

Использование нейронных сетей в задачах HDR

Р. С. Ковтун

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
vasabi.ph@gmail.com

С. С. Чупров

Rochester Institute of Technology
sc1723@rit.edu

Р. И. Гатауллин

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
rusfiner@mail.ru

А. Д. Ручкан

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
aruchkan@yandex.ru

А. Алхасан

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
aliyossefahasan@gmail.com

И. И. Виксин

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
wixnin@mail.ru

Аннотация. Статья представляет обзор современных методов исследований и применения нейронных сетей в задаче захвата и отображения широкого динамического диапазона сцен (HDR). Освещается суть проблемы и актуальность изучения данной темы. Рассматривается классический подход к решению задачи HDR и выявляются его ограничения, которые открывают путь для внедрения нейронных сетей в эту область. В работе представлен обзор применения различных типов архитектур машинного обучения в обработке HDR изображений. Рассматриваются свёрточные (CNN), генеративно-состязательные (GAN) и рекуррентные нейронные сети (RNN). Также анализируются методы с применением обучения без учителя. На основе обзора разработаны рекомендации по использованию различных методов обработки HDR изображений в различных сценариях.

Ключевые слова: HDR, широкий динамический диапазон, тональная компрессия, CNN в задаче HDR, GAN в задаче HDR, RNN в задаче HDR

I. ВВЕДЕНИЕ

В области цифровой обработки изображений существует постоянная потребность сохранять высокую степень детализации в сценах с широким динамическим диапазоном. Динамический диапазон (dynamic range) в изображениях представляет собой контраст между самой светлой и самой тёмной областями сцены. Этот параметр измеряется в ступенях экспозиции, также известных как «стопы» (stops, EVs). Один стоп соответствует удвоению (или делению надвое) количества света на сенсоре камеры.

Динамический диапазон человеческого глаза составляет около 12–14 стопов, когда большинство камер работают в диапазоне 6–8 ступеней. Разница проявляется в момент, когда камера снимает контрастную сцену (яркое окно в тёмном помещении, яркая лампа в ночи и т. п.). В таких условиях динамический диапазон сцены выходит за пределы

возможности сенсора, что приводит к пересветам и/или провалам в тенях на итоговом снимке (рис. 1). Таким образом, ограниченность сенсоров является причиной потери деталей сцены на снимке.

Один из способов решения проблемы сохранения деталей в сценах с высоким динамическим диапазоном – использование алгоритмов High Dynamic Range (HDR). Основным этапом в создании HDR изображений является брекетинг экспозиции (exposure bracketing). Этот процесс заключается в съёмке нескольких кадров одной сцены с различной экспозицией.



Рис. 1. Пример съёмки контрастной сцены. Стены видны хорошо, но окно полностью пересвечено

Так формируется набор изображений, где каждый снимок имеет разный уровень яркости. Объединение этих кадров позволяет создать одно HDR изображение с более широким динамическим диапазоном.

Для дальнейшего обсуждения необходимо более подробно рассмотреть процесс формирования HDR изображения с описанием возможных проблем на

каждом из этапов (схема всех этапов и их возможных огрехов показана на рис. 2).

А. Брекетинг экспозиции

Съёмка нескольких кадров одной сцены с различной экспозицией. При этом необходимо выбрать количество снимков и разницу ступеней между кадрами. Эти кадры имеют низкий динамический диапазон (low dynamic range или LDR)

В. Выравнивание изображений

Прежде чем объединять снимки, их необходимо выровнять, чтобы избежать сдвигов границ объектов на итоговом изображении. При выборе алгоритма выравнивания, следует учитывать скорость и качество результата работы алгоритма (при этом алгоритм должен быть достаточно устойчив к разнице экспозиций). Быстрым, но не точным является алгоритм битовой карты медианного порога (MTB) [1, с.155–160], относительно медленным, но более точным является алгоритм улучшенного коэффициента корреляции (ЕСС) [2].

