и результаты диссертационной работы можно сформулировать в следующем виде:

1. На основе анализа современных научно-технических источников в качестве правила смешивания в алгоритме формирования изображений с расширенной глубиной резкости выбран математический аппарат клеточных автоматов.
2. Разработан и протестирован алгоритм формирования изображений с расширенной глубиной резкости на основе клеточных автоматов.
3. Проведен эксперимент по сравнению различных метрик сфокусированности, в рамках которого предложены идеализированные характеристики метрик сфокусированности для рассматриваемой задачи.
4. Установлено, что наиболее точной метрикой оценки сфокусированности является метрика на основе соотношения коэффициентов вейвлет-преобразования, для которой коэффициент корреляции с идеализированными метриками $r = 0,88$.
5. Установлено, что наиболее эффективной с точки зрения скорости, а также весьма эффективной с точки зрения точности ($r = 0,75$) является метрика сфокусированности на основе расчета вариации уровня серого.
6. Среди алгоритмов, основанных на сверточных операциях, наилучшие результаты по выделению сфокусированных областей показывает алгоритм вычисления метрики Тененграда с использованием маски Робертса ($r = 0,62$). Этот алгоритм примерно сопоставим по вычислительной сложности с остальными сверточными алгоритмами и может использоваться в задачах реального времени.

7. Установлено, что диапазон $[0, 03; 0, 12]$ значения пороговой бинаризации в алгоритме смешивания на основе клеточного автомата соответствует наиболее качественному смешиванию изображений исходной серии.
8. Разработана и протестирована модификация алгоритма смешивания изображений с расширенной глубиной резкости на основе пирамид гауссианов и лапласианов.
9. Проведено исследование работы алгоритма при наличии аддитивного белого гауссовского шума. Предложены подходы к уменьшению его влияния на результат смешивания.
10. Установлено что подход, подразумевающий предфильтрацию исходной серии изображений, показывает наилучший результат среди рассмотренных подходов.
11. Проведен эксперимент по сравнению неэталонных метрик оценки качества изображений с расширенной глубиной резкости на основе ранговой корреляции Спирмена.
12. Установлено, что существующие автоматические неэталонные метрики оценки качества имеют средний уровень корреляции ($\rho = 0,4 - 0,6$) с экспертными оценками.
13. Можно выделить характерные особенности сцен, при наличии которых та или иная метрика демонстрирует существенно лучшие результаты, чем иные. Например, взаимная информация на 73% лучше справляется со сценой, имеющей малое количество высоких значений градиента, чем $Q^{AB/F}$ -метрика.
14. Установлено, что разработанный модифицированный алгоритм на основе пирамид изображений имеет в среднем наилучший показатель (4,5 балла), согласно экспертным оценкам.

15. Цель и задачи диссертационной работы успешно выполнены. Разработанные алгоритмы реализованы на языках Matlab и C++ в программе RicFocus 2.0. Получены три свидетельства о государственной регистрации программы для ЭВМ.

Список сокращений и условных обозначений

ISO	—	светочувствительность снимающего устройства в соответствии со стандартом ISO 12232:2006.
NLM	—	Non-Local Means, фильтрация на основе нелокального среднего.
BM3D	—	Block-Matching and 3D filtering.
DOF	—	Depth Of Field, глубина резко изображаемого пространства.
ДКП	—	дискретное косинусное преобразование.
КПБД	—	контурлет-преобразование без децимации.
ИНН	—	импульсная нейронная сеть.
ДВП	—	дискретное вейвлет-преобразование.
СКО	—	средняя квадратичная ошибка.
КСКО	—	корень из СКО.
ПОСШ	—	пиковое отношение сигнал/шум.
ВИ	—	взаимная информация.
КСП	—	коэффициент структурного подобия.
ВСИ	—	взаимная смешанная информация.
СПРВ	—	совместная плотность распределения вероятностей.
МУИК	—	модифицированный универсальный индекс качества.
ИКС	—	индекс качества смешивания.
SIMD	—	Single Instruction for Multiple Data, методы оптимизации вычислений одной инструкции для большого набора данных.
RGB	—	Red, Green, Blue. Система цветового кодирования на основе красного, зеленого и синего цветов.
АЦМ	—	абсолютный центральный момент.
ВУС	—	вариация уровня серого.
ОКВП	—	отношение коэффициентов вейвлет-преобразования.
ТИТ	—	тестовая испытательная таблица

АБГШ — аддитивный белый гауссовский шум.

