

На правах рукописи



НЕНАХОВ ИЛЬЯ СЕРГЕЕВИЧ

**НЕЭТАЛОННАЯ ОЦЕНКА КАЧЕСТВА ТЕЛЕВИЗИОННЫХ
ИЗОБРАЖЕНИЙ НА ОСНОВЕ ЛОКАЛЬНЫХ БИНАРНЫХ
ШАБЛОНОВ И АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ**

Специальность 05.12.04

Радиотехника, в том числе системы и устройства телевидения

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Владимир – 2016

Работа выполнена на кафедре динамики электронных систем ФГБОУ ВО «Ярославский государственный университет им. П.Г. Демидова».

Научный руководитель: **Брюханов Юрий Александрович**
доктор технических наук, профессор, заведующий кафедрой динамики электронных систем ФГБОУ ВО «Ярославский государственный университет им. П.Г. Демидова», г. Ярославль.

Официальные оппоненты: **Орлов Алексей Александрович**
доктор технических наук, доцент, заведующий кафедрой физики и прикладной математики Муромского института (филиала) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», г. Муром, Владимирская область.

Никитин Анатолий Евгеньевич
кандидат технических наук,
ведущий разработчик систем видеоаналитики
ООО «Точка зрения», г. Ярославль.

Ведущая организация: ОАО «Ярославский радиозавод», г. Ярославль.

Защита диссертации состоится 28 декабря 2016 г. в 14-00 часов на заседании диссертационного совета Д 212.025.04 при ФГБОУ ВО Владимирском государственном университете имени Александра Григорьевича и Николая Григорьевича Столетовых по адресу: 600000, г. Владимир, ул. Горького, д. 87, ВлГУ, корп. 3, ФРЭМТ, ауд. 301.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых и на сайте <http://diss.vlsu.ru/>.

Автореферат разослан «25» октября 2016 г.

Отзывы на автореферат, заверенные печатью, просим направлять по адресу: 600000, г. Владимир, ул. Горького, д. 87, ВлГУ, ФРЭМТ.

Ученый секретарь диссертационного совета доктор технических наук, профессор



А.Г. Самойлов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Цифровая обработка визуальной информации в настоящее время охватывает различные виды инфокоммуникационных 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).

Актуальность разработки алгоритмов обработки визуальной информации объясняется в том числе происходящим в настоящее время в России переходом к цифровому телевидению на основе стандартов DVB первого и второго поколений. Программа по переводу эфирного телевидения на цифровые технологии началась в 2009 году. С 2012 по 2016 годы федеральное государственное унитарное предприятие «Российская телевизионная и радиовещательная сеть» (РТРС) проводило строительство объектов телерадиовещания. Сегодня построено 4075 объектов первого мультиплекса и 2612 – второго. Первым мультиплексом охвачено 91,55% населения страны, вторым – 63,62%. Построены 83 центра формирования мультиплексов для организации регионального вещания. Развитие цифрового эфирного вещания в РФ сделает доступными 98,1% населения не менее 20 телеканалов в хорошем цифровом качестве. При этом треть населения страны сегодня подключена к кабельным ТВ-сетям, более четверти россиян – к спутниковому телевидению, которое позволяет принимать большее количество цифровых каналов, в том числе и в формате высокой четкости (HDTV).

Значительный вклад в данную область науки и техники внесли как отечественные ученые М.И. Кривошеев, В.А. Сойфер, 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.

Алгоритмы обработки, сжатия и кодирования визуальной информации неизбежно вносят искажения различного типа, что делает актуальным вопрос об оценке параметров этих искажений и объективного оценивания качества искаженного изображения. На современном этапе развития все методы оценки качества классифицируются на две большие категории:

1. Субъективные критерии оценки качества изображений. Такая оценка осуществляется человеком (оценки MOS по рекомендации ITU-R BT.500-11).

2. Объективные критерии оценки качества изображений. Здесь оценка производится алгоритмически. Такой алгоритм может использовать оригинал изображения для оценки качества, то есть быть эталонным, или может определять качество без наличия неискаженной копии изображения, то есть быть неэталонным.

В связи с необходимостью унификации исследований в данной области в последние годы предложены две общедоступные тестовые базы изображений с экспертными оценками MOS: LIVE (LIVE Image Quality Assessment Database, University of Texas at Austin, USA) и TID2008/2013 (Tampere Image Database, Tampere, Finland).

Задача неэталонной оценки качества изображений вызывает в последние годы повышенный интерес исследователей. Большая часть предлагаемых неэталонных алгоритмов разработана для конкретных типов априори известных искажений в изображении (блочность, размытие и др.), что не всегда удобно для использования в практических приложениях. Исследовательские лаборатории в области оценки качества в задачах цифровой обработки изображений и речевых сигналов созданы во многих крупных университетах мира. Среди них стоит выделить MIT, Stanford University, UCLA, University of Texas at Austin, Tampere University of Technology, ведущие университеты Японии и Сингапура. Практическими разработками занимаются как крупные глобальные компании: Intel, Panasonic, Samsung, так и мелкие технологичные компании-стартапы, созданные при ведущих университетах.

Обзор существующих неэталонных критериев оценки качества изображений и видеопоследовательностей показал, что универсальных решений в данной области на настоящий момент не существует. Перспективным подходом представляется решение данной проблемы на основе современных алгоритмов машинного обучения. Для этого следует, во-первых, определить признаки, отвечающие за качество изображений, во-вторых, определить эффективный классификатор, способный разделять изображения разного качества на классы.

Другой важной областью, где решение задачи неэталонной оценки качества изображений представляет большой практический интерес, является биометрическая идентификация человека по изображению его лица. Существует несколько стандартов, определяющих качество изображений лиц – ISO/IEC 19794-5, ICAO 9303, которые содержат описание характеристик, влияющих на принятие решения о пригодности изображения для систем автоматического распознавания.

Все стандартизированные характеристики можно разделить на два класса: текстурные (резкость, контрастность, освещенность, степень компрессии, наличие прочих искажений) и относящиеся непосредственно к признакам лица (симметричность, поза, поворот, видимость глаз, наличие бликов или теней на лице). Для их автоматического определения в научной литературе описаны следующие методы – определение позы, не фронтальности освещения и поворота.

