

<u>ИПМ им.М.В.Келдыша РАН</u> • <u>Электронная библиотека</u> <u>Препринты ИПМ</u> • <u>Препринт № 83 за 2016 г.</u>



ISSN 2071-2898 (Print) ISSN 2071-2901 (Online)

Зипа К.С., Игнатенко А.В., <u>Волобой А.Г.</u>

Алгоритм тональной компрессии видеопоследовательностей

Рекомендуемая форма библиографической ссылки: Зипа К.С., Игнатенко А.В., Волобой А.Г. Алгоритм тональной компрессии видеопоследовательностей // Препринты ИПМ им. М.В.Келдыша. 2016. № 83. 32 с. doi:<u>10.20948/prepr-2016-83</u> URL: <u>http://library.keldysh.ru/preprint.asp?id=2016-83</u> Ордена Ленина ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ имени М.В.Келдыша Российской академии наук

К.С. Зипа, А.В. Игнатенко, А.Г. Волобой

Алгоритм тональной компрессии видеопоследовательностей

Зипа К.С., Игнатенко А.В., Волобой А.Г.

Алгоритм тональной компрессии видеопоследовательностей

рассматривается решение B статье задачи тональной компрессии изображений широкого динамического диапазона для случая видеоряда. Производится рассмотрение существующих алгоритмов, решающих поставленную задачу, а также методов оценки качества алгоритмов тональной компрессии. Производится сравнение алгоритмов тональной компрессии для фотографий и выбирается лучший из рассмотренных алгоритмов. Этот алгоритм модифицируется для применения к видео-данным. Также в статье предлагается объективная метрика мерцания кадра, позволяющая оценивать степень временных искажений различных версий модифицированного алгоритма. На основании значений метрики был произведен ряд модификаций, позволивший уменьшить временные колебания яркости в результирующих данных.

Ключевые слова: тональная компрессия, восприятие человека.

Zipa K.S., Ignatenko A.V., Voloboy A.G.

Tone mapping for video data

This paper presents a new solution to the tone mapping problem for HDR video sequences. Brief overview of existing algorithms as well as methods for tone mapping quality assessment is presented. We compare tone mapping algorithms for photos and select the best one as a baseline. This algorithm is modified for application to video. Also, we propose an objective flickering metric that allows to evaluate the degree of temporal distortion of different algorithm modifications. On the basis of the metric values a number of modifications are proposed, which allows to reduce the flickering.

Key words: tone mapping, human perception.

Оглавление

Введение	3
1 Обзор существующих алгоритмов тональной компрессии видео	4
2 Субъективное тестирование алгоритмов тональной компрессии фотографий	i
в системе GML ePublish	8
3 Алгоритм тональной компрессии видеопоследовательностей	15
4 Корректировка глобальных параметров в алгоритме тональной компрессии	16
5 Локальная корректировка для подавления эффекта мерцания	21
6 Метрика мерцания для видео последовательностей	24
7 Тестирование алгоритма	27
Заключение	29
Библиографический список	29

Введение

Динамический диапазон яркости окружающего мира очень широк, в то же время диапазоны современных устройств вывода ограничены всего двумятремя порядками. Это ограничение привело к практике получения изображений с диапазоном яркости в 256 градаций, что может приводить к потере информации в некоторых областях изображения.

При наблюдении сцены в естественных условиях наблюдатель бы смог различать детали в ярких и тёмных областях в более широких пределах, чем по фотографии узкого диапазона. Это связано с тем, что человеческий глаз обладает более широким диапазоном чувствительности, чем тот, который позволяют достичь стандартные форматы хранения и передачи изображений.

С помощью обычной бытовой камеры [1] можно получить изображение широкого динамического диапазона (HDRI – high dynamic range image), то есть такое изображение, количество градаций яркости в котором превышает 256 значений. Напротив, изображение с меньшим количеством градаций называют изображением узкого динамического диапазона (LDRI – low dynamic range image).

Чтобы отобразить HDRI на экране монитора, нужно осуществить сжатие диапазона яркости. Такое преобразование называют алгоритмом тональной компрессии. Качество алгоритма тональной компрессии определяется степенью похожести результата на результат восприятия наблюдателем сцены. Поэтому алгоритмы тональной компрессии связаны с изучением восприятия человека.

Алгоритмы тональной компрессии (TM – tone mapping) развиваются со временем. В последнее время наблюдается смещение фокуса исследований в направлении задачи тональной компрессии для видеопоследовательностей.

Решение этой задачи востребовано в кинематографе, так как ведется постепенное совершенствование записывающей аппаратуры, позволяющей получать HDR-изображения в реальном времени [2]. При создании алгоритма для работы с видео исследователи, как правило, выбирают путь адаптации существующего алгоритма к новому типу данных. Для этого его иногда модифицируют, чтобы обеспечивалась устойчивость к малым изменениям входных данных, а результаты алгоритма были монотонны по времени и не вызывали эффекта мерцания (когда меняется яркость последовательных кадров, см. рис. 1).



Рис. 1. Пример последовательных кадров с различиями по яркости.

В этой работе проведено сравнение нескольких существующих алгоритмов тональной компрессии для статических изображений, выбран наилучший из них. Выбранный алгоритм модифицирован, применен к базе HDRV (High Dynamic Range Video) Stuttgart [3], и произведена оценка качества произведенных изменений.

1 Обзор существующих алгоритмов тональной компрессии видео

Исследования и разработка алгоритмов тональной компрессии ведутся уже два десятилетия. Разработчики операторов тональной компрессии ставили для цели: ОДНИ алгоритмы создавались ДЛЯ максимально них различные передачи приблизиться реалистичной яркости, стремились другие к художественному представлению реальных яркостей на устройствах С ограниченным диапазоном видео.

В последнее время стали появляться новые возможности для получения HDRV[4,5,6]. В связи с этим появилась возможность расширить область применения существующих алгоритмов для изображений, нацелив их на применение к видео.

Несмотря на обилие существующих операторов, только малая их часть подходит для обработки видеопоследовательностей. Это связано с тем, что необходимо отдельно озаботиться устойчивостью алгоритма ко времени. Подробный обзор применения алгоритмов тональной компрессии к видеоданным был произведен в работе [7]. Авторы проанализировали качество 11 оптимизации параметров операторов. После были исследованы И отсортированы по значимости обнаруженные артефакты (мерцание, временное несоответствие цветов, эффект «призраков») и проблемы цветопередачи насыщенности). После этого в эксперименте по (яркости, контраста и попарному сравнению были отсортированы наилучшие современные операторы.

Производя обзор алгоритмов, следует отдельно обратить внимание на методы и метрики для сравнения существующих алгоритмов. Этот вопрос очень важен, так как в настоящее время существует несколько различных подходов к проверке качества алгоритмов.

Согласно [7] всего можно выделить 4 больших группы методов проверки качества алгоритмов тональной компрессии:

- 1. соответствие реальности,
- 2. соответствие HDR-представлению,
- 3. метод сопоставления внешности,
- 4. оценка без эталона.

Каждый из методов имеет свои преимущества и недостатки. Наиболее часто используется последний вид тестирования, который применяется также и в этой работе.

В методе соответствия реальности [8,9] результат алгоритма сравнивается с физической сценой. Для этого необходимо одновременно предъявить наблюдателю результат алгоритма и саму сцену. Эта задача сложна для наблюдателей, поскольку отображаемый результат алгоритма не может быть в принципе тождественен сцене (например, максимальная яркость экрана никогда не равна яркости солнца, есть отличия по фокусу, полю зрения и т.д.). Эти факторы невозможно полностью устранить, однако с помощью такого сравнения можно напрямую проверить качество алгоритма.

В методе соответствия HDR-представлению [10,11] результат сравнивается с эталонным изображением на HDR-дисплее. В этом случае подразумевается, что на дисплее отображена наилучшая возможная визуализация. Хотя потенциально у таких дисплеев диапазон широкий, но даже для представления на нём уже нужно произвести некоторые преобразования входных данных (например, растяжение и обрезание яркости), а значит всё-таки эталонное изображение не даёт того же представления, что реальная сцена. Преимущество же этого подхода над предыдущим в том, что эксперименты могут осуществляться в хорошо контролируемом окружении.