Список литературы

1. Sroubek F., Gabarda S., Redondo R. et al. Multifocus Fusion with Oriented Windows // Academy of Sciences, CSIC.
2. Malik A. S., Choi T.-S., Nisar H. Depth Map and 3D Imaging Applications // Algorithms and Technologies. 2011. no. 285.
3. Wang J., Cohen M. Optimized color sampling for robust matting // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, USA, 2007. Minneapolis, MN, USA: IEEE, 2007. P. 1–9.
4. Kekre H. B., Sarode T., Dhannawat R. Implementation and Comparison of different Transform Techniques using Kekre's Wavelet Transform for Image Fusion // International Journal of Computer Applications. 2012. Vol. 44, no. 10. P. 41–48.
5. Kekre H. B., Athawale A., Sadavarti D. Algorithm to Generate Kekre's Wavelet Transform from Kekre's Transform // International Journal of Engineering Science and Technology. 2010. Vol. 2. P. 756–767.
6. Xydeas C., Petrovic V. Objective image fusion performance measure // Electronics Letters. 2000. no. 36. P. 308–309.
7. Stathaki T. Image Fusion: Algorithms and Applications. Academic Press, 2008.
8. Fergus R., Singh B., Hertzmann A. et al. Removing camera shake from a single image // Proceedings of ACM Transactions on Graphics (SIGGRAPH), Boston, USA, 2006. Boston, USA, 2006: 2007. P. 787–794.
9. Shan Q., Jia J., Agarwala A. High-quality motion deblurring from a single image // ACM Transactions on Graphics (SIGGRAPH). New York, NY, USA: 2008.
10. Xu L., Jia J. Two-phase kernel estimation for robust motion deblurring // Computer Vision–ECCV 2010. 2010. P. 157–170.
11. Harmeling S., Hirsch M., Scholkopf B. Space-variant single-image blind deconvolution for removing camera shake // Advances in Neural Information

- Processing Systems. 2010. Vol. 6. P. 829–837.
12. Whyte O., Sivic J., Zisserman A., Ponce J. Non-uniform deblurring for shaken images // *International journal of computer vision*. 2012. Vol. 98, no. 2. P. 168–186.
 13. Born M., Wolf E. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. CUP Archive, 1999.
 14. Favaro P., Soatto S. A geometric approach to shape from defocus // *IEEE Trans. Pattern Analysis and Machine Intelligence*. 2005. Vol. 27. P. 406–417.
 15. Favaro P. Seeing beyond occlusions (and other marvels of a finite lens aperture) // In *IEEE CVPR*. 2003. P. 579–586.
 16. Middleton W. E. K. *Vision Through the Atmosphere*. Canada: University of Toronto Press, 1958.
 17. Roggemann M. C., Welsh B. M. *Imaging through turbulence*. Boca Raton, Fla.: CRC Press, 1996.
 18. Li D., Mersereau R. M., Simske S. Atmospheric turbulence-degraded image restoration using principal components analysis // *IEEE Geoscience and Remote Sensing Letters*. 2007. Vol. 4, no. 3. P. 340–344.
 19. Shimizu M., Yoshimura S., Tanaka M., Okutomi M. Super-resolution from image sequence under influence of hot-air optical turbulence // *Computer Vision and Pattern Recognition*. 2008. P. 1–8.
 20. Tubbs R. N. *Lucky exposures: Diffraction limited astronomical imaging through the atmosphere: Ph.D. thesis / Cambridge University*. 2003.
 21. Tian Y., Narasimhan S. Seeing through water: Image restoration using model-based tracking // *IEEE International Conference on Computer Vision (ICCV)*. 2009. P. 2303–2310.
 22. Vorontsov M. A., Carhart G. W. Anisoplanatic imaging through turbulent media: image recovery by local information fusion from a set of short-exposure images // *IEEE International Conference on Computer Vision (ICCV)*. 2009. P. 2303–2310.