Еще большую актуальность данному вопросу придает планируемая в РФ разработка национальной биометрической платформы и решение других перспективных задач на рынках Safenet Национальной Технологической Инициативы (НТИ).

Все вышеперечисленное доказывает, что разработка неэталонных критериев оценки качества изображений представляет собой актуальную научно-техническую задачу как для области телевидения, так и для ряда смежных научно-технических областей.

Целью работы является расширение возможностей контроля качества систем вещательного и прикладного телевидения.

Объектом исследования являются радиотехнические системы фильтрации и сжатия визуальной информации, биометрической идентификации и прикладного телевидения.

Предметом исследования являются неэталонные алгоритмы оценки качества телевизионных изображений на основе локальных бинарных шаблонов и методов машинного обучения.

Задачи диссертационной работы

- проведение исследований в части выбора признаков, характеризующих качество изображения в пространственной области;
- проведение исследований по выбору алгоритма машинного обучения для классификации изображений различного качества;
- разработка и анализ алгоритма неэталонной оценки качества на основе выбранных признаков и алгоритмов классификации;
- проведение исследований по оценке корреляции между разработанным неэталонным алгоритмом и результатами субъективной экспертизы оценки качества на специализированных базах изображений LIVE и TID2013;
- проведение анализа использования предложенного алгоритма неэталонной оценки качества изображений в ряде задач вещательного и прикладного

телевидения и в задачах биометрической идентификации/верификации личности.

Методы исследования. При решении поставленных задач использовались современные методы цифровой обработки изображений, машинного обучения, теории вероятностей и математической статистики. Для практической реализации алгоритмов применялись современные численные методы и язык программирования python (библиотека scikit-learn).

Научная новизна полученных результатов

В рамках работы получены следующие новые научные результаты:

- Предложено использование локальных бинарных шаблонов в качестве признаков в задаче неэталонной оценки качества изображений.
- Предложено использование рандомизированных деревьев в качестве классификатора изображений различного качества.
- Разработан алгоритм неэталонной оценки качества телевизионных изображений на основе локальных бинарных шаблонов и методов машинного обучения.
- Разработан модифицированный алгоритм неэталонной оценки качества с использованием статистики естественных изображений.

Практическая значимость полученных результатов

1. Неэталонный характер работы предложенных алгоритмов оценки качества телевизионных изображений позволяет использовать их в широком спектре практических приложений, где, как правило, исходное (эталонное) изображение не доступно. При этом достигаемое значение коэффициента корреляции Спирмена 0,93–0,96 (база LIVE), 0,6–0,92 (база TID2013) соответствует используемым на практике эталонным критериям.
2. Разработанный алгоритм является вычислительно эффективным, так как он работает в пространственной области и использует вычислительно не затратный аппарат локальных бинарных шаблонов.
3. Проведенный анализ использования предложенного алгоритма в ряде прикладных задач показывает, что он может быть использован в системах обработки и передачи визуальной информации, прикладного телевидения, биометрической идентификации личности по изображению лица и в других задачах цифровой обработки изображений и технического зрения.

4. На основании проведенного анализа времени работы модифицированного алгоритма можно сделать вывод о том, что он применим для задач неэталонной оценки качества изображений в офлайн-приложениях.

Результаты работы внедрены в соответствующие разработки ООО «А-ВиЖН» г. Ярославль и ООО «Технологии поиска» 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 (№ 15-07-08674-А и № 15-08-99639-А). Получено свидетельство о государственной регистрации программы для ЭВМ.

Достоверность полученных научных результатов обусловлена применением адекватного математического аппарата, подтверждается их согласованностью с результатами проведенного компьютерного моделирования и сопоставлением ряда полученных результатов с научно-техническими данными, полученными на унифицированных тестовых базах изображений LIVE и TID2013.

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

- 18-я международная конференция «Цифровая обработка сигналов и ее применение», DSPA (Москва, 2016);
- 70-я и 71-я международные конференции «Радиоэлектронные устройства и системы для инфокоммуникационных технологий», РЭУС (Москва, 2015–2016);
- 4-я всероссийская конференция «Радиоэлектронные средства получения, обработки и визуализации информации», РСПОВИ (Москва, 2015);
- 13-я международная научно-техническая конференция «Телевидение: передача и обработка изображений» (Санкт-Петербург, 2016);
- 17-th conference of FRUCT association (Yaroslavl, 2015);
- 14-th IEEE EAST-WEST DESIGN & TEST SYMPOSIUM (Yerevan, 2016).

Публикации. По теме диссертации опубликовано 13 научных работ, из них 3 статьи в журналах, рекомендованных ВАК, и 10 докладов на научных конференциях.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка использованных источников, содержащего 111 наименований, и

2 приложений. Она изложена на 132 страницах машинописного текста, содержит 30 рисунок и 30 таблиц.

Основные научные положения, выносимые на защиту

1. Алгоритм неэталонной оценки качества телевизионных изображений на основе локальных бинарных шаблонов и рандомизированных деревьев.
2. Модифицированный алгоритм неэталонной оценки качества телевизионных изображений с использованием статистики естественных изображений.
3. Результаты сопоставления предложенных объективных неэталонных оценок качества телевизионных изображений со средними экспертными оценками.
4. Результаты анализа работы неэталонного алгоритма оценки качества изображений в условиях искажений, характерных для задач биометрической идентификации/верификации личности по изображению лица.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность выбранной темы, сформулированы цель и задачи исследования, изложены основные положения, выносимые на защиту, показаны научная новизна и практическая значимость работы.

В первой главе рассмотрены современные задачи, методы и алгоритмы в области оценки качества цифровых изображений. Представлен обзор стандартных баз данных для тестирования точности таких алгоритмов – LIVE, TID2013. Рассмотрены современные алгоритмы машинного обучения на основе бустинга, машины опорных векторов, рандомизированных деревьев и сверточных нейронных сетей.