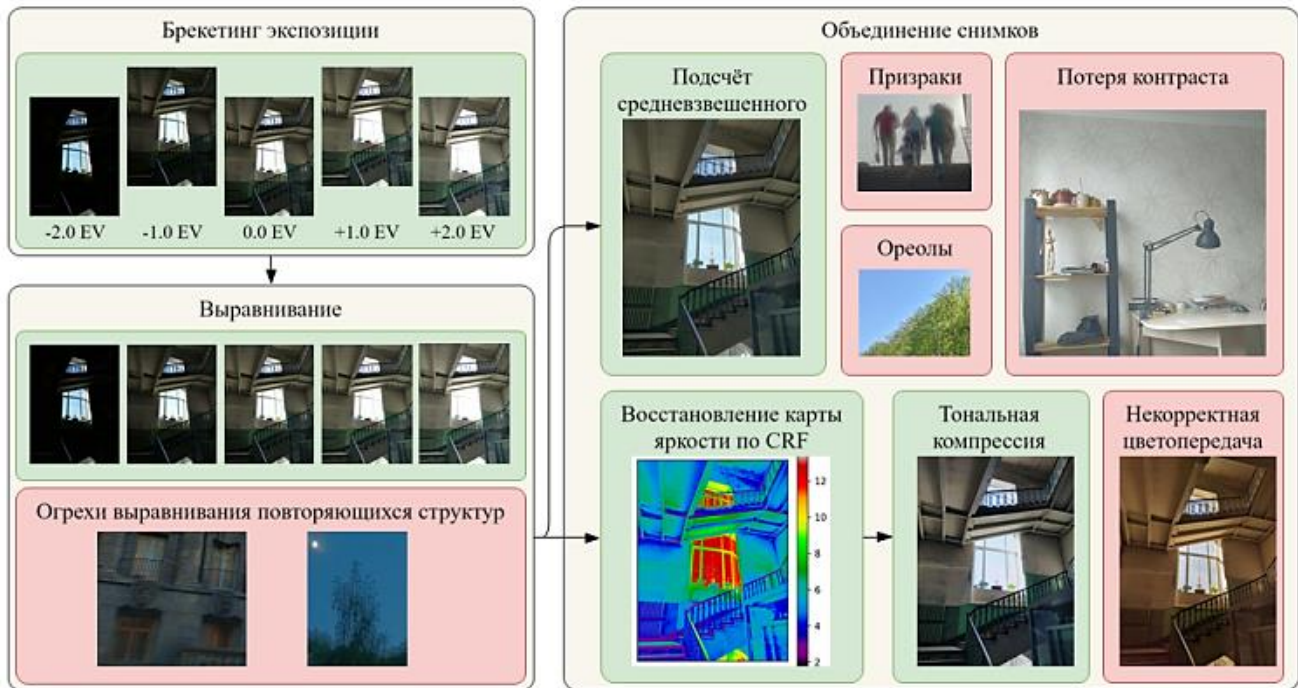


Рис. 2. Схема получения изображения с расширенным динамическим диапазоном. Стрелками указана последовательность обработки (начиная с брекетинга экспозиции). Зелёным цветом выделены результаты этапов, красным – артефакты, присущие этим шагам (для этого и последующий рисунок поощряется масштабирование читателем для детального рассмотрения изображений)

С. Объединение снимков

Ключевой шаг в создании HDR-изображений. На данном этапе требуется понять, какой сегмент каждого из снимков несёт полезную информацию и собрать из этих частей итоговую картину. Существует несколько подходов к реализации данного этапа:

- Подсчёт средневзвешенного между снимками [3].
- Объединение снимков с помощью восстановленной обратной функции отклика камеры (CRF) в файл hdr-формата [4], [5]. Этот файл способен хранить снимки с динамическим диапазоном до 255 стопов. В такой реализации необходим дополнительный шаг, чтобы отобразить снимок на экране, так как экраны чаще всего имеют 8 бит на канал и, соответственно, могут отобразить лишь 8 стопов. Цель *тональной компрессии* (tone mapping) в сжатии диапазона снимка с минимальными потерями [6], [7].

Для обоих подходов характерны артефакты в виде призрачных силуэтов (ghosts), ореолов (halos), потери контраста, некорректной цветопередачи, которые можно

увидеть на рис. 2. Особенно часто артефакты возникают, когда объект съёмки не неподвижен (например, идущий человек).

Далее предлагается ознакомиться с современными методами получения HDR-изображений с использованием различных архитектур машинного обучения, выделить их преимущества и недостатки, а также сравнить классический и современный подходы и сделать выводы о сфере применимости рассмотренных архитектур.

II. CNN в ЗАДАЧЕ HDR

Свёрточные нейронные сети (convolutional neural networks) помогают извлечь базовые признаки изображения, поэтому слои свёртки используются во всех методах HDR, которые применяют нейронные сети.

Одним из основных способов применения CNN в задаче HDR является шаг объединения экспозиций. CNN позволяют учитывать информацию о контексте и структуре изображения при объединении экспозиций, что помогает избежать появления «призраков» при объединении изображений. Контекст необходим чтобы

определить, насколько сдвинулся объект съёмки (или сдвинулся ли он вообще), при этом один из кадров серии выбирается в качестве опорного – сдвиги объектов считаются относительно него (чаще всего, это кадр с автоматическим замером экспозиции, то есть при относительном экспозиционном числе, равном 0.0).