23. Roggemann M. C., Stoudt C. A., Welsh B. M. Image-spectrum signal-to-noise-ratio improvements by statistical frame selection for adaptive-optics imaging through atmospheric turbulence // *Optical Engineering*. 2009. Vol. 33, no. 10. P. 3254–3264.
24. Fried D. L. Probability of getting a lucky short-exposure image through turbulence // *Optical Society of America*. 1978. Vol. 68. P. 1651–1658.
25. Tomasi C., Manduchi R. Bilateral filtering for gray and color images // *Proceeding of the International Conference of Compute Vision*. 1998. Vol. 68. P. 836–846.
26. Morillas S., Gregori V., Sapena A. Fuzzy bilateral filtering for color images // *Image analysis and recognition, ser. Lecture Notes in Computer Science*. 2006. P. 138–145.
27. Kao W.-C., Chen Y.-J. Multistage bilateral noise filtering and edge detection for color image enhancement // *IEEE Transactions on Consumer Electronics*. 2005. Vol. 51, no. 4. P. 1346–1351.
28. Chatterjee P., Milanfar P. Generalization of non-local mean via kernel regression // *SPIE conference on Computational Imaging*. 2008. P. 3–12.
29. Dabov K., Foi A., Katkovnik V., Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering // [IEEE Transactions on Image Processing](#). 2007. Vol. 16, no. 8. P. 2080–2095.
30. Nosratinia A. Postprocessing of jpeg-2000 images to remove compression artifacts // *Signal Processing Letters*. 2003. Vol. 10, no. 10. P. 296–299.
31. Wang Z., Sheikh H. R., Bovik A. C. No-reference perceptual quality assessment of jpeg compressed images // *International Conferencing on Image Processing*. 2002. Vol. 1. P. 477–480.
32. Shen C.-H., Chen H. H. Robust focus measure for low-contrast images // [Consumer Electronics](#). 2006. P. 69–70.
33. Li H., Chai Y., Yin H., Liu G. Multifocus image fusion and denoising scheme based on homogeneity similarity // *Optics Communications*. 2012. no. 285.

- P. 91–100.
34. Firestone L., Cook K., Culp K. et al. Comparison of autofocus methods for automated microscopy // *Cytometry*. 1991. no. 12.
 35. Santos A., de Solorzano C. O., Vaquero J. et al. Evaluation of autofocus functions in molecular cytogenetic analysis // *Journal of Microscopy*. 1997. no. 188. P. 264–272.
 36. Liu X. Y., Wang W. H., Sun Y. Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear // *Journal of Microscopy*. 2007. no. 227. P. 15–23.
 37. Nixon M., Aguado A. *Feature Extraction and Image Processing*. 2008.
 38. Geusebroek J. M., Cornelissen F., Smeulders A. W. M., Geerts H. Robust autofocusing in microscopy // *Cytometry*. 2000. no. 39. P. 1–9.
 39. Groen F., Young I., Ligthart G. A comparison of different focus functions for use in autofocus algorithms // *Cytometry*. 1985. no. 6. P. 81–91.
 40. Yousefi S., Rahman M., Kehtarnavaz N. A new auto-focus sharpness function for digital and smartphone cameras // *IEEE Trans. Consum. Electron.* 2011. Vol. 57, no. 6. P. 1003–1009.
 41. Agrawal D., Al-Dahoud A., Singhai J. Modified partition fusion technique of multifocus images for improved image quality // *UbiCC Journal*. 2009. Vol. 4, no. 3. P. 120–128.
 42. R. Hong C. W., Wang M., Sun F. Saliency preserving multi focus image fusion with dynamic range compression // *International journal of innovative*. 2009. Vol. 5, no. 8. P. 35–43.
 43. Li S., Kwok J. T., Wang Y. Multifocus image fusion using artificial neural networks // *Pattern Recognition Letters*. 2002. no. 23. P. 985–997.
 44. Subbarao M., Choi T., Nikzad A. Focusing techniques // *Optical Engineering*. 1993. no. 32. P. 2824–2836.
 45. Flamm D., Schulze C., Bruning R. et al. Fast M2 measurements for fiber beams based on modal analysis // *Application Optics*. 2012. no. 51. P. 987–993.