Во второй главе рассматриваются неэталонные алгоритмы оценки качества цифровых изображений. Особое внимание уделяется применению локальных бинарных шаблонов в задаче оценки качества.

Локальный бинарный шаблон (ЛБШ) – оператор, используемый для классификации текстур в техническом зрении. Имеет сложность вычисления $O(n)$, где n – число пикселей в изображении. Он представляет собой бинарный код, описывающий окрестность заданного пикселя. Если значение яркости пикселя из окрестности больше значения яркости рассматриваемого пикселя, то в соответствующее место бинарного кода ставится 1, иначе 0. Для вычисления локального бинарного шаблона необходимо задать значение двух параметров – P и r . Первый параметр отвечает за то, сколько точек из окрестности будет участвовать

в расчетах, а второй – радиус, определяющий размер окрестности непосредственно (рис. 1).

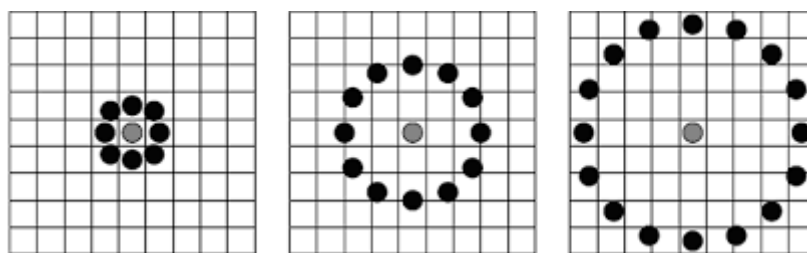


Рис. 1. Выбор точек из окрестности оператором ЛБШ с параметрами P, r : (8, 1), (12, 2), (16, 4)

Локальные бинарные шаблоны определенного вида несут в себе больше информации, чем остальные. Так ЛБШ называется равномерным, если он содержит не более двух переходов между «0» и «1» (например, 00000011, 00111000, 11111111). Равномерные локальные бинарные шаблоны отражают важные структурные особенности изображения, такие как концы линий, углы, пятна.

Поворот изображения приводит к повороту пикселей относительно центрального пикселя внутри окрестности, что ведет к изменению бинарных кодов шаблонов. Для того чтобы набор ЛБШ признаков для изображения не изменялся с поворотом, вводится специальная процедура присваивания бинарных кодов соответствующим ЛБШ шаблонам – каждый бинарный код приводится к своему минимальному значению путем битового циклического сдвига.

В работе показано, что ЛБШ при $r=1$ чувствительны к изменению направления освещения и угла обзора. Более устойчивые признаки можно получить, используя многомасштабный оператор ЛБШ. Под многомасштабностью здесь понимается одновременное вычисление ЛБШ для заданного пикселя в окрестностях разного радиуса.

Для описания изображения в пространстве признаков ЛБШ используют гистограмму, столбцы которой соответствуют десятичным значениям бинарных кодов шаблонов. На рис. 2 и рис. 3 представлено соответствие гистограммы равномерных инвариантных к повороту ЛБШ для изображений из базы LIVE, подвергнутых искажениям различных типов.



а)



б)

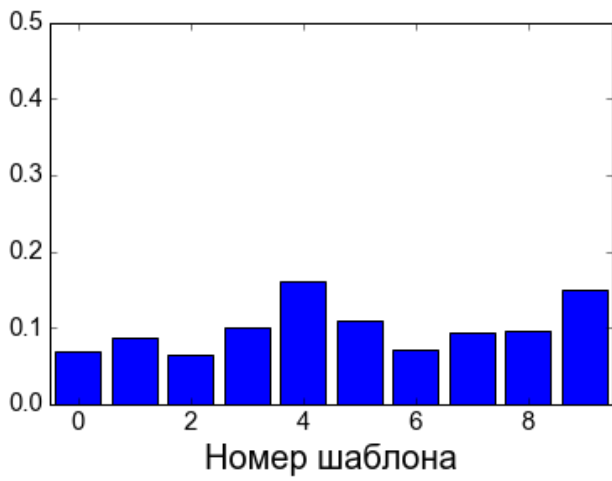


в)

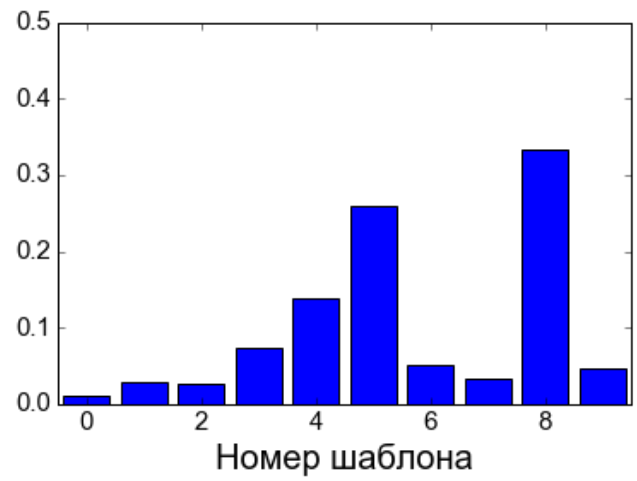


г)

Рис. 2. Изображения с различными типами искажений:
а) оригинал; б) сжатие jpeg; в) сжатие jpeg2000; г) размытие



а)



б)