Суть метода сопоставления внешности [12] состоит в сопоставлении атрибутов оригинальной сцены и репродукции, например, можно измерить яркость квадратной области в сцене и на дисплее. Наилучший алгоритм тональной компрессии будет обеспечивать соответствие атрибутов. Однако трудность состоит в выборе надёжных атрибутов. Допустим, измерение яркости неравномерной области – очень сложная задача из-за наличия блестящих поверхностей, градиентов, текстур. Во-вторых, соответствие в частности не гарантирует соответствия в общем [7]. В-третьих, если стоит задача обеспечить сохранение результата с точки зрения человеческого восприятия, то метод объективных измерений яркости областей некорректен и требуется измерять субъективные характеристики, как делалось в работе [13].

Оценка без эталона [14,15,16] состоит в том, что наблюдателей просят сравнить результаты операторов без эталона либо же сопоставить нескольким характеристикам результатов конкретного алгоритма некоторый балл. Хотя оценка без эталона кажется наименее «строгим» из всех методов, на практике она используется наиболее часто. У этого метода есть несколько плюсов:

- 1) субъективная оценка заключает в себе использование аппарата зрения человека;
- 2) большинство наблюдателей имеет большой опыт, связанный с фотографиями и репродукциями и может уверенно оценивать различные параметры получившегося контента;
- 3) сравнительная простота организации: не требуются специальные условия съёмки, можно использовать однажды отснятый контент, не нужно использовать дорогостоящее оборудование.

Вместе с тем есть и большой недостаток, связанный с тем, что можно получить неконсистентные результаты, если использовать малое число

наблюдателей, ставить вопросы о свойствах результата, не до конца понятные наблюдателям, использовать малую выборку. Однако все эти параметры поддаются контролю.

После краткого перечисления методов оценки качества операторов тональной компрессии перейдём к их систематизации. Для систематизации операторов можно использовать различные подходы, например, в работе [7] предлагается такая классификация: симуляторы зрительной системы, операторы репродукции сцены и операторы субъективного качества.

Симуляторы зрительной системы моделируют ограничения и свойства зрительной системы. Примеры: эмуляция ореола вследствие ограниченности ночного зрения человека или уменьшение насыщенности и контраста в темных регионах сцены. Операторы репродукции сцены пытаются сохранить атрибуты исходной сцены: контраст, четкость и цвета. Операторы субъективного качества нацелены на наилучший в терминах субъективных предпочтений и артистических эффектов результат.

Можно производить и классификацию по принципу действия оператора. В этом случае можно выделить глобальные (пространственно-равномерные, TRC – tone reproduction curve) и локальные (TRO – tone reproduction operator, пространственно-неравномерные) операторы. Глобальные операторы используют идентичную кривую сжатия для всех пикселей изображения. Локальные операторы используют информацию о пикселе и его окружении для того, чтобы скорректировать значение яркости вокруг этого пикселя.

Глобальные операторы [17,18,19] имеют простой вид и очень высокую скорость работы, однако при любом выборе кривой происходит потеря либо в бликах, либо в объектах, расположенных преимущественно в тени, ведь чем больше диапазон, тем больше значений нам нужно отобразить. Заметный плюс глобальных операторов – отсутствие артефактов на полученном изображении, так как изменения с пикселями происходят гладко, плавно и локальные соотношения яркости между соседними пикселями не изменяются. Глобальные методы плохо работают с изображениями, где области интереса (детали) равномерно расположены по всему диапазону яркости.

Среди пространственно-зависимых операторов можно выделить несколько подгрупп. Во-первых, это локальные операторы [20], действующие в ограниченной области вокруг данного пикселя. Во-вторых, это частотные операторы [21], использующие разложение изображения на две и более частотных компонент. И, в-третьих, это градиентные операторы [22], модифицирующие изменения яркости на изображениях.

В дополнение к сжатию диапазона яркости, тональная компрессия может подражать воспринимаемых характеристикам, воспроизводя изображения, возбуждающие в зрительной системе тот же отклик, что и реальная сцена, например, пытаясь сохранять контраст, яркость или четкость. Для корректного сжатия диапазона нужно разрабатывать комплексные психофизиологические модели восприятия. Часто методы тональной компрессии сосредотачиваются на единичных аспектах, таких как восприятие яркости. Это происходит, потому что отсутствуют глубокие знания об устройстве зрительной системы человека, поэтому часть исследователей не рассматривает устройство зрительной системы человека вообще и работает исключительно с экспериментальными данными. Резюмируя, можно выделить вторую плоскость классификации алгоритмов тональной компрессии: экспериментальные и теоретические.

Вообще говоря, алгоритмы могут зависеть от абсолютной яркости. Если требуется восстановить той или иной задачи точные значения ДЛЯ освещенности для реальной сцены, то для этого, как правило, необходимо иметь дополнительную информацию об устройстве, на котором производилась (фотоаппарате). Каждому фотоаппарату соответствует съемка свой коэффициент, связывающий значения на матрице с физической энергией. Этот коэффициент можно рассчитать по специальным калибровочным объектам. В случае если изображение создается в результате синтеза, необходимо аккуратно учитывать все физические преобразования яркости в результате отражений и преломлений, чтобы получить на результирующем изображении корректные физические значения. Соответственно, часть алгоритмов тональной компрессии (обычно из числа теоретических) требует физических значений яркости на вход, а другая часть не требует.

Классификация алгоритмов тональной компрессии верна и для временных алгоритмов тональной компрессии. На практике большинство временных алгоритмов тональной компрессии основывается на некотором аналоге для статических изображений.

Алгоритм визуальной адаптации [23] основан на работе [17]. Алгоритм временной адаптации [24] – на работе [25]. Алгоритм [26] – на работе [18]. Алгоритм визуальных выдержек [27] основан на билатеральной фильтрации отдельно во временной и пространственной плоскостях. Адаптивный к дисплею оператор [28] также использует явно выраженную процедуру фильтрации, используя модель отображения пикселя на дисплее. Оператор модели сетчатки [29] основан на пространственно-временной модели, но также явно использует размытие по времени. Оператор временной согласованности [30], применяющий к статическим операторам алгоритм постобработки, наиболее приспособлен к алгоритму [20]. Оператор цветового представления [31] использует вычисления по алгоритму медианного среза и явную процедуру накопления параметров для обеспечения временной устойчивости. Алгоритм локальной адаптации [32] декомпозирует общую модель на статическую часть, связанную с моделированием колбочек и палочек, и динамическую, связанную с адаптацией.

Однако есть и исключения, например, алгоритм на модели колбочек [33] основан на замерах работы реальных колбочек во времени. В исследовании [7] этот алгоритм занял средние позиции относительно других.

Можно заключить, что качество алгоритма для видео в первую очередь определяется качеством базового алгоритма, так как большая часть

предложенных алгоритмов де-факто использует разделение на обработку каждого кадра в отдельности и временную фильтрацию\сглаживание параметров используемой модели.

Для выбора наилучшего базового алгоритма было произведено сравнение существующих алгоритмов для статических изображений.

2 Субъективное тестирование алгоритмов тональной компрессии фотографий в системе GML ePublish

Субъективное сравнение алгоритмов тональной компрессии производилось в системе <u>GMLePublish¹</u>, созданной для сбора и агрегации данных тестирования таких алгоритмов [34]. Эта онлайн-система позволяет испытуемым удаленно проходить тестирования в любое удобное для них время. Исследователю же она предоставляет функционал для загрузки, сбора и анализа данных. Реализована система на базе технологии Microsoft ASP.NET, обмен данными производится с помощью базы данных Microsoft SQL Server.