46. Perez-Vizcaino J., Mendoza-Yero O., Martinez-Cuenca R. et al. Free-motion beam propagation factor measurement by means of a liquid crystal spatial light modulator // *Display Technologies*. 2012. no. 8. P. 539–545.
47. Schulze C., Flamm D., Duparre M. et al. Beamquality measurements using a spatial light modulator // *Optics Letters*. 2012. no. 8. P. 4687–4689.
48. Nayar S. K., Nakagawa Y. Shape from focus // *IEEE Translation Pattern Analysis Machine intellect*. 1994. no. 16. P. 824–831.
49. Pentland A. A new sense for depth of field // *IEEE Translation Pattern Analysis Machine intellect*. 1987. no. 9. P. 523–531.
50. K. Konnai Y. H. Method of capturing image of object passing through focus area // *Proceeding of SPIE – The International Society for Optical Engineering*. 2001. P. 86–94.
51. Xie H., Ronga W., Sun L. Wavelet-Based Focus Measure and 3-D Surface Reconstruction Method for Microscopy Images // *RSJ International Conference on Intelligent Robots and Systems*. 2006. P. 229–234.
52. Mallat W. *A Wavelet Tour of Signal Processing*. Academic Press.
53. Candes E. J., Donoho D. L. *Curvelets – a surprisingly effective nonadaptive representation for objects with edges*. Saint-Malo: Vanderbilt University Press.
54. Mahmood M., Choi W., Choi T. PCA-based method for 3D shape recovery of microscopic objects from image focus using discrete cosine transform // *Microscopic Research Techniques*. 2008. no. 71. P. 897–907.
55. Xu X., Wang Y., Zhang X. et al. A comparison of contrast measurements in passive autofocus systems for low contrast images // *Multimedia Tools Applications*. 2012. no. 71. P. 139–156.
56. da Cunha A. L., Zhou J., Do M. N. *The Nonsampled Contourlet Transform: Theory, Design, and Applications* // *Image Processing*. 2006. Vol. 15. P. 3089–3101.
57. Johnson J., Padgett M. *PCNN models and applications* // *Neural Networks*. 1999. Vol. 10. P. 480–498.

58. Maass W. Networks of Spiking Neurons: The Third Generation of Neural Network Models // *Neural Networks*. 1997. no. 10. P. 1659–1671.
59. Maass W. *Computing with spiking neurons*. MIT Press, Cambridge.
60. Qu X., Yan J. Multi-focus Image Fusion Algorithm Based on Regional Firing Characteristic of Pulse Coupled Neural Networks // [Bio-Inspired Computing: Theories and Applications](#). 2007. P. 62–66.
61. Тоффоли Т., Марголюс Н. *Машины клеточных автоматов*. М: Мир, 1991.
62. Hoekstra A., J. Kroc P. S. *Simulating complex systems by cellular automata*. Springer, 2010.
63. Wolfram S. *A New Kind of Science*. Wolfram Media, 2002.
64. Nuttall A. H. Some windows with very good sidelobe behavior // *IEEE Transactions on Acoustics, Speech, Signal Processing*. 1981. — February. P. 84–91.
65. Ando T. Majorization relations for hadamard products // *Linear Algebra and its Applications*. 1995. P. 57–64.
66. Abramowitz M., Stegun I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1972.
67. Shirvaikar M. An optimal measure for camera focus and exposure // [South-eastern Symposium on System Theory](#). 2004. P. 472–475.
68. Xie H., Rong W., Sun L. Wavelet-based focus measure and 3-d surface reconstruction method for microscopy images // [International Conference on Intelligent Robots and Systems](#). 2006. P. 229–234.
69. Otsu N. A Threshold Selection Method from Gray-Level Histograms // [Systems, Man and Cybernetics, IEEE Transactions](#). 1979. — Jan. Vol. 9, no. 1. P. 62–66.
70. Burt P., Adelson E. The laplacian pyramid as a compact image code // [IEEE Transactions on Communications](#). 1983. P. 532–540.
71. Katkovnik V., Foi A., Egiazarian K., Astola J. From local kernel to nonlocal multiple-model image denoising // *Int. J. Computer Vision*. 2010. Vol. 86, no. 8. P. 1–32.