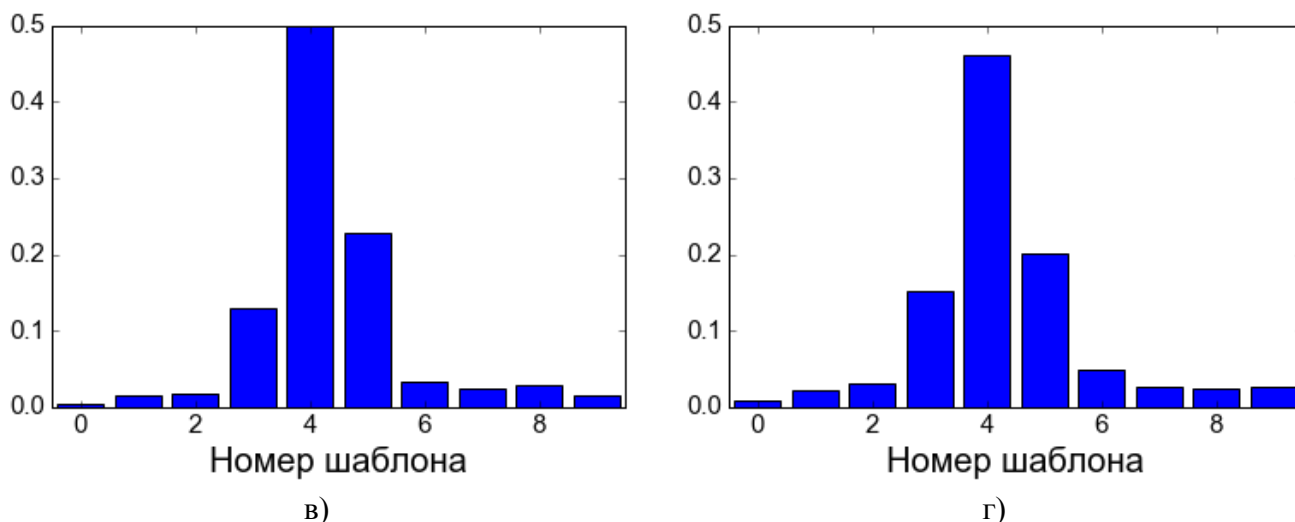


Рис. 3. Гистограмма равномерных инвариантных к сдвигу ЛБШ для искаженных изображений: а) оригинал; б) сжатие jpeg; в) сжатие jpeg2000; г) размытие

Предложен новый алгоритм неэталонной оценки качества с использованием локальных бинарных шаблонов и алгоритмов машинного обучения (НОК ЛБШ-1). Схема его работы показана на рис. 4.

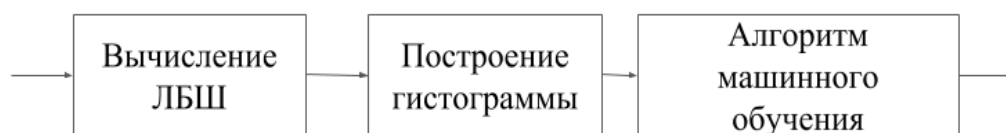


Рис. 4. Алгоритм неэталонной оценки качества на основе локальных бинарных шаблонов и методов машинного обучения

В качестве вектора признаков для предложенного алгоритма выступают столбцы гистограмм многомасштабных равномерных локальных бинарных шаблонов, инвариантных к сдвигу, с $r=1,2,3$ и значениями $P=8,16,24$. Гистограммы для каждого масштаба строятся отдельно, после чего формируется результирующая гистограмма путем склейки гистограмм всех масштабов. Всего для каждого изображения рассчитывается 54 признака. Выбор многомасштабных локальных бинарных шаблонов обусловлен их повышенной точностью по сравнению с одномасштабными шаблонами.

В третьей главе приведены результаты исследований для алгоритма НОК ЛБШ-1 с различными классификаторами, такими как бустинг (AdaBoost), машина опорных векторов, градиентный бустинг над решающими деревьями и рандомизированные деревья (табл. 1).

Полученные результаты показывают, что для алгоритма НОК ЛБШ-1 по всем трем метрикам наилучшие результаты показывает классификатор на основе рандомизированных деревьев.

Таблица 1. Коэффициенты корреляции Пирсона (верхнее значение) и Спирмена (нижнее значение) при использовании многомасштабных ЛБШ для различных классификаторов на базе изображений LIVE

Тип искажения \ Классификатор	JPEG	JPEG 2000	Белый шум	Размытие	Битовые ошибки	Все
Бустинг	0,923	0,943	0,950	0,945	0,886	0,924
	0,901	0,932	0,967	0,942	0,926	0,926
Машина опорных векторов	0,809	0,884	-0,484	0,933	0,797	0,618
	0,854	0,883	-0,364	0,931	0,796	0,567
Градиентный бустинг	0,962	0,962	0,981	0,967	0,931	0,959
	0,937	0,945	0,964	0,945	0,904	0,947
Рандомизированные деревья	0,970	0,968	0,987	0,971	0,936	0,965
	0,952	0,957	0,974	0,955	0,912	0,956

Для улучшения результатов также разработан модифицированный алгоритм неэталонной оценки качества, основанный на комбинации признаков из НОК ЛБШ-1 и алгоритма неэталонной оценки качества в пространственной области (НОКПО), в дальнейшем такой алгоритм будем называть НОК ЛБШ-2. Алгоритм НОКПО работает с усредненными нормализованными коэффициентами контраста (УНКК-изображение):

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C} \quad (\text{УНКК-изображение}),$$

$$\mu(i, j) = \sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} I_{k,l}(i, j) \quad (\mu\text{-изображение}),$$

$$\sigma(i, j) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} (I_{k,l}(i, j) - \mu(i, j))^2} \quad (\sigma\text{-изображение}),$$

где $i \in 1, 2, \dots, M$, $j \in 1, 2, \dots, N$ – пространственные индексы, M, N – высота и ширина изображения, $C = 1$ – константа, необходимая для исключения случаев деления на нуль, а $\{w_{k,l} | k = -K, \dots, K; l = -L, \dots, L\}$ – двумерная циклично-симметричная взвешенная функция Гаусса. Параметры K, L – небольшие целые числа. На рис. 5 представлены визуальные примеры оригинального изображения из базы LIVE и соответствующих ему УНКК-изображения и σ -изображения.

Экспериментально было установлено, что внесение различных типов искажений в изображение влияет на распределение УНКК, причем между типом искажения и распределением УНКК существуют определенные закономерности. Например, белый шум уменьшает значения «хвостов» распределения, а размытие искажает распределение так, что оно становится похожим по форме на распределение Лапласа.