Архитектуры с использованием CNN имеют структуру, близкую к представленной на рис. 3. Примером может служить архитектура DeepHDR [8], которая использует CNN, для обработки HDR изображений именно с целью устранения «призраков». В этом методе предложена архитектура, которая обучается на пакете LDR в качестве входа (с указанием опорного

кадра) и соответствующим эталонном HDR в качестве выхода (эта архитектура не реализует блок Tonemapper, оставляя тональную компрессию на пользователя). Отличительной особенностью архитектуры является наличие skip-соединений от кодировщиков к декодировщику. Во время тренировки сети такие соединения позволяют ошибке распространяться на более глубокие слои, а во время прямого прохода на этапе декодирования (повышения разрешения) – предоставлять информацию об изначальной структуре изображения.

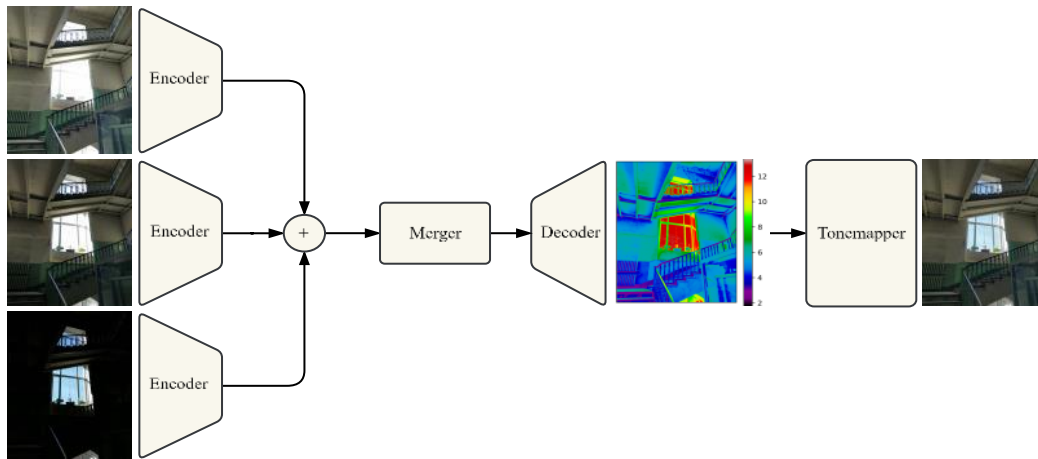


Рис. 3. Схема типичной HDR-архитектуры с использованием CNN. На вход подаются выровненные снимки брекетинга с указанием опорного кадра. Блоки **Encoder** понижают разрешение входных кадров (down-sampling), чтобы облегчить вычислительную нагрузку нахождения базовых признаков. Далее слои конкатенируются в один тензор и объединяются в одно HDR-изображение блоком **Merger**. После этого блок **Decoder** повышает разрешение (up-sampling) и остаётся провести тональную компрессию **Tonemapper**

Другой сферой применения CNN в контексте получения HDR изображений является тональная компрессия. Нелинейные зависимости в нахождении баланса между сохранением деталей и улучшением визуального восприятия изображения могут быть найдены с использованием свёрточных сетей.

Например, метод тональной компрессии с использованием Local Laplacian Filters (LLF) [9] предлагает адаптивный подход к коррекции яркости с учётом локальной структуры изображения. В данном методе анализ исходного HDR-файла происходит на различных масштабах, каждый из которых содержит информацию об уровнях деталей на разных пространственных частотах. Такой подход позволяет исключить (или сильно уменьшить) эффект «ореола».

III. RNN в задаче HDR

Рекуррентные нейронные сети (RNN) представляют собой класс нейронных сетей, способных учитывать последовательности данных и зависимости между ними, что делает их полезными для обработки изображений с учетом их структуры и контекста.

В данном случае, имеется в виду контекст информации в изображении при принятии решений о коррекции яркости или контраста. Такие сети могут учитывать окружающие пиксели для определения наилучшего уровня яркости или контраста для каждого

пикселя. Этот подход позволяет получить HDR-изображение, сгенерировав недостающие участки из единственного кадра (так называемые Single-exposure HDR методы). Или же можно сгенерировать с помощью RNN несколько экспозиций из *единственного* снимка и далее объединять их традиционными методами [10]. И хотя ни в том, ни в другом случае не выходит «честного» расширения динамического диапазона (то есть, когда детали сцены берутся из реальных снимков этой сцены, а не генерируются), такие методы дают естественный результат и вычисляются быстро.