72. Marsland S. Machine learning: an algorithmic perspective. Chapman and Hall, 2009.
73. Muresan D. D., Parks T. W. Adaptive principal components and image denoising // Proc. IEEE Int. Conf. Image Processing. 2003. Vol. 1. P. 101–104.
74. Priorov A., Tumanov K., Volokhov V. Efficient denoising algorithms for intelligent recognition systems. Vol. 75.
75. Волохов В. А., Сергеев Е. В., Приоров А. Л. Двухэтапная процедура фильтрации изображений на основе анализа главных компонент // Проектирование и технологии электронных средств. 2011. № 2. С. 40–44.
76. Сэломон Д. Сжатие данных, изображений и звука. М.: Техносфера, 2004.
77. Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image quality assessment: from error visibility to structural similarity. 2004. Vol. 13, no. 4. P. 600–612.
78. Матросов М. А. Методы построения изображений расширенной глубины резкости: Кандидатская диссертация / МГУ им. М. В. Ломоносова. 2009.
79. Petrovic V., Xydeas C. Sensor noise effects on signal-level image fusion performance // Information Fusion. 2003. Vol. 4. P. 167–183.
80. Petrovic V., Xydeas C. Objective evaluation of signal-level image fusion performance // [Optical Engineering](#). 2005. — Aug. no. 44.
81. Xydeas C., Petrovic V. Objective pixel-level image fusion performance measure // [Proceedings of SPIE 4051](#). 2000. P. 89–99.
82. Wang Z., Bovik A. C. A universal image quality index // [IEEE Signal Processing Letters](#). 2002. — Mar. Vol. 9, no. 3. P. 81–84.
83. Piella G. New quality measures for image fusion // Proceedings of the Seventh International Conference on Information Fusion / Ed. by P. Svensson, J. Schubert. Vol. I. Mountain View, CA: International Society of Information Fusion, 2004. — Jun. P. 542–546.
84. Дьяконов В. П. MATLAB и SIMULINK для радиоинженеров. М.: «ДМК-Пресс», 2011. ISBN: [978-5-94074-492-4](#).
85. Кобзарь А. И. Прикладная математическая статистика. М: Физматлит,

2006.

86. Носков А. А., Волохов В. А., Приоров А. Л., Апальков И. В. Синтез полностью сфокусированных изображений при наличии шумов // Цифровая обработка сигналов. 2017. № 3. С. 67–72.
87. Носков А. А., Аминова Е. А., Приоров А. Л. Анализ метрик оценки сфокусированности для задачи формирования полностью сфокусированных изображений // Успехи современной радиоэлектроники. 2016. № 6. С. 22–29.
88. Носков А. А., Аминова Е. А., Приоров А. Л. Формирование полностью сфокусированных изображений на основе клеточных автоматов и пирамид изображений // Успехи современной радиоэлектроники. 2016. С. 39–46.
89. Трапезников И. Н., Приоров А. Л., Носков А. А., Аминова Е. А. Комбинированный алгоритм детектирования автомобильных регистрационных знаков // Электромагнитные волны и электронные системы. 2014. № 1. С. 32–37.
90. Носков А. А., Аминова Е. А., Приоров А. Л., Трапезников И. Н. Система определения повторного детектирования объектов на основе нейронной сети // Нейрокомпьютеры: разработка, применение. 2014. № 3. С. 36–43.
91. Noskov A., Volokhov V., Aminova E. Multi-focus image fusion based on cellular automata method // Proceeding of the 17th conference of FRUCT association. 2015. P. 136–141.
92. Noskov A., Aminova E., Volokhov V. Comparison of image focus assessment methods for multi-focused image construction // Proceeding of the 18th conference of FRUCT association. 2016. P. 226–232.
93. Noskov A., Priorov A. Application of rank correlation at multi-focused image fusion quality assessment // Proceeding of the Systems of Signal Synchronization, Generating and Processing in Telecommunications. 2017. P. 1–5.
94. Носков А. А., Аминова Е. А. Синтез сфокусированных изображений на основе бинарных масок // Материалы XX международной научно-техни-