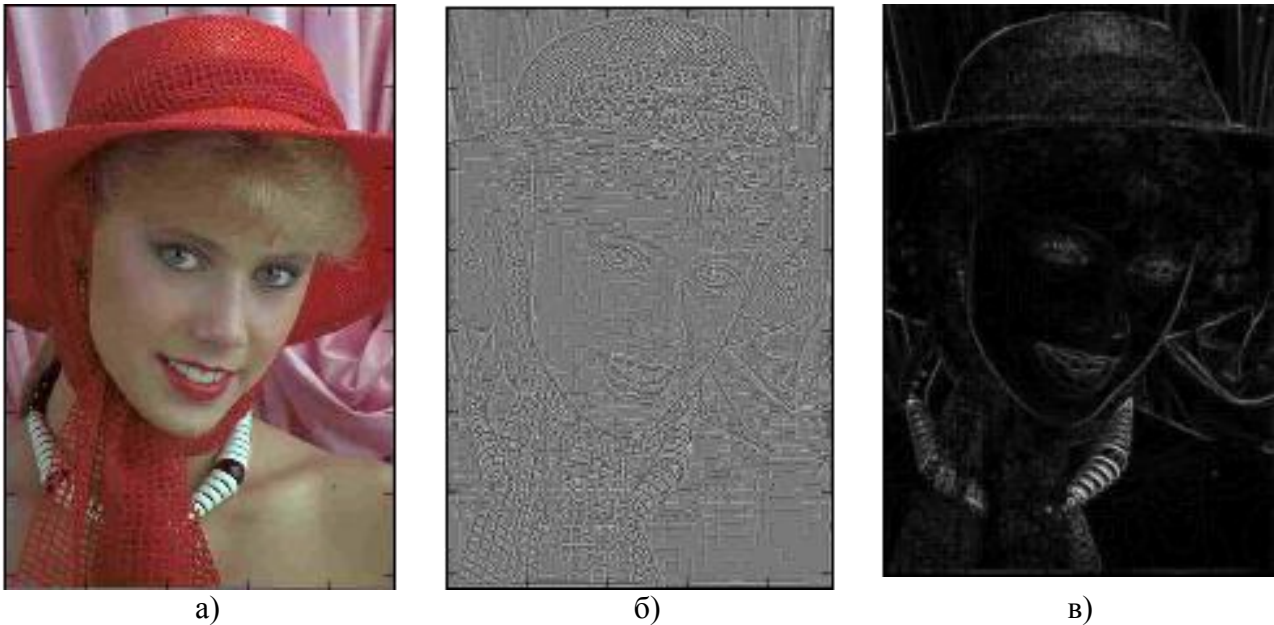


Рис. 5. Изображения, используемые алгоритмом НОКПО: а) оригинальное; б) нормализованные коэффициенты контраста (УНКК-изображение); в) σ -изображение

Множество всевозможных распределений искаженных изображений удобно записать в виде обобщенного распределения Гаусса:

$$f(x, \alpha, \sigma^2) = \frac{\alpha}{\beta 2\Gamma(1/\alpha)} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right),$$

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}},$$

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt \quad \alpha > 0,$$

где α – параметр формы, σ^2 – дисперсия.

Во время работы алгоритма НОК ЛБШ-2 для каждого изображения оценивается два параметра (α, σ^2) к которым добавляются признаки НОК ЛБШ-1, рассчитанные по 3 изображениям – оригинальному, УНКК-изображению и σ -изображению. В табл. 2 представлены результаты сравнения различных вариантов комбинированных алгоритмов с алгоритмом НОК ЛБШ-1.

Приведенные результаты показывают, что алгоритм, вычисляющий признаки для трех изображений и первый набор признаков НОКПО, обладает повышенной точностью по сравнению с алгоритмом НОК ЛБШ-1.

Проведен эксперимент по измерению времени работы алгоритмов НОК ЛБШ-1, НОК ЛБШ-2 на изображениях различного разрешения (рис. 6).

Таблица 2. Корень из среднеквадратической ошибки при использовании комбинированных алгоритмов на тестовой базе LIVE

Изображения, по которым рассчитываются признаки	Тип признаков	JPEG	JPEG 2000	Белый шум	Размытие	Битовые ошибки	Все
УНКК-изобр.+ σ -изобр.+ориг.	ЛБШ+ α + σ^2	5,74	5,59	2,91	4,26	6,54	5,31
УНКК-изобр.+ σ -изобр.+ориг.	ЛБШ	5,79	5,6	3,42	4,31	6,76	5,44
ориг.	ЛБШ	5,95	6,22	3,40	5,63	7,56	6,02