Рекуррентные сети тоже способны решить задачу объединения снимков с движущимися объектами. Prabhakar et al. предлагают собственную архитектуру рекуррентной ячейки – self-gated-memory (SGM) [11], которая позволяет эффективно и вне зависимости от размера и количества артефактов получить результат без «призраков». На рис. 4 представлена предлагаемая архитектура такой сети. При этом блок RNN может быть представлен в виде ячейки краткосрочной и долговременной памяти (LSTM), управляемым рекуррентным блоком (GRU), ячейкой саморегулируемой памяти (SGM) и другими решениями.

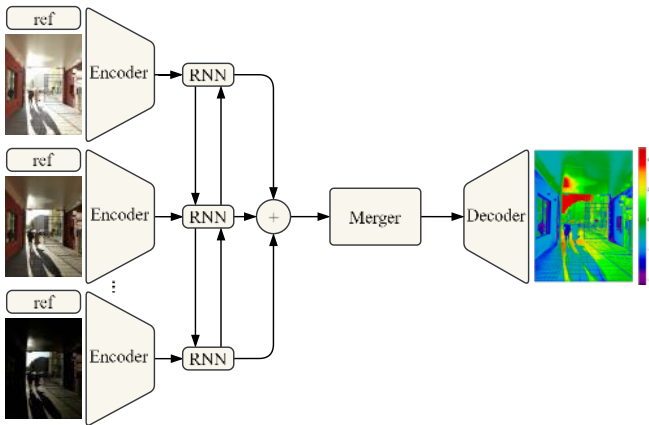


Рис. 4. Архитектура рекуррентных сетей применительно к задаче HDR, предложенная Prabhakar et al. [11]. Вход каждого кодировщика – это пара из кадра с определённой экспозицией и опорного снимка (**ref**), при этом количество входных изображений не фиксировано. RNN принимает сжатые данные и «помня» опорный снимок обрабатывает экспозиции так, чтобы они соответствовали **ref**. Все блоки RNN последовательно связаны между собой прямой и противоположной связью

IV. GAN В ЗАДАЧЕ HDR

Генеративно-сопоставительные сети (GAN) представляют собой обширный класс алгоритмов глубокого обучения, который нашел применение и в задачах обработки изображений, включая задачи HDR.

GAN – так же, как и RNN, – могут использоваться для задачи Single-exposure HDR. При этом, одна сеть создаёт изображение, а другая – судит о корректности. Это достигается путем обучения генеративной сети на парах, состоящих из LDR и HDR изображений. Генеративная сеть пытается создать изображение с широким динамическим диапазоном из единственного входного LDR с использованием контекста HDR изображений.

Но GAN используются и в традиционном расширении динамического диапазона. Авторы статьи [12] предложили подход, в котором GAN обучается на непарных данных. На вход такой архитектуре подаётся пакет выровненных снимков различной экспозиции, но выход, который сеть получает для оценки своего результата во время тренировки, не сопоставим с входными данными. Точнее, не полностью сопоставим – это может быть та же сцена с теми же условиями освещения, но с иным ракурсом (или с иным положением объектов). Архитектура состоит из следующих блоков (рис. 5):

- **Генератор (Generator).** Создаёт HDR изображения из входов (похожим образом, как на рис. 3).
- **Дискриминатор (Discriminator).** Для обучения модели используется дискриминатор, который различает сгенерированные и реальные изображения HDR. Дискриминатор обучается классифицировать изображения как реальные, созданные генератором или размытые.
- **Мини-патч модуль (Min-Patch Module).** Этот модуль внедрен для борьбы с артефактами «призраков» на изображениях HDR. Мини-патч-модуль концентрируется на аномальных участках сгенерированных изображений и помогает избежать появления нежелательных артефактов в итоговых изображениях.

V. ОБУЧЕНИЕ БЕЗ УЧИТЕЛЯ В ЗАДАЧЕ HDR

В задачах создания HDR изображений, не имеющих точных парных данных для обучения, подходы обучения без учителя представляют собой мощный инструмент для автоматического извлечения и использования информации из непарных изображений различных экспозиций.

Чаще всего, алгоритмы без учителя реализуют ветвь объединения изображения на основе средневзвешенного (рис. 2). При этом нейросеть нужна, чтобы гибко определить вес каждого участка каждого снимка.

Авторы работы MEFLUT [13] предлагают строить одномерные таблицы поиска (lookup table), которые позволяют объединить информацию о яркости и контрасте различных экспозиций. При этом архитектура решения осуществляет следующие шаги (рис. 6):

- Извлечение гистограмм яркости каждого снимка.
- Нормализация гистограмм.
- Построение таблиц поиска веса в зависимости от значения пикселя для каждой из экспозиций.
- Подсчёт средневзвешенного на основе построенных таблиц поиска.