Функционал GMLePublish достаточно широк, возможно использование разных способов тестирования. Наибольшую надежность, тем не менее, имеет подход с попарным сравнением результатов работы различных алгоритмов тональной компрессии. Именно он использовался в этой работе для субъективной оценки качества.

Главный недостаток попарного сравнения – трудоёмкость. Чтобы его скомпенсировать, часть анализа была проведена автоматически и полуавтоматически, поэтому размер данных для массовых тестов уменьшился.

Прежде всего, было отобрано для сравнения 11 алгоритмов тональной компрессии. Все реализации этих алгоритмов являются авторскими либо промышленными, поэтому исключен риск неправильной реализации этих алгоритмов. Итак, для сравнения были выбраны следующие тональные операторы: Reinhard'02 [31], Reinhard'05 [35], Shan'10 [36], Mantiuk'06 [37], Mantiuk'08 [38], Fattal'02 [22], Durand'02 [21], Drago'03 [19], Pattanaik'98 [39], Ashikhmin'02 [40], Zipa'15 [41].

Для предварительного анализа использовалась внешняя база HDRизображений, состоящая из 33 элементов в формате OpenEXR (Stuttgart[3]), а также база из найденных в сети интернет и отснятых сотрудниками Лаборатории компьютерной графики и мультимедиа 33 изображений в формате HDR (Radiance RGBE). Заметим, что специфика проведения тестирования подразумевает масштабирование результирующих изображений на экране монитора, что затрудняет сравнение деталей изображений. Чтобы повысить надёжность тестирования, было решено расширить базу, добавив к сравнению увеличенные фрагменты изображений, которые отображались бы в исходном разрешении. Фрагменты содержали наибольшие отличия, например, туда входили источники света, сложные текстурированные поверхности и т.д. Всего

¹ http://epublish.gml-team.ru

было дополнительно выбрано 54 фрагмента размером 400х400 пикселей, таким образом, общий размер базы составил 120 изображений.

К каждому изображению из базы был применен каждый из одиннадцати операторов тональной компрессии. Затем производилось попарное сравнение всех восьми алгоритмов для каждого изображения из базы. Поскольку попарное сравнение восьми алгоритмов требует 55 сравнений на изображение то общее число пар сравнения для всей базы изображений составило 6600 сравнений. Для выбора наилучшего метода человеку требуется не менее 5 секунд, а значит, минимальное время для выбора для одного человека составило бы около девяти часов работы за компьютером. Поскольку это очень сложная процедура, то было решено сузить число рассматриваемых алгоритмов. Для этого 1 испытуемый просмотрел всевозможные сравнения для всевозможных изображений и методов из базы.

Обратимся к результатам, полученным попарным сравнением. Каждая победа в сравнении засчитывалась алгоритму за 1 балл, ничья – за 0.5 баллов. Максимальный балл – 10, именно в стольких сравнениях участвует любой алгоритм. Затем для каждого алгоритма был посчитан средний балл среди всех изображений базы (см. рис. 2).



Рис. 2. График, иллюстрирующий рейтинг (среднее значение баллов) рассматриваемых алгоритмов после этапа предварительного ранжирования на двух базах. Отступами отмечены стандартные отклонения.

Оказалось, что некоторые алгоритмы выдают абсолютно некорректный результат (полностью черное или полностью белое изображение) для некоторых экземпляров из исходной базы 66 полномасштабных изображений. Как видно из табл. 1, лишь 5 из 11 рассматриваемых алгоритмов не имели сбоев

на базе. Из 6 алгоритмов со сбоями только Durand'02 имеет 1 сбой, оставшиеся алгоритмы имеют от 13 до 14 сбоев на 66 изображениях.

Таким образом, по совокупности субъективных и объективных значений, было выбрано всего 4 метода. Это Mantiuk'06 (7.35 баллов), Mantiuk'08 (7.39 баллов), Durand'02 (5.87 баллов) и Zipa'15 (8.86 баллов). Остальные методы было решено не проверять, потому что, во-первых, методы Reinhard'02 и Shan'10 имеют разную сортировку на двух базах, во-вторых, метод Shan'10 уступает алгоритму Zipa'15, что было показано в статье [41], в-третьих, результаты остальных алгоритмов намного хуже первых четырёх.

Таблица 1

Алгоритм	Число сбоев	Алгоритм	Число сбоев	Алгоритм	Число сбоев
Ashikhmin'02	0	Pattanaik'98	14	Mantiuk'06	0
Drago'03	16	Reinhard'02	14	Mantiuk'08	0
Durand'02	1	Reinhard'05	14	Zipa'15	0
Fattal'02	13	Shan'10	0		

Число некорректных результатов на базе из 66 изображений

Алгоритмы из тройки лидеров (Mantiuk'06, Mantiuk'08, Zipa'15) основаны на моделировании зрительной системы человека, такие алгоритмы способны лучше решать задачу тональной компрессии. В результате более детального анализа четырёх алгоритмов у них были выявлены достоинства и недостатки (см. табл. 2). Иллюстрации недостатков см. на рис. 3.

Таблица 2

Метод	Достоинства	Недостатки		
Durand'02	 Быстрый. Выдаёт приемлемое качество для широкого класса изображений. 	 Часто пересвечивает или недосвечивает ту или иную область изображения. Плохо передает детали. 		
Mantiuk'08	Очень быстрый.Основан на моделировании ЗСЧ.	 Плохо обрабатывает источники света. Усиливает шум. 		
Mantiuk'06	 Локальный метод. Основан на моделировании ЗСЧ. Передаёт максимальное количество деталей. 	 Очень медленный (до 30 секунд на изображение). Добавляет невидимые черты, придаёт сюрреалистический эффект. 		

Достоинства и недостатки алгоритмов Mantiuk'06, Mantiuk'08, Durand'02 и Zipa'15

Метод	Достоинства	Недостатки
Zipa'15	 Локальный метод. Средняя скорость работы. Хорошо обрабатывает источники в сцене. Сохраняет детали. Не усиливает шум. 	 Плохой глобальный контраст на тёмных сценах.



Рис. 3. Примеры, где алгоритмы дают сбои. Вверху – пересвеченные области на изображениях (Durand'02). По центру – некорректная обработка ярких источников и бликов (Mantiuk'08). Внизу – гиперреализм и шум (Mantiuk'06).

На обозначенных 4 алгоритмах производилось массовое тестирование. Поскольку испытуемые не являлись специалистами в области разработки алгоритмов тональной компрессии, то задача была упрощена – испытуемым не требовалось оценивать результаты алгоритмов тональной компрессии по некоему обобщенному критерию качества, а нужно лишь было оценивать какую-то одну характеристику изображения.

Тестирование было разбито на 5 частей (категорий) по числу исследуемых в изображении признаков. Каждая категория содержала в себе 5 примеров. На каждую из характеристик был создан отдельный опрос в системе GMLePublish (см. рис. 4).

						nome	ADO
Selec	tap	oreset:					
	id	description	type	pres	ets sc <u>er</u>	nes created	
Выбор	19	2015 Шум	preference	4	5	21.01.2016 12:32:02	
Выбор	20	2015 Источник	preference	4	5	21.01.2016 12:32:02	
Выбор	21	2015 Детали	preference	4	5	21.01.2016 12:32:02	
Выбор	22	2015 Контраст	preference	4	5	21.01.2016 12:32:02	
<u>Выбор</u>	23	2015 Естественность	preference	4	5	21.01.2016 12:32:02	
	N						

Рис. 4. Скриншот страницы с доступными тестовыми наборами в системе GMLePublish.

Исследуемые признаки и заданный вопрос:

- 1) Уровень шума На каком изображении меньше выражен шум?
- 2) Детализация источников Где меньше потери цветовой или структурной информации в области источника освещения или блика?
- 3) Общая детализация На каком изображении выше детализация (различимость деталей)?
- 4) Яркость и контрастность У какого изображения лучше общий контраст и яркость?
- 5) Естественность Какое изображение выглядит естественнее?