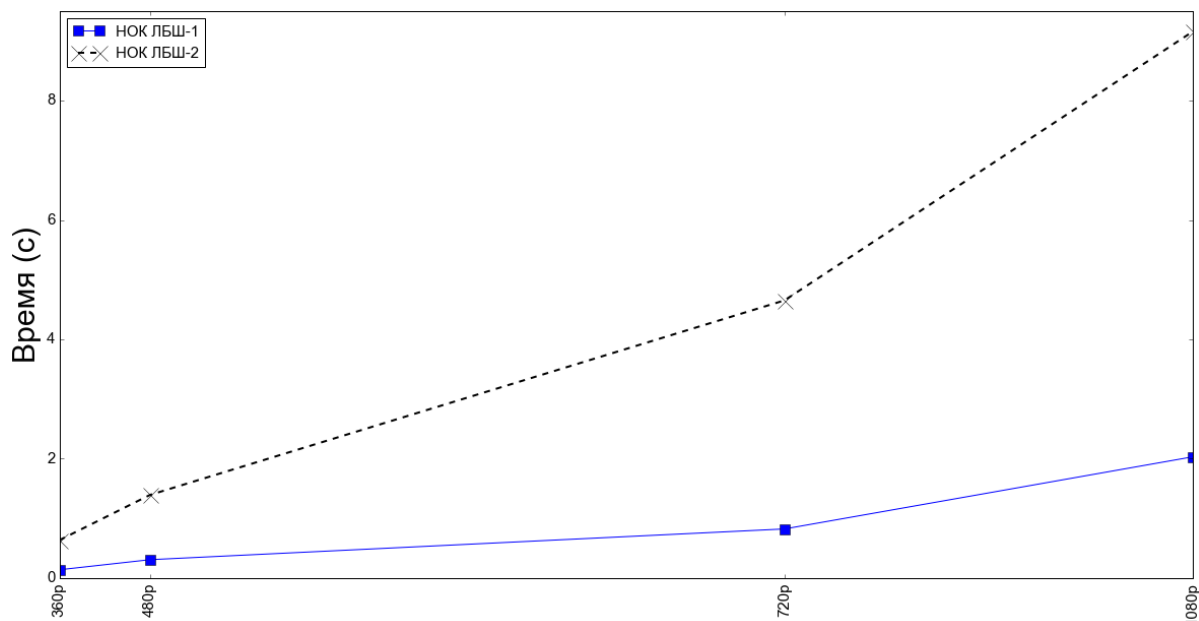


Рис. 6. Время работы алгоритмов НОК ЛБШ-1, НОК ЛБШ-2 для изображений различного разрешения

Приведенные результаты показывают, что время работы алгоритма НОК ЛБШ-2 при росте разрешения изображения увеличивается существенно, поэтому использование его целесообразно только в тех приложениях, когда требуется оценка качества изображения в режиме офлайн.

Получены результаты по тестированию точности алгоритмов НОК ЛБШ-1 и НОК ЛБШ-2 в сравнении с известными эталонными и неэталонными метриками на стандартных базах изображений LIVE и TID2013 (табл. 3 и табл. 4).

Таблица 3. Коэффициенты корреляции Пирсона (верхнее значение) и Спирмена (нижнее значение) алгоритмов оценки качества изображений на тестовой базе LIVE

Название алгоритма	JPEG	JPEG 2000	Белый шум	Размытие	Битовые ошибки	Все типы
ПОСШ (PSNR)	0,903	0,876	0,917	0,780	0,880	0,859
	0,883	0,865	0,941	0,752	0,874	0,864
КСП (SSIM)	0,946	0,941	0,982	0,900	0,951	0,907
	0,947	0,939	0,964	0,905	0,939	0,913

МКСП (MSSIM)	0,979	0,975	0,988	0,918	0,949	0,954
	0,979	0,963	0,977	0,914	0,939	0,951
LBIQ	0,935	0,910	0,976	0,910	0,838	0,909
	0,929	0,904	0,970	0,898	0,822	0,906
LD-TS	0,851	0,827	0,954	0,932	0,878	0,877
	0,833	0,820	0,957	0,925	0,886	0,883
GRNN	0,880	0,828	0,989	0,825	0,819	0,837
	0,872	0,816	0,979	0,833	0,735	0,827
НОКПО	0,973	0,923	0,985	0,951	0,903	0,942
	0,965	0,914	0,979	0,951	0,877	0,940
НОК ЛБШ-1	0,971	0,969	0,988	0,972	0,946	0,966
	0,956	0,960	0,977	0,960	0,924	0,957
НОК ЛБШ-2	0,973	0,974	0,992	0,981	0,960	0,974
	0,957	0,966	0,985	0,970	0,937	0,967

Таблица 4. Сравнительные результаты работы алгоритмов НОК ЛБШ-1 и НОК ЛБШ-2 на тестовой базе TID2013

Алгоритм	Коэффициент корреляции Спирмена	Коэффициент корреляции Кендалла
НОК ЛБШ-2	0,859	0,670
FSIM _x	0,851	0,667
PSNR-NA	0,819	0,643
PSNR-HMA	0,813	0,632
FSIM	0,801	0,630
НОК ЛБШ-1	0,790	0,600
МКСП (MSSIM)	0,787	0,608
PSNR _c	0,687	0,508
VSNR	0,681	0,508
PSNR-HVS	0,654	0,496
ПОСШ (PSNR)	0,640	0,482
КСИ (SSIM)	0,637	0,470

Результаты тестирования на базе LIVE показывают, что предложенный алгоритм НОК ЛБШ-2 показывает наилучшие результаты для следующих типов искажений: белый шум, размытие, а также в среднем по всей базе LIVE.

Результаты тестирования на базе TID2013 показывают, что алгоритм НОК ЛБШ-2 опережает по точности известные эталонные и неэталонные метрики. Алгоритм НОК ЛБШ-1 находится на одном уровне с известным алгоритмом МКСП (MSSIM).

В четвертой главе проведен анализ практических применений разработанных неэталонных алгоритмов оценки качества для телевизионных изображений.

Проведено неэталонное оценивание качества видеоданных по метрикам НОК ЛБШ-1 и НОКПО относительно субъективных экспертных оценок. Результаты получены на тестовой базе ReTRiEVED, содержащей экспертные оценки.

Усредненные по 8 тестовым видеопоследовательностям результаты приведены на рис. 7. Их анализ показывает, что неэталонные метрики НОКПО и НОК ЛБШ-1 коррелируют с визуально воспринимаемым качеством изображения (характеризуемым средними экспертными оценками MOS). При падении оценок MOS с 3 до 1 – уровень неэталонных метрик возрастает на величину порядка 20%.