Преимущество такого подхода в том, что нейросеть нужна, только чтобы получить таблицы преобразования (натренировать их), но после окончания тренировки можно пользоваться лишь таблицами, что даёт ощутимый прирост к скорости вычислений.

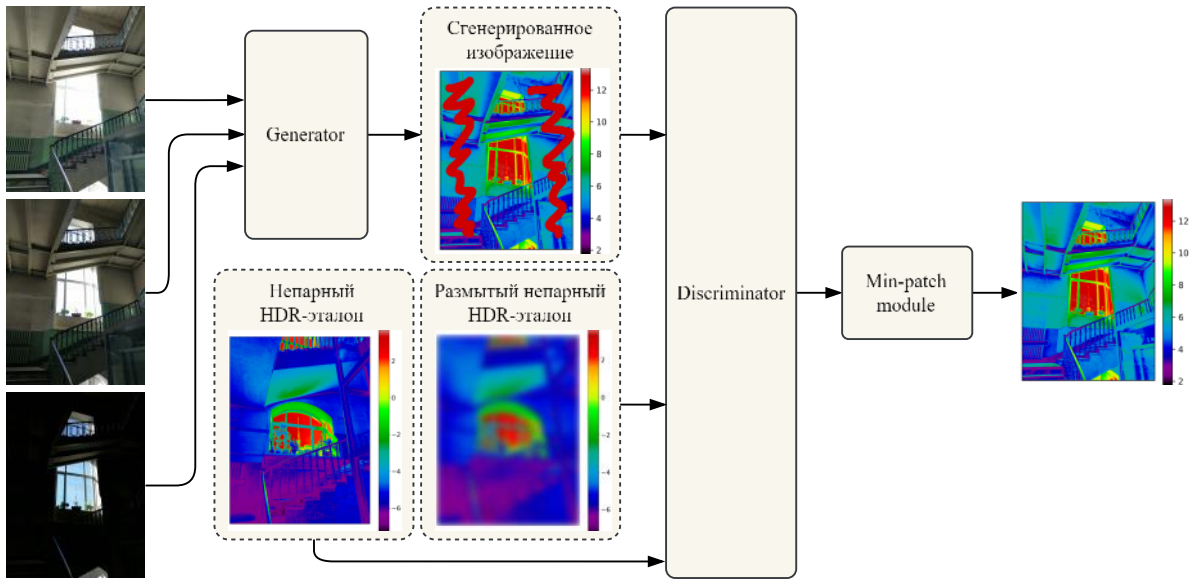


Рис. 5. Архитектура состязательной сети [12]. Предлагаемый метод направлен на получение качественных результатов HDR с непарными наборами данных. **Генератор** извлекает признаки из входных данных с мультиэкспозицией и генерирует HDR (архитектура сопоставима с рис. 3). **Дискриминатор** поочередно различает сгенерированное и размытое изображения с непарным эталоном. **Модуль min-patch** концентрируется на «странной» части ненастоящих (сгенерированного и размытого) изображений по сравнению с эталоном и помогает избежать призраков