Для испытуемых была создана инструкция, в которой иллюстрировались недочеты реальных алгоритмов тональной компрессии (на других примерах, нежели в тестировании). Также испытуемые были информированы о минимальных требованиях к размеру и разрешению экрана, на котором они проходят тестирование: запрещалось производить сравнение на разрешении ниже FullHD и диагонали монитора менее 17 дюймов. Такие ограничения были введены, чтобы избежать неверных оценок вследствие пропавшей детализации.

Всего в тестировании приняли участие 31 человек.

В каждый момент времени респондент сравнивал между собой результаты двух случайно выбранных алгоритмов (см. рис. 5) и выбирал из них тот, который лучше отвечал на поставленный вопрос, или же ни один, если изображения оказывались неразличимы по заданному критерию. Респонденты не знали о том, какие именно алгоритмы сравниваются в данный момент.



Рис. 5. Скриншот страницы с попарным тестированием.

После попарного сравнения для каждого алгоритма был посчитан его балл для каждого примера. За каждую победу в попарном сравнении начислялся 1 балл, за проигрыш – 0 баллов, за ничью – 0.5 балла. Таким образом, максимальное теоретическое значение баллов – 3, а минимальное – 0.

Баллы, выставленные каждым пользователем, были усреднены². В результате каждому алгоритму по каждому изображению из категории был выставлен свой балл. Средние ответы пользователей по каждой категории в виде гистограммы см. на рис. 6.

Исходя из критерия Стьюдента, статистически не различимы по уровню шума алгоритмы Mantiuk'06 и Mantiuk'08; статистически не различимы по естественности алгоритмы Durand'02 и Mantiuk'06, а также алгоритмы Mantiuk'08 и Zipa'15; все прочие соотношения между качеством рассмотренных алгоритмов статистически достоверны.

Таким образом, можно сделать выводы о качестве алгоритмов.

Алгоритм Zipa'15 превосходит алгоритм Durand'02 по всем параметрам.

Алгоритм Mantiuk'06 превосходит алгоритм Zipa'15 по общей детализации и детализации источников, но алгоритм Zipa'15 превосходит алгоритм Mantiuk'06 по реалистичности (так как присутствует гипердетализация относительно человеческого восприятия).

² Усреднённые ответы пользователей и изображения, которые они оценивали, можно посмотреть по ссылке http://epublish.gml-team.ru/public/report-jan-2016/averuser.html



Рис. 6. Средние ответы пользователей по каждой категории в виде гистограммы (засечками отмечены стандартные отклонения величин).

Алгоритм Mantiuk'08 сравним с алгоритмом Zipa'15 по уровню естественности, но уступает по всем другим параметрам.

Алгоритм Zipa'15 лидирует по категориям 1 и 4, занимает второе место по категориям 2 и 3, делит первое место с другим алгоритмом в категории 5. Алгоритм Mantiuk'06 лидирует по категориям 2 и 3, занимает третье место в категориях 1 и 4 и последнее место в категории 5. Алгоритм Mantiuk'08 делит первое место в категории 5 и вторые-третьи места в других категориях. Алгоритм Durand'02 занимает последнее место во всех категориях. Итоговый рейтинг алгоритмов по категориям и задачам см. в табл. 3.

Таблица 3

Категория	Durand'02	Mantiuk'06	Mantiuk'08	Zipa'15
Шум	4	3	2	1
Детализация источников	4	1	3	2
Общая детализация	4	1	3	2
Контраст и яркость	4	3	2	1
Естественность	3	4	2	1
Контроль качества = детализация источников+общая детализация	4	1	3	2
Фотография и визуализация = все категории	4	2	3	1

Итоговый рейтинг алгоритмов по категориям и задачам

В задачах, где нужна детализация изображения (например, в задаче контроля качества), наилучшим алгоритмом является Mantiuk'06. Однако в задачах, где важна реалистичность (фотография), а также уровень шума

(визуализация), из рассмотренных алгоритмов лучше всего справится алгоритм Zipa'15, а алгоритм Mantiuk'06 лучше не использовать из-за нереалистичности итогового результата. В задачах, где важна скорость работы и не очень важна детализация, лучшим решением будет алгоритм Mantiuk'08 – он дает довольно реалистичный, хороший по яркости и контрасту результат.

Таким образом, в качестве базового алгоритма для модификации был выбран алгоритм Zipa'15[41]. Далее будут описаны модификации алгоритма, делающие его применимым для работы с видеопоследовательностями.

3 Алгоритм тональной компрессии видеопоследовательностей

Нужно изменить алгоритм [41] таким образом, чтобы при покадровом применении к видеопоследовательности не возникало эффекта мерцания. Математически правильно было бы рассматривать трёхмерные данные и производить вычисление и минимизацию функционала разности в трехмерном окне, однако такая операция требует очень больших вычислительных ресурсов.

Формализуем общую задачу, которую бы потребовалось решать в случае трёхмерных данных. Необходимо минимизировать функционал разности:

$$f(q) = \sum_{p \in w(q)} (I(p) - u(p)Q - v(p))^2 + \varepsilon C(p)^{-2} (u(p) - C(p))^2$$

q – пиксель изображения, w(q) – множество окон, в которое входит q,

p – пиксель окна, *I* – излучательность узкого диапазона,

Q — карта излучательности, *С* — карта рекомендации,

и, v – параметры линейного преобразования,

 $\varepsilon-$ коэффициент регуляризации.

Для трёхмерного случая окно представляет собой трёхмерную структуру, состоящую из 27 (3x3x3) точек. Поскольку общий вид функционала не изменяется для трёхмерного случая, то аналогично статье [41] можно свести задачу его минимизации к решению СЛАУ следующего вида:

$$S * I = b,$$

$$s_{kj} = \sum_{i \mid \{k,j\} \in w} \delta_{kj} - \frac{1}{m_i \Delta_i} \Big((Q(j) - \mu_i) \big(Q(k) - \mu_j \big) + \Delta_i \Big)$$

$$b_k = \sum_{i \mid k \in i} \frac{\varepsilon}{m_i \Delta_i C_i} \Big(Q(k) - \mu_j \Big), \qquad \Delta_i = \sigma^2 + \frac{\varepsilon C_i^{-2}}{m_i}$$

 $m_i = \sum_{j \in w_i} 1$ – число точек в окне, $\mu_i = \frac{1}{m_i} \sum_{j \in w_i} Q(j)$ – среднее излучательности в окне.

Аналогично двумерному случаю, число окон, в которое входит данная точка, равно $(2 * n - 1)^d$, где n – размер окна в каждом измерении, d – число измерений. Для двумерного случая и размера окна 3 это число равно 25, для трёхмерного – 125.

Размерность рассматриваемой разреженной матрицы (*M* * *N*)², где N – число пикселей в изображении, М – длина видеопоследовательности.

Предельное число точек (без учета уменьшения размерности окон на границах изображения) в разреженной матрице для хранения определяется формулой $(n^3)^2 * (M * N) = 729 * (M * N)$. Для Full HD изображения (1920х1280 пикселей) при предварительном уменьшении размерности карты излучательности в 4 раза (рекомендованной авторами статьи) получим число точек M= 153600. При N = 1 (один кадр) получим 1 1197 4400 – а нужно хранить в 4 раза больше значений float для разреженной матрицы.

Таким образом, для реализации подобного алгоритма даже для небольшого числа кадров требуется большой объем оперативной памяти. Поэтому хотя такой алгоритм и может быть, теоретически, разработан, его использование конечными пользователями будет в любом случае затруднительным. Исходя из этих соображений было решено сосредоточиться на алгоритме подавления мерцания при имеющихся результатах покадровой работы алгоритма тональной компрессии.

4 Корректировка глобальных параметров в алгоритме тональной компрессии

Прежде всего необходимо провести глобальную корректировку параметров алгоритма, обеспечив их согласованность на последовательных кадрах одной сцены. Перечислим значения с предыдущего кадра, которые предлагается учитывать при обработке следующего.