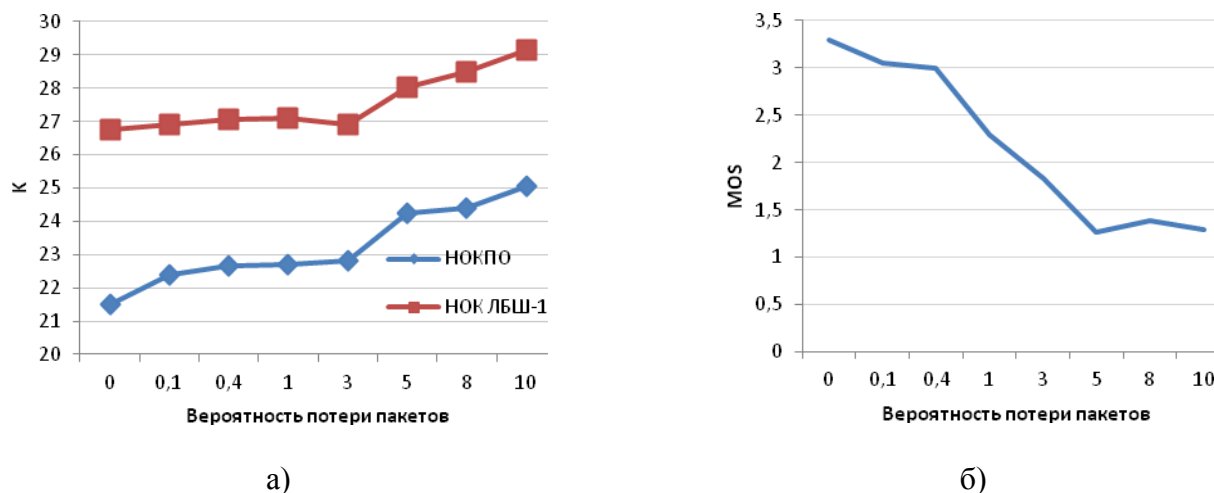


Рис. 7. Зависимость метрик качества от вероятности потери видео в канале: а) НОКПО, НОК ЛБШ-1; б) экспертные оценки MOS

Проведена оценка качества изображений лиц в задаче идентификации при использовании алгоритма распознавания на основе сверточной нейронной сети. Цель проведенного эксперимента – анализ точности системы распознавания при внесении во входные изображения различных типов искажений: размытие по Гауссу, белый шум, сжатие JPEG. Требовалось оценить корреляцию между точностью распознавания и значениями метрики НОК ЛБШ-1 для искаженных изображений.

Анализируя полученные результаты, можно сделать следующие выводы:

- искажения типа размытие по Гауссу и сжатие JPEG оказывают незначительное влияние на уровень распознавания. Искажения типа белый шум, импульсный шум, сжатие JPEG2000 существенно влияют на уровень распознавания;
- значения метрики НОК ЛБШ-1 ниже 50 соответствуют при всех типах искажений рабочему диапазону для работы алгоритмов распознавания на сверточной нейронной сети.

Рассмотрена задача выбора наилучшего лица для задач биометрической идентификации на основе алгоритмов неэталонной оценки качества.

Типичная схема системы распознавания с модулем оценки качества изображений лиц приведена на рис. 8. Качество лиц здесь оценивается на этапе

предобработки. Изображения низкого качества отбрасываются или архивируются, распознавание происходит только по высококачественным изображениям. В научно-технической литературе показано, что использование системы распознавания с модулем оценки качества лиц приводит к значительному приросту производительности всей системы видеонаблюдения.



Рис. 8. Система биометрической идентификации/верификации лиц с модулем оценки качества

Проведено исследование ряда метрик оценки качества изображения лица применительно к задаче выбора наилучшего изображения для биометрической идентификации личности. Результаты моделирования показывают, что при низком уровне освещенности лучшую корреляцию с субъективными экспертными оценками показывает метрика размытия, что вызвано низким средним качеством соответствующих видеопоследовательностей. При нормальном и высоком уровне освещенности лучшую корреляцию показывает метрика на основе симметричности особых точек. В эксперименте с выбором трех лучших изображений наилучший результат показывает метрика на основе обучения ранжированию. Показано, что

использование алгоритма НОК ЛБШ-1 целесообразно при уровне освещения лица выше 100 лк. При низком уровне освещения целесообразно использование специализированных метрик оценки качества.

Полученные в 4-й главе результаты практического применения неэталонного алгоритма НОК ЛБШ-1 будут полезны инженерам в области построения систем прикладного телевидения и биометрической идентификации личности по изображению лица.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

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

1. На основе анализа современных научно-технических источников выбран метод для разработки неэталонного алгоритма оценки качества изображений. В качестве его использован алгоритм на основе выделения признаков и алгоритмов машинного обучения. Обоснован выбор локальных бинарных шаблонов в качестве признаков в задаче неэталонной оценки качества изображений.
2. Разработан и протестирован алгоритм неэталонной оценки качества изображений на основе локальных бинарных шаблонов и классификации на основе методов машинного обучения (НОК ЛБШ-1).
3. Установлено, что использование многомасштабных локальных бинарных шаблонов позволяет увеличить коэффициент корреляции Пирсона на 0,006, коэффициент корреляции Спирмена на 0,009 и метрику корня из среднеквадратической ошибки на 0,51 при тестировании на базе LIVE.
4. Установлено, что использование рандомизированных деревьев оказывается наилучшим выбором из всех тестируемых алгоритмов машинного обучения. Их использование совместно с многомасштабными локальными бинарными шаблонами позволяет достичь следующих показателей для алгоритма НОК ЛБШ-1 на базе LIVE: коэффициент корреляции Пирсона – 0,966, коэффициент корреляции Спирмена – 0,957, корень из среднеквадратической ошибки – 6,02.
5. Разработан и протестирован модифицированный алгоритм неэталонной оценки качества (НОК ЛБШ-2). Он использует смешанные признаки и несколько изображений для их расчета. Его использование позволяет достичь следующих показателей для алгоритма НОК ЛБШ-2 на базе LIVE:

коэффициент корреляции Пирсона – 0,974, коэффициент корреляции Спирмена – 0,967, корень из среднеквадратической ошибки – 5,31.