- ческой конференции студентов и аспирантов "Радиоэлектроника, электроника и энергетика". 2014. Т. 1. С. 120.
95. Носков А. А., Аминова Е. А., Трапезников И. Н., Приоров А. Л. Применение метода бинарных масок для синтеза полностью сфокусированных изображений // Тезисы научно-технической конференции "Техническое зрение в системах управления – 2014". 2014. С. 153–155.
96. Носков А. А. Синтез полностью сфокусированных изображений на основе клеточных автоматов // Тезисы докладов 67-й всероссийской научно-технической конференции студентов, магистрантов и аспирантов ВУЗов с международным участием. 2014. С. 296.
97. Носков А. А. Применение клеточных автоматов в задаче синтезирования полнофокусных изображений // Сборник научных трудов по материалам международной научно-практической конференции. 2014. С. 114–116.
98. Носков А. А. Алгоритм формирования полнофокусных изображений на основе клеточного автомата // Тезисы научно-технической конференции "Техническое зрение в системах управления – 2015". 2015. С. 56–57.
99. Носков А. А., Волохов В. А. Обзор методов смешивания полнофокусных изображений и их применение // Сборник докладов XV Всероссийской научно-практической конференции «Проблемы развития и применения средств противовоздушной обороны на современном этапе». 2014. С. 84–92.
100. Носков А. А. Алгоритм формирования полностью сфокусированных изображений на основе клеточных автоматов // Цифровая обработка сигналов и ее применение (DSPA–2015) : докл. 17-й междунар. конф. 2015. С. 724–728.
101. Носков А. А., Аминова Е. А., Волохов В. А. Сравнение метрик оценки сфокусированности для задачи формирования полностью сфокусированных изображений // Цифровая обработка сигналов и ее применение (DSPA–2016): докл. 18-й междунар. конф. 2016. С. 729–734.
102. Носков А. А. Система формирования полностью сфокусированных изображений

- ражений на основе клеточного автомата и пирамид изображений // Цифровая обработка сигналов и ее применение (DSPА–2017): докл. 19-й международн. конф. 2017. С. 296.
103. Носков А. А., Аминова Е. А., Трапезников И. Н. PicFocus – программа для формирования полнофокусных цифровых изображений // Свидетельство о государственной регистрации программы для ЭВМ № 2014615039 от 15 мая 2014.
104. Носков А. А., Аминова Е. А., Приоров А. Л. PicFocus2.0 – программа для формирования и оценки качества полнофокусных цифровых изображений // Свидетельство о государственной регистрации программы для ЭВМ № 2015617434 от 9 июля 2015.
105. Носков А. А., Аминова Е. А., Приоров А. Л. FocusExpert – программа для оценки качества сфокусированности цифровых изображений // Свидетельство о государственной регистрации программы для ЭВМ № 2016613017 от 15 марта 2016.