Во-первых, нужно хранить коэффициент нормализации светлоты, рассчитанной по модели СІЕСАМ02. Во-вторых, параметры выравнивания гистограммы (см. рис. 7), которое производится трижды – сначала по шкале яркости для преобразованной светлоты, потом для полноцветного изображения в формате LMS, а затем для реконструированного изображения.



Рис. 7. На примере гистограммы красными линиями отмечены позиции почтиминимального и почти-максимального значения, соответствующие квантилям в 0.5% и 99.5%.

Выравниванию гистограммы соответствует линейное преобразование изображения (с последующим обрезанием значений, выходящих за пределы рассматриваемого диапазона [0,1]), поэтому его параметры нужно плавно изменять. Пусть I – входное изображение, J – результирующее, p – пиксель изображения, m и M – квантили яркости, найденные по гистограмме, a и b – параметры линейного преобразования. Тогда линейное преобразование яркости можно описать следующими формулами:

 $J(p) = \frac{I(p)-m}{M-m} = \frac{I(p)}{M-m} - \frac{m}{M-m} = I(p) * a + b; a = \frac{1}{M-m}; b = \frac{m}{M-m}.$

Итак, нужно плавно изменять коэффициент нормализации светлоты и параметры линейных преобразований. Под плавным изменением подразумевается схема с накоплением с некоторой константой α , то есть если предыдущее (полученное на предыдущем кадре) смешиваемое значение было равно *X*, а рассчитанное на этом кадре значение равно *Y*, то текущее значение будет равно $\alpha Y + (1 - \alpha)X$.

В экспериментах достаточным оказалось значение параметра $\alpha = 0.1$. При этом рассчитывались объективные метрики (описаны в разделе 6), получаемые на базе Stuttgart при разных значениях параметра α . Поскольку оптимальные по устойчивости результаты будут получаться при тривиальном $\alpha = 0$, то выбирать этот параметр следует исходя из соображений результирующего качества при увеличении параметра. На рис. 8 даны значения метрик для α с шагом 0.1. Видно, что обе метрики монотонно увеличиваются при увеличении параметра α и достигают минимума при $\alpha = 0$.



Рис. 8. Значения метрик инверсии и немонотонности временного контраста для значений параметра *α* от 0 до 1 с шагом 0.1.

Чтобы выбрать порог, нужно сочетать метрики временной устойчивости с метриками качества результирующего изображений. Если оставить параметр α на уровне 0, то есть использовать для всей сцены параметры с первого кадра, то в случае усиления освещения при старых параметрах будут получаться сильно недосвеченные изображения, а при ослабевании освещения — наоборот. В качестве примера этого недостатка приведём некоторые кадры из

видеопоследовательности fireplace_01 (см. рис. 9). При $\alpha = 0$ (посередине) в определенный момент яркость изображения становится слишком высокой, а детали пропадают. Этого недостатка, однако, лишены результаты алгоритма с параметром $\alpha = 0.1$ (справа), который постепенно и плавно адаптирует яркость кадра. Пример скачка яркости, который ликвидируется с помощью сглаживания параметров, представлен на рис. 10.



Рис. 9. Кадры 89,133,177 из последовательности fireplace_01. Слева – результат применения индивидуального набора параметров к каждому изображению, по центру – результат использования константы α=0, справа – результат использования константы α=0.1.



Puc. 10. Вверху кадр fireplace_01_000783, внизу fireplace_01_000784. Слева результат алгоритма без коррекции, справа – результат с коррекцией с параметром α=0.1.

Для формализации этого наблюдения можно рассмотреть средние яркости кадров входящей видеопоследовательности и средние яркости результирующих видеопоследовательностей (см. рис. 11). При использовании абсолютного сглаживания ($\alpha = 0$) не наблюдается мерцания, однако, несмотря на изменения входных данных, яркость не адаптируется к новому уровню (см. рис. 12). При отсутствии временного сглаживания ($\alpha = 1$) наблюдаются скачки яркости, не связанные с изменением яркости в исходной видеопоследовательности.



Puc. 11. Трек средней яркости результирующих кадров для двух видеопоследовательностей при разных параметрах α. Яркость в HDRI умножена на 500 слева и на 300 справа.



Рис. 12. Одинаковые входные данные, α=1.0 (вверху), 0.1, 0.0 (внизу).

При $\alpha = 0.1$ на изменения яркости в исходной видеопоследовательности производится её коррекция в результирующей видеопоследовательности, а затем яркость постепенно возвращается к оптимальному значению. Этот эффект подобен тому, что происходит при адаптации человеческого глаза, однако заметим, что никакого явного моделирования процесса адаптации в системе человека производится. Bo зрительной не второй видеопоследовательности с менее устойчивой средней яркостью, колебания яркости в итоге незначительны при использовании абсолютного сглаживания и менее выражены, чем в итоговой последовательности, при использовании сглаживания с параметром $\alpha = 0.1$.

Однако, чтобы численно оценить этот эффект на всей базе, потребуется рассчитать некоторую метрику. Поскольку основной недостаток жесткой фиксации параметра – наличие недосвеченных/пересвеченных областей, то можно использовать в качестве такой метрики качества % недосвеченных и пересвеченных пикселей на всём временном отрезке видео. Усреднённые значения этой метрики на базе для различных значений параметра α см. на рис. 13 а).

Можно также оценить и время, через которое параметры начнут изменяться при резком изменении кадра. Например, если $\alpha = 0.1$, то при резком изменении какой-то из регулируемых настроек к противоположному значению (от 0 до 1) пройдёт 67 шагов, прежде чем значение параметра будет отличаться на 10^{-3} . При увеличении параметра время срабатывания будет увеличиваться (см. рис. 13 б).





Зависимость числа кадров отлика можно аппроксимировать функцией $N = \frac{p}{\alpha+q}$ (см. рис. 14). Параметр α выражается исходя из заданного времени обновления параметра t[c] и частоты кадров в видео $v[\frac{\text{кадр}}{c}]$ следующим образом: $\alpha = \frac{p}{t*v} - q, p = 6.778, q = 0.00017.$



Рис. 14. Аппроксимация числа кадров отклика в зависимости от параметра α

Для видео с частотой $v = 24 \left[\frac{\kappa a d p}{c}\right]$, времени обновления t = 3[c] будет соответствовать значение параметра $\alpha = 0.094 \approx 0.1$ (значение по-умолчанию). Таким образом, предложенное компромиссное решение с параметром $\alpha = 0.1$, с одной стороны, позволяет уменьшить мерцание, а с другой – учесть изменения яркости и отобразить их в результирующей последовательности для их компенсации. Параметр можно регулировать и уменьшать его значение для видео с высокой частотой кадров (60 $\left[\frac{\kappa a d p}{c}\right]$).

5 Локальная корректировка для подавления эффекта мерцания

Для локальной корректировки используется алгоритм, основанный на вейвлет-разложении изображений. В каждый момент рассматривается 2 последовательных кадра. К ним последовательно применяется процедура коррекции, начиная с первой пары кадров и заканчивая последней (см. рис. 15).



Рис. 15. Общая схема алгоритма локальной коррекции.

Пусть L_p – предыдущий скорректированный результат работы алгоритма L текущий результат, компрессии, a _ который тональной нужно скорректировать. Пусть этим результатам соответствуют входные данные (HDR-изображения) *H*_pи *H*. Требуется преобразовать изображение *L* в локально скорректированное изображение L'. При этом будем считать, что существует некий набор параметров θ , характеризующий произведенное на предыдущем шаге локальное преобразование. После работы алгоритма необходимо передать новый набор параметров на следующую итерацию (следующей паре кадров). Кроме того, обозначим некоторые неизменяемые настройки алгоритма φ .

Таким образом будем описывать алгоритм *DeF* (сокр. от deflickering – устранение мерцания):



Рис. 16. Схема алгоритма DeF.