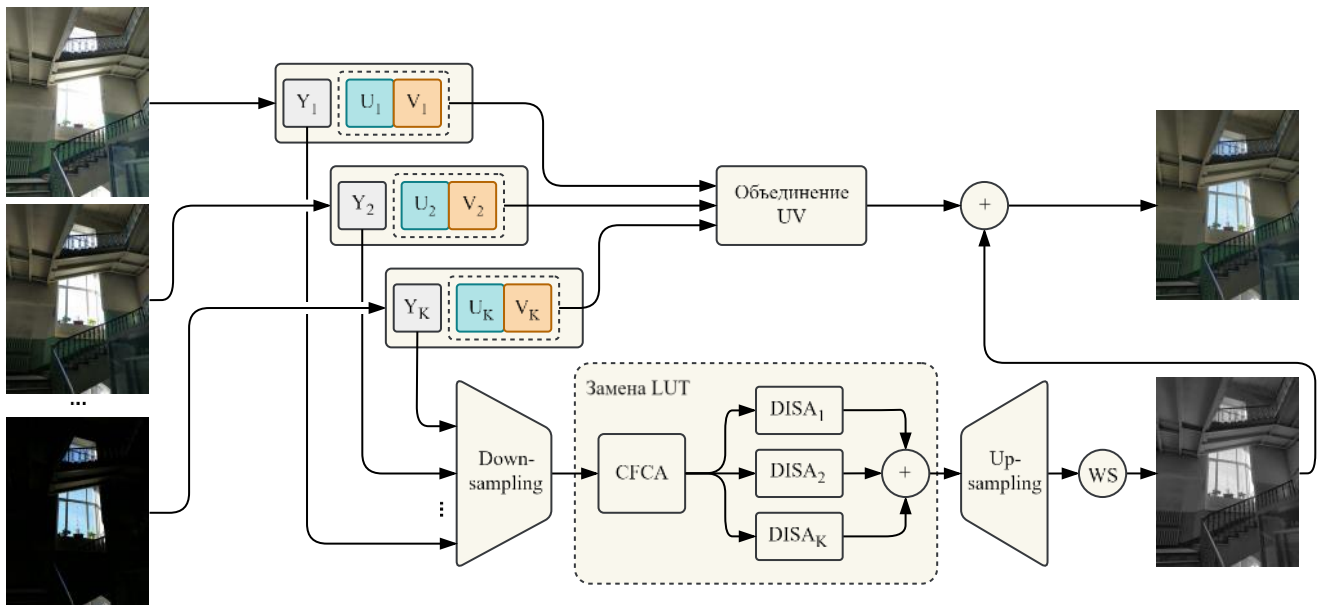


Рис. 6. Упрощённая архитектура сети обучения без учителя [13]. Ключевой блок – это блок тренировки таблиц поиска (**замена LUT**). Внутри имеются следующие модули: свёртка с кадровым и каналным вниманием (**CFCA**), которая определяет, какие пиксели и каналы наиболее полезны; расширенная свёртка с пространственным вниманием (**DISA**). Анализ полезных участков производится лишь по Y-каналу YUV-пространства. После повышения разрешения вычисляется взвешенная сумма (**WS**) всех экспозиций в соответствии с найденными таблицами поиска и результат конкатенируется с объединёнными слоями UV

VI. АРХИТЕКТУРЫ МАШИННОГО ОБУЧЕНИЯ В ЗАДАЧЕ ВЫРАВНИВАНИЯ ИЗОБРАЖЕНИЙ

Задача выравнивания имеет вспомогательную функцию в создании HDR изображений, тем не менее, без этого шага итоговый результат может быть сильно испорчен (рис. 2). Основная проблема в выравнивании снимков в задаче HDR заключается в чувствительности методов к различному уровню экспозиции – алгоритмам сложно сопоставить опорные точки, так как на снимках с повышенным или пониженным количеством света точка может быть засвечена или затенена. Поэтому алгоритмы,

основанные на составлении дескрипторов, инвариантных к поворотам (но не инвариантных к экспозиции) [14], [15], в чистом виде не могут быть использованы. Ещё одной причиной некорректного результата классических алгоритмов могут быть крупные передвижения объектов между кадрами брекетинга.

Современный подход к выравниванию изображений подразумевает использование вспомогательных нейронных сетей в связке с классическими алгоритмами. Это могут быть модели, основанные на:

- оптическом потоке ([16], [17]);

- поиске корреляции между изображениями через карты внимания (attention maps [18]) ([19]);
- итеративном пирамидальном выравнивании [20].

Выравнивание с помощью таких моделей происходит дольше, чем по сравнению с классическими методами, но выдаёт лучший результат.

VII. РЕКОМЕНДАЦИИ

Суммируя всё вышесказанное можно выделить следующие рекомендации по использованию многообразия нейронных сетей в задачах HDR:

- Для борьбы с артефактами «призраков» подходят простые CNN и RNN решения.
- «Ореолы» можно убрать с помощью CNN на этапе тональной компрессии.
- Если имеется единственный кадр, а расширить динамический диапазон всё же нужно (или же просто требуется повысить качество изображения) подходят Single-exposure методы на основе RNN и GAN.
- Для случая, когда нет эталонов и/или важен баланс скорости и качества, лучше использовать LUT методы на основе обучения без учителя. По окончании обучения можно больше не использовать громоздкую нейросеть и оперировать лишь легковесной таблицей преобразований.
- Для выравнивания целесообразно использовать нейросети, в случае если брекетинг осуществлялся без штатива и/или имеются крупные перемещения объектов между экспозициями. Иначе подойдут классические методы – их качество будет достаточным, а скорость высокой.

Недостатком всех решений с использованием нейронных сетей является их медлительность по сравнению с алгоритмическими решениями: тренировка занимает несколько часов (а иногда и дней [12]) на современных видеокартах, а извлечение при размере картинки 4096x2160 занимает от 10G до 10T FLOPs [21], в то время как алгоритмам не нужна тренировка, а извлекаются они быстрее.