Приложение А

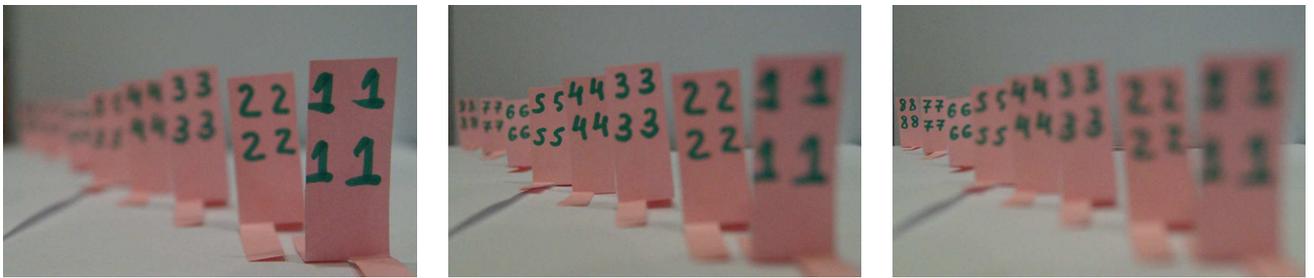
Примеры изображений тестовых серий



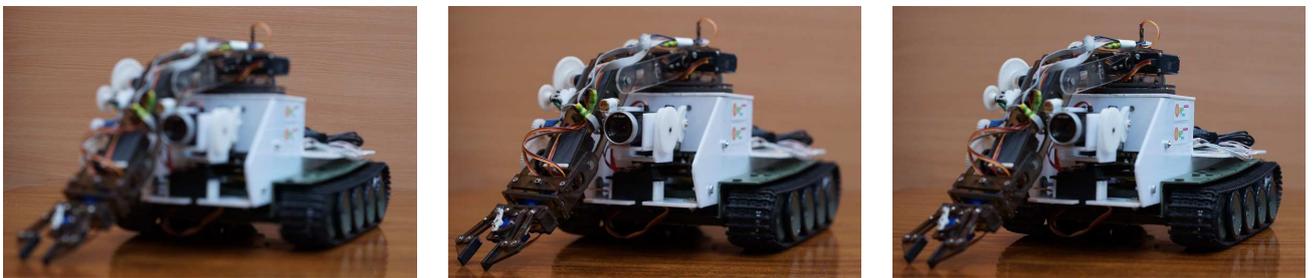
а)



б)



в)



г)

Рис. А.1. Примеры изображений из собственных тестовых серий: а) «toys»; б) «soldiers»; в) «numbers»; г) «robot»

Приложение Б

Свидетельства о государственной регистрации
программ для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ
№ 2014615039

**PicFocus - программа для формирования полнофокусных
цифровых изображений**

Правообладатели: *Аминова Елена Альбертовна (RU), Носков Андрей
Александрович (RU), Трапезников Илья Николаевич (RU)*

Авторы: *Аминова Елена Альбертовна (RU), Носков Андрей
Александрович (RU), Трапезников Илья Николаевич (RU)*

Заявка № **2014612915**
Дата поступления **27 марта 2014 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **15 мая 2014 г.**

Руководитель Федеральной службы
по интеллектуальной собственности



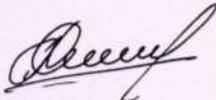
 **Б.П. Симонов**

Рис. Б.1. Свидетельство о государственной регистрации программ для ЭВМ №2014615039



Рис. Б.2. Свидетельство о государственной регистрации программ для ЭВМ №2015617434

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2016613017

**FocusExpert - программа для оценки качества
сфокусированности цифровых изображений**

Правообладатели: *Аминова Елена Альбертовна (RU), Носков Андрей
Александрович (RU), Приоров Андрей Леонидович (RU)*

Авторы: *Аминова Елена Альбертовна (RU), Носков Андрей
Александрович (RU), Приоров Андрей Леонидович (RU)*

Заявка № 2016610377

Дата поступления 19 января 2016 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 15 марта 2016 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Ивлиев Г.П. Ивлиев



Рис. Б.3. Свидетельство о государственной регистрации программ для ЭВМ №2016613017

Приложение В

Акты внедрения результатов работы



Рис. В.1. Акт внедрения в научно-исследовательские работы ЯрГУ

«УТВЕРЖДАЮ»

Первый проректор Ярославского

Государственного университета

И.М.Т. Дежидова

С.А. Кащенко

29 _____ 2017 г.



АКТ

внедрения результатов диссертационной работы Носкова Андрея Александровича на тему «Формирование изображений с расширенной глубиной резкости для систем прикладного телевидения» в учебный процесс

Мы, нижеподписавшиеся, заведующий кафедрой инфокоммуникаций и радиофизики, профессор, д.т.н. Брюханов Ю.А. и доцент кафедры инфокоммуникаций и радиофизики, к.т.н. Волохов В.А. составили настоящий акт о том, что результаты диссертационной работы Носкова А.А. внедрены в учебный процесс на кафедре инфокоммуникаций и радиофизики физического факультета ЯрГУ (направление «Радиотехника»)

- в курсе «Цифровая обработка изображений» – алгоритм формирования изображений с расширенной глубиной резкости на основе клеточных автоматов;
- в курсе «Системы технического зрения» – результаты работы алгоритма формирования полностью сфокусированных изображений в условиях наличия аддитивного белого гауссовского шума.