Сначала разложим все изображения на частотные компоненты, используя дискретное вейвлет-преобразование.

$$\{w_{H}(s, o, p)\}_{s=\overline{1:N}, o=\overline{1:3}, p\in P} = DWT(H), \{w_{H_{p}}(s, o, p)\}_{s=\overline{1:N}, o=\overline{1:3}, p\in P} = DWT(H_{p})$$

$$\{w_{L}(s, o, p)\}_{s=\overline{1:N}, o=\overline{1:3}, p\in P} = DWT(L), \{w_{L_{p}}(s, o, p)\}_{s=\overline{1:N}, o=\overline{1:3}, p\in P} = DWT(L_{p})$$

 s, o – текущий масштаб и ориентация, $p \in P \in \mathbb{N}^{W \times H}$ – пиксель изображения, $N = log2(\max(W, H))$ – максимально доступный уровень детализации, W, H – ширина и высота изображений).

Далее последовательно рассматриваем все уровни детализации, начиная с наименьшего, и ищем то значение, начиная с которого уровень различия между текущим и предыдущим HDR-изображением превосходит некоторое наперёд заданное значения. Назовём этот уровень (пороговый индекс детализации – threshold index of detailing) *TIoD*. Измерять степень отличия между двумя изображениями будем по метрике SSIM[42]. Установим некоторое минимальное значение детализации, ниже которого не будем производить изменений в изображении, назовём его *minTIoD*. Также предположим, что на предыдущем шаге был найден пороговый индекс детализации *oldTIoD*.

Если oldTIoD > TIoD, то для всех уровней детализации, больших max(*minTIoD*, *TIoD*), осуществляем альфа-смешивание. Если oldTIoD < TIoD, то для всех уровней детализации, больших max(*minTIoD*, *TIoD*), осуществляем альфа-смешивание, а для всех уровней детализации, больших max(*minTIoD*, *oldTIoD*), осуществляем частичное альфа-смешивание (только для определенных пикселей), остальные пиксели оставляем без изменений.

В конце, после смешивания (где необходимо) всех компонент, производим обратную сборку из модифицированных компонент разложения в изображение $L' = IDWT(\{w_{L'}(s, o, p)\}_{s=\overline{1:N}, o=\overline{1:3}, p \in P}).$

Опишем чуть подробнее некоторые детали (см. рис. 17).



Рис. 17. Детализация алгоритма DeF.

Чтобы получить метрику SSIM для изображения, нужно усреднить значение функционала в каждом пикселе, основанного на статистиках в некотором квадратном окружающем окне. Размер окна выбран равным 11. Пусть мы рассматриваем пиксель *q*. Этот пиксель принадлежит одновременно

изображению H_p и H. Тогда мы можем посчитать средние и стандартные отклонения по окнам в одном и другом изображении, обозначим их $\mu_H(q), \sigma_H(q), \mu_{H_p}(q), \sigma_{H_p}(q)$. Обозначим ковариацию между ними как $\sigma_{HH_p}(q)$.Тогда значение функционала в пикселе будет вычисляться по формуле:

$$F_{ssim}(q) = \frac{\left(2\mu_H(q)\mu_{H_p}(q) + c_1\right)\left(\sigma_{HH_p}(q) + c_2\right)}{\left(\mu_H^2(q) + \mu_{H_p}^2(q) + c_1\right)\left(\sigma_H^2(q) + \sigma_{H_p}^2(q) + c_2\right)},$$

c1 = 0.01 * *L*, c2 = 0.03 * *L* – некоторые константы,
L – динамический диапазон пикселя.

Динамический диапазон пикселя *L* выбирается равным не 255, как для изображения узкого диапазона, а разнице между почти максимальным и почти минимальным уровнем яркости, соответствующим квантилям в 1% и 99% гистограммы яркости.

Порог *TIoD* соответствует уровню вейвлет-разложения, начиная с которого метрика SSIM превосходит наперёд заданное число T_a =0.999.

Полное альфа-смешивание применяется ко всем пикселям слоя (\bar{s} , o):

 $\forall o, p \, w_{L'}(\bar{s}, o, p) = \beta w_L(\bar{s}, o, p) + (1 - \beta) w_{L_p}(\bar{s}, o, p) \}.$

Частичное альфа-смешивание применяется только к тем пикселям p, в которых есть локальные отличия ($F_{ssim}(s, p)$ больше порога $T_l = 0.99$).

Итого, глобальные параметры $\varphi = \{T_g, T_l, minTloD, \beta\}$. Значения по умолчанию для этих параметров: $T_g = 0.999, T_l = 0.99, minTloD = 4, \beta = 0.2$. Изменяемый набор локальных параметров θ состоит только из одного параметра $oldTloD = \beta TloD + (1 - \beta)oldTloD$.

6 Метрика мерцания для видео последовательностей

В обзоре [7] для выявления мерцания не использовалась какая-то автоматическая метрика, вместо этого авторы продемонстрировали два примера трека яркости в видеопоследовательности, когда алгоритмы давали мерцание. При таком ручном анализе авторы показывали, что если в исходном треке наблюдается монотонность, а в результирующем – скачки, то алгоритм неустойчив. Такой подход вполне годится, когда задача исследования – проведение сравнения существующих алгоритмов, однако при разработке и модификации нового алгоритма необходимо более удобное автоматическое средство для определения степени мерцания и наличия артефактов. Поэтому в этом исследовании был использован более формальный автоматический подход, обобщающий идею [7].

Метрика мерцания алгоритма тональной компрессии ДЛЯ видеопоследовательностей принимает на ВХОД две последовательности изображений: исходную последовательность широкого динамического диапазона и результирующую (обработанную тестируемым алгоритмом тональной компрессии) последовательность узкого динамического диапазона. На выходе метрика выдает вещественное число, соответствующее степени выраженности мерцания. Метрика для видеопоследовательности – усреднение результата метрики для трека пикселя изображения:

$$F(V,V') = \frac{1}{N} \sum_{p=1}^{N} f(T_p,T_p'),$$

$$V = \{T_p | p = 1..N\} - HDRV, V' = \{T_p' | p = 1..N\} - LDRV,$$

N—число пикселей на каждом кадре видео,

 T_p – исходный трек пикселя p, T'_p – результирующий трек пикселя p.

Метрика для треков пикселей определяет, имеет ли место сбой монотонности яркости в исходной и результирующей последовательности, и в случае различного поведения добавляет некоторый штраф на каждом кадре.

$$f(T,T') = \frac{1}{(M-2)} \sum_{t=3}^{M} change \mathbb{I}(T(t), T(t-1), T(t-2), T'(t), T'(t-1), T'(t-2), \alpha)$$

* penalty $(T'(t), T'(t-1), T'(t-2))$

T — исходный трек длины M, T′ — результирующий трек длины M

T(t) — значение трека в момент времени t.

Индикатор смены монотонности срабатывает в случае, если монотонность была в исходной видеопоследовательности, но исчезла в результирующей:

change
$$\mathbb{I}(T(t), T(t-1), T(t-2), T'^{(t)}, T'^{(t-1)}, T'^{(t-2)}, \alpha)$$

$$= mon \mathbb{I}(T(t), T(t-1), T(t-2), \varepsilon_{HDR}) \&! mon \mathbb{I}(T'^{(t)}, T'^{(t-1)}, T'^{(t-2)}, \varepsilon_{LDR}), \\ \varepsilon_{HDR} - допустимая погрешность в HDR - данных,$$

*ε*_{LDR} – допустимая погрешность в LDR – данных.

Индикатор монотонности работает с четырьмя числами: тремя последовательными значениями яркости и параметром – допустимой разницей между числами для признания их равными.

$$mon\mathbb{I}(X, Y, Z, \varepsilon) = \begin{cases} 1, & если \\ (X > Y + \varepsilon \& Y > Z + \varepsilon) \\ | \\ (X < Y - \varepsilon \& Y < Z - \varepsilon) \\ | \\ (|X - Y| < \varepsilon \& |Y - Z| < \varepsilon) \\ 0, & иначе \end{cases}$$

Допустимая погрешность для LDR и HDR данных должна выбираться согласованно, например, её можно брать в процентах диапазона входных и выходных данных, то есть пусть m и M – максимум и минимум видеопоследовательности или первого кадра, тогда погрешность зависит от единого параметра следующим образом: $\varepsilon = \alpha (M - m), \alpha -$ параметр метрики мерцания.