Проблемой всей области HDR (а значит, и проблемой методов с использованием нейросетей) является субъективность конечного результата – особенно это касается тональной компрессии. Такая ситуация обусловлена тем, что объективных метрик «естественности» или «реалистичности» нет – можно лишь оценить количество сохранённых деталей и скорость вычислений [22]. Поэтому эталоны многих датасетов ([16], [23], [24]) часто выглядят неестественно. Но всё же есть и такие ([13], [25]), которые выглядят более натурально (субъективно, конечно) – это можно объяснить тем, что эти датасеты используют Raw-изображения (которые имеют 12–16 бит на канал), снятые на профессиональные фотоаппараты. То есть эти изображения не сгенерированы из брекетинга – это буквально слепок яркости сцены (хотя, конечно, сцены больше, чем 10–14 EV даже такие камеры пока не могут

снять). Также, тональную компрессию этих изображений производила группа профессиональных фотографов, что повышает объективность результата. Проблема в том, что эти датасеты не содержат серии брекетинга, из которого можно было бы пытаться восстановить карту яркости, и, соответственно, нет движений между кадрами (а значит, эти датасеты не получится использовать для тренировки сетей, нацеленных на устранение «призраков»). В этом смысле, решения с обучением без учителя выглядят ещё более привлекательными.

VIII. ЗАКЛЮЧЕНИЕ

В итоге обзора современных подходов к решению задач съёмки и отображения кадров с широким динамическим диапазоном были выделены основные используемые типы архитектур машинного обучения и описаны сферы их применения в аспекте различных задач устранения артефактов HDR. На основе обзора были сформулированы рекомендации по реализации различных типов архитектур машинного обучения для конкретных случаев использования: Для борьбы с артефактами «призраков» подходят простые CNN и RNN решения, ореолы убираются с помощью CNN в тональной компрессии, Single-exposure методы на основе RNN и GAN подходят для увеличения динамического диапазона при единственном кадре, LUT методы на основе обучения без учителя обеспечивают баланс скорости и качества, и для выравнивания нейросети подходят при отсутствии штатива или крупных перемещений объектов, иначе – классические методы.

Дальнейшие исследования предполагают продолжение изучения архитектур нейросетей применительно к задачам HDR с упором на обучение без учителя. Также необходимы исследования с целью нахождения более объективных метрик, чем есть сейчас.

СПИСОК ЛИТЕРАТУРЫ

- [1] Reinhard E., Heidrich W., Debevec P., Pattanaik S., Ward G., Myszkowski K. High Dynamic Range Imaging, Second Edition Acquisition, Display, and Image-Based Lighting (Визуализация широкого динамического диапазона). Amsterdam: Morgan Kaufmann Publishers, 2010. 673 p.
- [2] Evangelidis G.D., Psarakis E.Z. Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization (Параметрическое выравнивание изображений с использованием максимизации повышенного коэффициента корреляции) // IEEE Transactions on pattern analysis and machine intelligence, 30(10) – oct. 2008
- [3] Mertens T., Kautz J., Van Reeth F. Exposure fusion (Объединение экспозиций) – 15th Pacific Conference on Computer Graphics and Applications (PG'07). IEEE, 2007. С. 382-390.
- [4] Debevec P., Malik J. Recovering High Dynamic Range Radiance Maps from Photographs (Восстановление карт яркости с широким динамическим диапазоном из фотографий) // Proc. of ACM SIGGRAPH 1997, 1997. p. 369–378.
- [5] Mitsunaga T., Nayar S.K. Radiometric self calibration (Самокалибровка яркости). Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149) // IEEE, 1999. Т. 1. p. 374-380.
- [6] Reinhard E., Stark M., Shirley P., Ferwerda J. Photographic tone reproduction for digital images (Фотографическое воспроизведение тона для цифровых изображений) // ACM transactions on graphics (TOG) 21.3, 2002. p. 267–276.
- [7] Durand F., Dorsey J. Interactive tone mapping (Итеративная тональная компрессия) // Rendering Techniques 2000: Proceedings