Заведующий кафедрой
инфокоммуникаций и радиофизики,
д.т.н., профессор

Ю.А. Брюханов

Доцент кафедры инфокоммуникаций
и радиофизики, к.т.н.

В.А. Волохов

Рис. В.2. Акт внедрения в учебный процесс ЯрГУ

ООО «А-Реал Консалтинг»

УТВЕРЖДАЮ

ИНН 7606047112,
КПП 760201001,
ОГРН 1037600807122

Директор ООО «А-Реал Консалтинг»

150044, г. Ярославль, Полушкина
роща, д.16, стр.67А, оф 3.

«13» октября 2017 г.

Сухарев И.Е.

АКТ

внедрения результатов диссертационной работы Носкова Андрея Александровича, выполненной в Ярославском государственном университете имени П.Г. Демидова (ЯрГУ), на тему «Формирование изображений с расширенной глубиной резкости для систем прикладного телевидения».

Результаты диссертационной работы Носкова А.А. нашли применение в работе ООО «А-Реал Консалтинг». Особый практический интерес представляют следующие результаты диссертации:

- рекомендации по выбору метрик сфокусированности с точки зрения точности. Метрика на основе отношения коэффициентов вейвлет-преобразования (ОКВП) имеет наибольший коэффициент корреляции с идеализированными характеристиками ($r = 0,88$), а также наименьший доверительный интервал;
- рекомендации по выбору метрик сфокусированности с точки зрения быстродействия. Метрика на основе вариации уровня серого имеет наименьшую вычислительную сложность (в 22 раза быстрее ОКВП) при этом сохраняя высокий уровень точности ($r = 0,76$).

Разработанные Носковым А.А. методики оценки сфокусированности областей изображений применяются в составе программного комплекса контентной фильтрации интернет трафика.

Начальник отдела разработки
ООО «А-Реал Консалтинг»


Красотин А.А.

Начальник отдела тестирования
ООО «А-Реал Консалтинг»


Комарова Д.А.

Рис. В.3. Акт внедрения в производственный процесс ООО "А-Реал Консалтинг"

ООО «Энергия-Инфо»

ОГРН: 1107603000130
КПП: 760301001
ИНН: 7603045036150008, г. Ярославль,
ул. Союзная, д.144

УТВЕРЖДАЮ

Директор ООО «Энергия-Инфо»

Попов П.Ю.

12 октября 2017



АКТ

внедрения результатов диссертационной работы Носкова Андрея Александровича, выполненной в Ярославском государственном университете имени П.Г. Демидова (ЯрГУ), на тему «Формирование изображений с расширенной глубиной резкости для систем прикладного телевидения».

Комиссия в составе: председатель комиссии – начальник отдела разработки Баракова А.Н., члены комиссии – инженер отдела разработки Балашов И.В., инженер отдела разработки Никитинский М.А., рассмотрев диссертационную работу Носкова А.А., составила настоящий акт о том, что ее результаты нашли применение в работе ООО «Энергия-Инфо». Особый практический интерес представляют следующие результаты работы:

- алгоритм формирования изображений с расширенной глубиной резкости на основе клеточного автомата;
- модифицированный алгоритм формирования изображений с расширенной глубиной резкости с использованием пирамид изображений, показывающий в среднем на 14% лучшие результаты по результатам экспертной оценки.

Разработанные Носковым А.А. алгоритмы применяются в составе программного комплекса технического зрения в качестве дополнительного компонента, повышающего эффективность распознавания образов.

Председатель комиссии

Баракова А.Н.

Члены комиссии

Балашов И.В.

Никитинский М.А.

Рис. В.4. Акт внедрения в производственный процесс ООО "Энергия-Инфо"