Было выбрано значение параметра $\alpha = 0.003$, то есть на LDR-данных регистрируется смена яркости даже в 1 значение из 255, а на HDR-данных регистрируется смена яркости в 0.3% диапазона.

Штраф (см. рис. 18) обратно пропорционален величине монотонности:

$$penalty(X,Y,Z) = \frac{\max(|X-Y|,|Y-Z|)}{|X-Z|}$$

Z, Y, X — последовательные значения яркости в результирующем треке .



Рис. 18. Расчет штрафа. Синяя линия – трек. В обоих случаях величина штрафа будет одинакова и равна 2.5, хотя в первом она будет вычисляться по формуле $\frac{a}{c}$, а во втором – $\frac{b}{c}$.

Таким образом, полностью определена метрика мерцания. Если преобразование полностью сохраняет монотонность (не производит новых скачков яркости), то метрика равна 0. Если исходная видеопоследовательность статична, монотонно возрастает или монотонно убывает (с точностью до ε_{HDR}), то штрафуется любой пикообразный скачок высоты больше ε_{LDR} . Если исходная видеопоследовательность немонотонна, то штраф не начисляется. Чем сильнее скачок в ожидаемо монотонном треке, тем больше штраф.

Аналогично можно определить и метрику инверсии временного контраста, которая позволяет оценивать выраженность некоторых артефактов. Эта метрика вычисляется намного проще, так как рассматривает только два последовательных значения в треке.

$$R(V,V') = \frac{1}{N} \sum_{p=1}^{N} r(T_p, T_p'),$$
$$V = \{T_p | p = 1..N\} - HDRV, V' = \{T_p' | p = 1..N\} - LDRV$$

N — число пикселей на каждом кадре видео,

 T_p — исходный трек пикселя p, T_p' — результирующий трек пикселя p,

$$r(T,T') = \frac{1}{(M-1)} \sum_{t=2}^{M} rev \mathbb{I}(T(t), T(t-1), T'(t), T'(t-1), \alpha) * |T'^{(t)} - T'(t-1)|$$

T – исходный трек длины M, T' – результирующий трек длины M, T(t) – значение трека в момент времени t.

Индикатор инверсии:

$$rev\mathbb{I}(T(t), T(t-1), T'(t), T'(t-1), \alpha) = \begin{cases} 1, если \\ ((T(t) > T(t-1) + \varepsilon_{HDR}) \& (T'(t) < T'(t-1) - \varepsilon_{LDR})) \\ | \\ ((T(t) < T(t-1) - \varepsilon_{HDR}) \& (T'(t) > T'(t-1) + \varepsilon_{LDR})) \\ 0, иначе \end{cases}$$

В практическом использовании эта метрика оказалась менее полезна, так как инверсия временного контраста также косвенно оценивается метрикой

мерцания. Примеры численной оценки степени мерцания для видеопоследовательностей см. на рис. 19 и 20.



Рис. 19. Примеры треков с максимальными значениями метрик для каждого алгоритма из алгоритмов Mantiuk'08 и Zipa'16 (предложенного).



Рис. 20. Представлены 3 кадра одной сцены. Вверху – исходные HDRI (логарифмическое отображение), посередине – результат работы алгоритма Mantiuk'08 (F = 0.0271, R = 0), внизу – результат работы предложенного алгоритма Zipa'16 (F = 0.0004, R = 0).

7 Тестирование алгоритма

Оценивалось влияние произведённых модификаций на качество работы алгоритма по метрикам из раздела 6. В табл. 4 приведены численные значения метрик инверсии *R* и мерцания *F* при изменении параметра α . В табл. 5 приведены численные значения тех же метрик, а также примеры карт инверсии и мерцания при использовании дополнительной локальной фильтрации (для значений параметра α =0.0 и α =0.1).

Как видно из приведённых таблиц, как глобальная, так и локальная коррекция позволяют уменьшить значения метрик. При использовании параметра α =0.1 и локальной коррекции достигается улучшение по метрике инверсии временного контраста в 6 раз относительно покадрового применения алгоритма (параметр α =1), а по метрике мерцания – в 11 раз.

Таблица 4

Метрики инверсии и немонотонности временного контраста при различных значениях параметра α

α	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
R	0.11	0.12	0.12	0.13	0.13	0.14	0.16	0.17	0.18	0.22	0.24
F	0.011	0.012	0.016	0.023	0.030	0.033	0.039	0.044	0.052	0.053	0.070

Таблица 5

Метрики инверсии и мерцания при использовании локальной фильтрации, примеры карт для примеров beerfest_lightshow_02, fishing_longshot и увеличенного фрагмента из bistro_1.

	α=1 (без	$\alpha=1$ (без $\alpha=0$, без $\alpha=0$, с $\alpha=0.1$, без		α=0.1, c	
	сглаживания)	локальной	локальной	локальной	локальной
		коррекции	коррекцией	коррекции	коррекцией
R	0.24	0.107	0.081	0.118	0.041
F	0.07	0.011	0.009	0.012	0.006
					and the second second

Заключение

В работе предложено решение задачи тональной компрессии изображений широкого динамического диапазона для видео. Это решение было реализовано и применено к набору видео-данных, содержащих сцены с широким динамическим диапазоном и представляющим интерес для обработки алгоритмами тональной компрессии.

Был произведён тщательный анализ существующих алгоритмов тональной компрессии, выявлены их достоинства и недостатки. Также был произведён обзор методик для оценки качества результатов работы алгоритмов тональной компрессии. По результатам обзора был выбран метод оценки без эталона и осуществлено субъективное тестирование существующих алгоритмов в системе GMLePublish.

По результатам тестирования был выбран алгоритм, продемонстрировавший наилучшие результаты по комплексу оцениваемых свойств. Этот алгоритм был модифицирован для устранения эффекта мерцания в случае применения к набору кадров (видеоряду).

Для оценки степени мерцания предложена пара объективных метрик, позволяющая численно сравнивать степень временных искажений по отношению к исходному сигналу. На основании этих метрик, а также объективных метрик оценки качества результирующего видео, выбрано оптимальное значение коэффициента гладкости глобальной коррекции α=0.1.

Помимо глобальной коррекции, предложен также алгоритм локальной коррекции результата, работающий по принципу пост-обработки результирующего видео.

Оба алгоритма коррекции, предложенные в работе, позволили уменьшить степень мерцания в результирующих изображениях. Алгоритм локальной коррекции может также быть применён к другим базовым алгоритмам тональной компрессии, так как не зависит от параметров базового алгоритма.