- of the Eurographics Workshop in Brno, Czech Republic, June 26–28, 2000 11. Springer Vienna, 2000. С. 219-230.
- [8] Wu S., Xu J., Tai Y., Tang C. Deep high dynamic range imaging with large foreground motions (DeepHDR для крупных движений на переднем плане) // Proceedings of the European Conference on Computer Vision (ECCV). 2018. С. 117-132.
- [9] Zhang F., Tian M., Li Z., Xu B., Lu Q., Gao C., Sang N. Lookup Table meets Local Laplacian Filter: Pyramid Reconstruction Network for Tone Mapping (Таблица поиска соответствует локальному фильтру Лапласа: пирамидальная сеть восстановления для тональной компрессии) // Advances in Neural Information Processing Systems. 2024. Т. 36.
- [10] Kim J., Lee S., Kang S. J. End-to-end differentiable learning to hdr image synthesis for multi-exposure images (Комплексное дифференцированное обучение синтезу HDR-изображений для кадров с мультиэкспозицией) // Proceedings of the AAAI Conference on Artificial Intelligence. 2021. Т. 35. №. 2. С. 1780-1788.
- [11] Prabhakar K.R., Agrawal S., Babu R.V. Self-gated memory recurrent network for efficient scalable HDR deghosting (Саморегулируемая рекуррентная сеть памяти для эффективного масштабируемого устранения призраков HDR) // IEEE Transactions on Computational Imaging. 2021. Т. 7. С. 1228-1239.
- [12] Li R., Wang C., Wang L., Liu G., Zhang H., Liu S. UPHDR-GAN: Generative adversarial network for high dynamic range imaging with unpaired data (UPHDR-GAN: Генеративно-сопоставительная сеть для отображения высокого динамического диапазона с непарными данными) // IEEE Transactions on Circuits and Systems for Video Technology. 2022. Т. 32. №. 11. С. 7532-7546.
- [13] Jiang T. et al. MEFLUT: Unsupervised 1D lookup tables for multi-exposure image fusion (MEFLUT: одномерные таблицы поиска для объединения кадров с мультиэкспозицией с обучением без учителя) // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. С. 10542-10551.
- [14] Lowe G. Sift-the scale invariant feature transform (SIFT – Масштабно-инвариантная трансформация признаков) // Int. J. 2004. Т. 2. №. 91-110. С. 2.
- [15] Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features (SURF: ускоренные устойчивые признаки) // Computer Vision-ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9. Springer Berlin Heidelberg, 2006. С. 404-417.
- [16] Kalantari N. K., Ramamoorthi R. Deep high dynamic range imaging of dynamic scenes (Глубокое обучение в отображении широкого динамического диапазона контрастных сцен) // ACM Trans. Graph. 2017. Т. 36. №. 4. С. 144:1-144:12.
- [17] Catley-Chandar S., Tanay T., Vandroux L., Leonardis A., Slabaugh G., Pérez-Pellitero E. FlexHDR: Modeling alignment and exposure uncertainties for flexible hdr imaging (FlexHDR: моделирование огрехов выравнивания и экспозиции для гибкой визуализации HDR) // IEEE Transactions on Image Processing. 2022. Т. 31. С. 5923-5935.
- [18] Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. Attention is all you need (Внимание – это всё, что вам нужно) // Advances in neural information processing systems. 2017. Т. 30.
- [19] Yan Q., Gong D., Shi Q., Hengel A. V. D., Shen C., Reid I., Zhang Y. Attention-guided network for ghost-free high dynamic range imaging (Сеть, управляемая вниманием, для получения HDR изображений без «призраков») // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019. С. 1751-1760.
- [20] Pu Z., Guo P., Asif M. S., Ma Z. Robust high dynamic range (HDR) imaging with complex motion and parallax (Высокоскоростное составление HDR изображений со сложным движением и параллаксом) // Proceedings of the Asian Conference on Computer Vision. 2020.
- [21] Yang Q., Liu Y., Yang J. Efficient HDR reconstruction from real-world RAW images (Эффективное восстановление HDR на основе реальных RAW-изображений) // arXiv preprint arXiv:2306.10311. 2023.
- [22] Wang L., Yoon K. J. Deep learning for HDR imaging: State-of-the-art and future trends (Глубокое обучение для обработки изображений HDR: современное состояние и тенденции) // IEEE transactions on pattern analysis and machine intelligence. 2021. Т. 44. №. 12. С. 8874-8895.
- [23] Prabhakar K.R., Arora R., Swaminathan A., K. Singh P., Babu R. V. A fast, scalable, and reliable deghosting method for extreme exposure fusion (Быстрый, масштабируемый и надежный метод устранения призраков для слияния экстремальных воздействий) // 2019 IEEE International Conference on Computational Photography (ICCP). IEEE, 2019. С. 1-8.
- [24] Cai J., Gu S., Zhang L. Learning a deep single image contrast enhancer from multi-exposure images (Усилитель контраста одного изображения на основе глубокого обучения на серии мультиэкспозиции) // IEEE Transactions on Image Processing. 2018. Т. 27. №. 4. С. 2049-2062.
- [25] Vychkovsky V., Paris S., Chan E., Durand F. Learning photographic global tonal adjustment with a database of input / output image pairs (Обучение фотографической глобальной тональной компрессии с использованием датасета пар входных/выходных изображений) // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2011.