6. Тестирование времени работы алгоритмов неэталонной оценки качества показывает, что вычислительная сложность алгоритма НОК ЛБШ-2 при росте разрешения изображения увеличивается существенно, поэтому использование его целесообразно только в тех приложениях, когда требуется оценка качества изображения в режиме офлайн. Для приложений реального времени рекомендуется использовать алгоритм НОК ЛБШ-1.
7. Итоговое тестирование известных эталонных и неэталонных алгоритмов на базе LIVE показывает, что предложенный алгоритм НОК ЛБШ-2 показывает абсолютно лучшие результаты для следующих типов искажений: белый шум, размытие, а также в среднем по всей базе LIVE.
8. Итоговое тестирование эталонных и неэталонных алгоритмов на базе TID2013 показывает, что предложенный алгоритм НОК ЛБШ-2 показывает абсолютно лучший результат, достигая следующих значений: коэффициент корреляции Спирмена – 0,859, коэффициент корреляции Кендалла – 0,670.
9. Анализ результатов оценивания качества видеопоследовательностей показывает, что неэталонный алгоритм НОК ЛБШ-1 хорошо коррелирует с визуально воспринимаемым качеством изображения (характеризуемым средними экспертными оценками MOS). При падении оценок MOS с 3 до 1 – уровень неэталонных метрик возрастает на величину порядка 20%.
10. Анализ применения алгоритма НОК ЛБШ-1 в задаче распознавания показал, что искажения типа размытие по Гауссу и сжатие JPEG оказывают незначительное влияние на уровень распознавания. Искажения типа белый шум, импульсный шум, сжатие JPEG2000 существенно влияют на уровень распознавания. Значение метрики НОК ЛБШ-1 ниже 50 соответствуют при всех 5 типах искажений рабочему диапазону для работы алгоритмов распознавания на сверточной нейронной сети.
11. В задаче выбора наилучшего лица для биометрической идентификации показано, что использование алгоритма НОК ЛБШ-1 целесообразно при уровне освещения лица свыше 100 лк. При низком уровне освещения целесообразно использование специализированных метрик оценки качества.
12. Цель и задачи диссертационной работы успешно выполнены. Разработанные алгоритмы неэталонной оценки качества изображений интегрированы в программу SkifVideo.QMonitoring – приложение для неэталонной оценки

качества видео. Получено свидетельство о государственной регистрации программы для ЭВМ.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах из перечня ВАК

1. Приоров А.Л., Ненахов И.С., Хрящев В.В. Разработка и анализ алгоритма неэталонной оценки качества изображений на основе локальных бинарных шаблонов // Успехи современной радиоэлектроники. 2016. № 7. С. 46–52.
2. Тараканов А.Н., Балашова О.Н., Ненахов И.С. Сравнительный анализ алгоритмов многодескрипторного видеокодирования с пространственным и временным разделением на субпотoki // Радиотехнические и телекоммуникационные системы. 2016. № 1. С. 76–86.
3. Хрящев В.В., Матвеев Д.В., Лебедев А.А., Ненахов И.С. Детектирование человека при анализе видеопоследовательностей с купольных камер // Успехи современной радиоэлектроники. 2016. № 8. С. 47–55.

Материалы российских и международных конференций

4. Гущина О.Н., Ненахов И.С., Тараканов А.Н. Анализ качества изображения в системе многодескрипторного кодирования видео с пространственным мультиплексированием // Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2015: Сб. докл. междунар. конф. – М., 2015. С. 317–321.
5. Приоров А.Л., Хрящев В.В., Ненахов И.С. Исследование метода неэталонной оценки качества изображений в пространственной области // Радиоэлектронные средства получения, обработки и визуализации информации (РСПОВИ-2015): Сб. докл. 3-й всерос. конф. – М., 2015. С. 147–151.
6. Ненахов И.С., Шемяков А.М., Хрящев В.В. Использование неэталонного алгоритма оценки качества изображений в задачах нелинейной фильтрации // Цифровая обработка сигналов и ее применение (DSPА-2016): Докл. 18-й междунар. конф. – М., 2016. Т. 2. С. 743–748.
7. Ненахов И.С., Хрящев В.В., Павлов В.А. Оценка качества изображения лица в задачах биометрической идентификации // Цифровая обработка сигналов и ее применение (DSPА-2016): Докл. 18-й междунар. конф. – М., 2016. Т. 2. С. 954–959.

8. Ненахов И.С. Разработка и анализ алгоритмов оценки качества изображений для задач распознавания // Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2016: Сб. докл. междунар. конф. – М., 2016. Т. 2. С. 445–450.
9. Ненахов И.С., Брюханов Ю.А., Хрящев В.В. Неэталонная оценка качества изображений на основе многомасштабных локальных бинарных шаблонов // Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2016: Сб. докл. междунар. конф. – М., 2016. Т. 2. С. 378-382.
10. Шемяков А.М., Хрящев В.В., Ненахов И.С., Лебедев А.А. Распознавание лиц на изображениях при потере фронтальности // Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2016: Сб. докл. междунар. конф. – М., 2016. Т. 2. С. 525–529.
11. Хрящев В.В., Ненахов И.С., Приоров А.Л. Использование многомасштабных локальных бинарных шаблонов в задаче оценки качества изображений // Телевидение: передача и обработка изображений: Тез. 13-й междунар. конф. – СПб., 2016. С. 119–123.
12. Gushina O., Nenakhov I., Tarakanov A. Influence of Packets Losses on Video Quality in Case of Using Multiple Description Coding with Using Multiple Description Coding with Time Division into Two and Three Substreams // Proceedings of the 17th conference of FRUCT association. 2015. P. 230–236.
13. Nenakhov I., Khryashchev V., Priorov A. No-Reference Image Quality Assessment based on Local Binary Patterns // Proceedings of the 14th IEEE EAST-WEST DESIGN & TEST SYMPOSIUM. 2016. P. 529–532.

Свидетельство о государственной регистрации программы для ЭВМ

14. Мочалов И.С., Ненахов И.С., Приоров А.Л. SkifVideo.QMonitoring – приложение для неэталонной оценки качества видео // Свидетельство о государственной регистрации программы для ЭВМ № 2016618772 от 17 июня 2016.

Ненахов Илья Сергеевич

**НЕЭТАЛОННАЯ ОЦЕНКА КАЧЕСТВА ТЕЛЕВИЗИОННЫХ
ИЗОБРАЖЕНИЙ НА ОСНОВЕ ЛОКАЛЬНЫХ БИНАРНЫХ
ШАБЛОНОВ И АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ**

Автореферат диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 21.10.2016.

Формат 60x84/16. Усл. печ. л. 1. Тираж 100 экз.

ИП Платонова И.В. 150000, Ярославль, ул. Б. Октябрьская 37/1.