Библиографический список

1. Debevec P., Jitendra Malik. Recovering high dynamic range radiance maps from photographs // Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. New York. 1997. P. 369-378. URL: http://dx.doi.org/10.1145/258734.258884

2. Lapray P.-J., Heyrman B. and Ginhac D. HDR-ARtiSt: an adaptive real-time smart camera for high dynamic range imaging // Journal of Real-Time Image Processing, 2014. P. 1-16. URL: <u>http://dx.doi.org/10.1007/s11554-013-0393-7</u>

3. Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays / Froehlich, J. [и др.] // Proc. SPIE Electronic Imaging, vol. 9023, 2014. p. 1-10. URL: <u>http://dx.doi.org/10.1117/12.2040003</u>

4. Unger J., Gustavson S. High-dynamic-range video for photometric measurement of illumination // Proc. SPIE, vol. 6501, 2007. P. 1-10. URL: <u>http://dx.doi.org/10.1117/12.703050</u>

5. A versatile HDR video production system / Tocci M. D. [и др.] // ACM Trans. Graph., vol. 30(4), 2011. P. 1-10. URL: http://doi.acm.org/10.1145/2010324.1964936

6. Unified HDR reconstruction from raw CFA data / Kronander J [и др.] // In Proceedings of the IEEE International Conference on Computational Photography. 2013. P. 1-10

7. Evaluation of Tone Mapping Operators for HDR-Video / Eilertsen, G. [и др.] // Computer Graphics Forum, vol. 32(7), 2013. P. 275-284. URL: http://dx.doi.org/10.1111/cgf.12235

8. Perceptual evaluation of tone mapping operators with real world scenes / Yoshida A [и др.] // Proc. SPIE Human Vision and Electronic Imaging X, vol. 5666. San Jose, USA. 2005. Pp. 192-203

9. Villa C., Labayrade R. Psychovisual assessment of tone-mapping operators for global appearance and colour reproduction // In Proc. of Colour in Graphics Imaging and Vision 2010. Joensuu, Finland. 2010. P. 189-196

10. Evaluation of tone mapping operators using a high dynamic range display / Ledda P. [и др.] // ACM Trans. Graph., vol. 24(3), 2005. P. 640-648. URL: <u>http://doi.acm.org/10.1145/1073204.1073242</u>

11. Kuang J., Heckaman R., Fairchild M. D. Evaluation of HDR tone-mapping algorithms using a high-dynamic range display to emulate real scenes // Journal of the Society for Information Display, vol. 18(7), 2010. P. 461-468

12. McCann J. J., Rizzi A. The Art and Science of HDR Imaging // Chichester, West Sussex, UK. : Wiley, 2012.

13. Zipa K., Ignatenko A. Estimation of object's integral brightness. // 2015 IEEE International Conference on Intelligent Computer Communication and Processing. Cluj-Napoca, Romania. 2015. P. 359-66.

14. Testing HDR image rendering algorithms. / Kuang J. [и др.] //In Proc. IS&T/SID 12th Color Imaging Conference. Scotsdale, USA. 2004. P. 315–320

15. Analysis of reproducing real-world appearance on displays of varying dynamic range / Yoshida A. [и др.] // Computer Graphics Forum, vol. 25(3), 2006. P. 415-426.

16. Petit J., Mantiuk R. K. Assessment of video tonemapping: Are cameras' Sshaped tone-curves good enough? // Journal of Visual Communication and Image Representation, vol. 24(7), 2013. P. 1020-1030 URL: http://dx.doi.org/10.1016/j.jvcir.2013.06.014

17. Ward G. A contrast-based scalefactor for luminance display. In Graphics gems V. San Diego: Academic Press Professional, Inc., 1994. P. 415-421

18. Larson G.W., Rushmeier H., Piatko C. A visibility matching tone reproduction operator for high dynamic range scenes // IEEE Transactions on

Visualization and Computer Graphics, vol. 3(4). 1997 P. 291-306 URL: http://dx.doi.org/10.1109/2945.646233

19. Adaptive Logarithmic Mapping For Displaying High Contrast Scenes / Drago F. [и др.] // Computer Graphics Forum, vol. 22(3), 2003. P. 419-426.

20. Photographic tone reproduction for digital images / Reinhard E.[и др.] // ACM Trans. Graph, vol. 21(3), 2002. P. 267-276. URL: <u>http://doi.acm.org/10.1145/566654.566575</u>

21. Durand F., Dorsey J. Fast Bilateral Filtering for the Display of High-Dynamic Range Images // ACM Trans. Graph., vol. 21(3), 2002. P. 257-266. URL: http://doi.acm.org/10.1145/566654.566574

22. Fattal R., Lischinski D., Werman M. Gradient domain high dynamic range compression // ACM Trans. Graph., vol. 21(3), 2002. P. 249-256. URL: <u>http://doi.acm.org/10.1145/566654.566573</u>

23. A model of visual adaptation for realistic image synthesis / Ferwerda J. A [и др.] // Proceedings of the 23th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA. 1996. P. 249-258

24. Time-dependent visual adaptation for fast realistic image display / Pattanaik S. N [и др.] // Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York. USA, 2000. P. 47-54. URL: <u>http://dx.doi.org/10.1145/344779.344810</u>

25. Hunt R. W. G. The Reproduction of Colour. England: Fountain Press, 1995.

26. Irawan P., Ferwerda J. A., Marschner S. R. Perceptually based tone mapping of high dynamic range image streams // In Proc. of Eurographics Symposium on Rendering. June 2005.

27. Bennett E. P., Mcmillan L. Video enhancement using per-pixel virtual exposures // ACM Trans. Graph, vol. 24(3), 2005. P. 845-852. URL: <u>http://doi.acm.org/10.1145/1073204.1073272</u>

28. Mantiuk R., Daly S., Kerofsky L. Display adaptive tone mapping // ACM Trans. on Graph, vol. 27(3), 2008. P.1-10. URL: http://doi.acm.org/10.1145/1360612.1360667

29. Spatio-temporal Tone Mapping Operator Based on a Retina Model. In Computational Color Imaging. / Benoit A. [и др.] // Berlin: Springer-Verlag, 2009.

30. Temporal coherency for video tone mapping / Boitard R.[и др.] // In Proc. of SPIE 8499, Applications of Digital Image Processing XXXV. October 2012. P.1-10

31. Calibrated image appearance reproduction. / Reinhard E.[и др.] //ACM Trans. on Graph., vol. 31(6), 2012. P.1-11. URL: http://doi.acm.org/10.1145/2366145.2366220

32. Ledda P., Santos L. P., Chalmers A. A local model of eye adaptation for high dynamic range images. // Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa. Stellenbosch, South Africa. 2004. P. 151-160. URL: http://doi.acm.org/10.1145/1029949.1029978 33. Van Hateren J. H.: Encoding of high dynamic range video with a model of human cones // ACM Trans. Graph, 25(4), 2006. P.1380-1399 URL: <u>http://doi.acm.org/10.1145/1183287.1183293</u>

34. Матросов М. А., Игнатенко А. В. GMLePublish: web-система оценки алгоритмов тональной компрессии hdr-изображений // Научно-практический семинар "Новые информационные технологии в автоматизированных системах-16". Москва. 2013.

35. Reinhard E., Devlin K. Dynamic Range Reduction Inspired by Photoreceptor Physiology //IEEE Transactions on Visualization and Computer Graphics, 11(1), 2005. P. 1-12. URL: <u>http://erikreinhard.com/papers/tvcg2005.pdf</u>

36. Shan Q., Jia J., Brown M.S. Globally Optimized Linear Windowed Tone Mapping // IEEE Transactions on Visualization and Computer Graphics, 16(4), 2010. P. 663-675. URL: <u>http://doi.ieeecomputersociety.org/10.1109/TVCG.2009.92</u>

37. Mantiuk R., Myszkowski K., Seidel H.-P. A Perceptual Framework for Contrast Processing of High Dynamic Range Images //ACM Transactions on Applied Perception, 3(3), 2006. P. 286-308. URL: http://doi.acm.org/10.1145/1166087.1166095

38. Mantiuk R., Daly S., Kerofsky L. Display Adaptive Tone Mapping // ACM Trans. Graph., 27(3), 2008. P. 1-10. URL: http://doi.acm.org/10.1145/1360612.1360667

39. A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display / S. N. Pattanaik [и др.] // Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. New York. 1998. P. 287-298 URL: <u>http://doi.acm.org/10.1145/280814.280922</u>

40. M. Ashikhmin. A tone mapping algorithm for high contrast images // In Proceedings of Eurographics Workshop on Rendering. 2002. P. 145-156

41. Зипа К., Игнатенко А. Алгоритм тональной компрессии на основе восприятия человека // Труды международной научной конференции GraphiCon'2015. Протвино, Россия. 2015. Р. 194-199. URL: <u>http://graphicon.ru/html/2015/papers/39.pdf</u>

42. Image quality assessment: from error visibility to structural similarity / Wang Z. [и др.] // IEEE Transactions on Image Processing, vol. 13 (4), 2004. P. 600–612.