

*Государственное образовательное учреждение высшего профессионального  
образования*

**«Московский государственный технический университет  
имени Н. Э. Баумана»  
(МГТУ им. Н.Э. Баумана)**

---

ФАКУЛЬТЕТ                      «Информатика и системы управления»  
КАФЕДРА                      «Программное обеспечение ЭВМ и информационные технологии»

**РАСЧЁТНО - ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**  
**к курсовой работе на тему:**

Разработка программы записыванию видеозаписи в HDR качестве

Студент	_____	Киселев А.М.
	(Подпись, дата)	
Руководитель курсового проекта	_____	Оленев А.А.
	(Подпись, дата)	

Москва 2018

## Содержание

Введение . . . . .	3
1 Аналитический раздел . . . . .	4
1.1 Различия LDR и HDR изображений . . . . .	4
1.2 Процесс получения HDR изображения . . . . .	5
1.2.1 Конвейер получения HDR изображения на камере со сла- бым сенсором . . . . .	6
1.2.2 Получение HDR видео на камере со слабым сенсором в режиме реального времени . . . . .	6
1.2.2.1 Способы получения разной экспозиции в видео . . .	6
1.3 Выравнивание изображений . . . . .	7
1.3.1 Алгоритм МТВ(Median Threshold Bitmap) . . . . .	7
1.4 Удаление артефактов изображения . . . . .	10
1.4.1 Алгоритм BMD(Bitmap Movement Detection) . . . . .	10
2 Конструкторский раздел . . . . .	13

## Введение

В 20 веке был принят стандарт CIE, который охарактеризовывал то, как должна отображаться картинка на восьмибитных мониторах. Технологии того времени позволяли передавать только сильно ограниченный диапазон цветовой информации, поэтому картинка, предоставляемая монитором, была далека от более сочных и ярких цветов, которые человек видит в повседневной жизни. Такое явление может быть охарактеризовано следующими терминами: LDR(low dynamic range) или SDR(standart dynamic range) image – маленький динамический диапазон или стандартный динамический диапазон изображения. Мониторы LDR и SDR не могли передавать широкий спектр цветов. К тому же не все сенсоры цифровых камер могут позволить себе захватить широкий диапазон оттенков. Из-за этого возникает проблема засвеченных и затемненных участков на фотографии или видео.

С развитием технологий, стали появляться так называемые HDR мониторы, которые позволяют передавать цветовую информацию с глубиной в 10 битов. По причине того, что LDR и SDR мониторы обладают маленькой глубиной цвета, они не в состоянии корректно отображать картинку с широким диапазоном цветов.

Из-за того, что не все сенсоры могут запечатлеть широкий спектр цветов, были придуманы специальные алгоритмы и методы получения HDR изображений с обычных и самых распространенных сенсоров(цифровые камеры, мобильные телефоны, планшеты, веб-камеры и т.п.). Полученные HDR снимки или видео с камер, сенсор которых рассчитан на LDR и SDR изображения, могут быть отображены на HDR мониторах.

Получение HDR кадра - нетривиальная задача, которая делится на несколько этапов: получения серии снимков с разной длиной экспозиции, выравнивание кадров по отношению друг к другу, реконструкция и удаление движущихся объектов полученных изображений, объединение изображений(merging), проведение цветовой коррекции изображения.

## 1 Аналитический раздел

*SDR(Standart Dynamic Range) изображение* – изображение, пиксели которого содержат цвета и яркость, соответствующую глубине монитора.

*LDR(Low Dynamic Range) изображение* – изображение, пиксели которого хранят ограниченный диапазон цветов и яркости, предназначенное для отображении на старых мониторах.

*HDR(High Dynamic Range) изображение* – изображение, пиксели которого содержат более широкие значения цвета и яркости в сравнении с изображениями стандартного диапазона(SDR).

*Глубина цвета* – количество бит, приходящихся на один пиксель

### 1.1 Различия LDR и HDR изображений

Несмотря на то, что технологии за последние несколько лет быстро развиваются и качество полученных кадров и устройств, их отображающих, увеличивается, получение картинки, сопоставимой с реальным окружением, остается нелегкой задачей. Яркость, диапазон цветов, которые видит человек в повседневной жизни, невозможно отобразить на большинстве мониторов, используемых во всем мире.

Хотя уже начинают появляться так называемые HDR мониторы и существуют камеры, сенсоры которых позволяют запечатлеть широкий спектр цветов, яркостей и деталей, цена таких устройств может достигать огромных значений, поэтому большинство мониторов остаются SDR формата и не в состоянии передать картинку с большим количеством цветовой информации.

Глубина цвета на большинстве мониторов составляет 8 бит. Глубина цвета самого распространенного формата изображений JPEG так же составляет 8 бит, в котором используется цветовое пространство  $Y C_r C_b$ . Это цветовое пространство позволяет использовать лишь малую часть видимых человеком цветов. Так же это цветовое пространство не способно передать большую часть яркостей, которых способен распознать человеческий глаз.

В противовес формату JPEG существует так называемый RAW формат. В отличие от JPEG он способен содержать гораздо больше информации. Глубина цвета такого формата может достигать 12-16 бит(Значение может варьироваться в зависимости от возможностей сенсора камеры), который доступен в большинстве современных камер. Чаще всего этот формат автоматически переводят в JPEG во время съемки, что приводит к потере многих деталей без возможности восстановления. Однако, используя специальные инструменты RAW изображение в дальнейшем можно преобразовать в LDR изображение, получив на выходе кадр без потери нужной информации(провести tonemapping).

На цветовом пространстве 1.1 наглядно показан охват возможных отображаемых цветов в RGB пространстве или на тех же LDR мониторах и в HDR с глубиной цвета хотя бы в 12 бит.

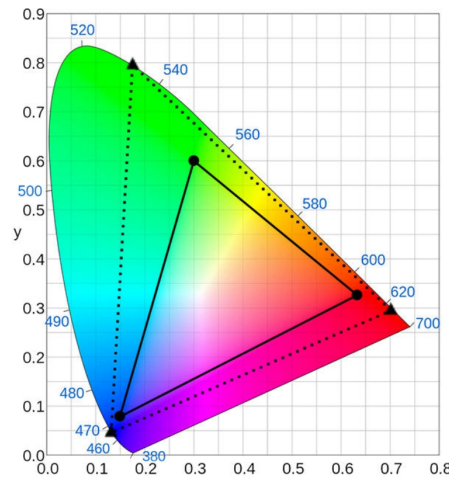


Рисунок 1.1 — Охват видимых человеком цветов цветовым пространством RGB и HDR.

Главное различие HDR и традиционных LDR изображений в том, что HDR картинка не зависит от какого-либо устройства и содержит максимально большое количество информации. Количество отображаемой информации на устройстве уже зависит непосредственно от его собственных характеристик и возможностей. LDR изображение же зависит от устройства, на котором оно будет отображаться и спроектировано специально под мониторы или экраны, которые отображают определенное количество цветовой информации.

## 1.2 Процесс получения HDR изображения

HDR изображение можно получить несколькими способами:

- с помощью объединения снимков с разной длинной экспозиции,
- с помощью камеры, сенсор которой позволяет захватить широкий объем данных,
- при помощи перевода LDR изображения в HDR специальными алгоритмами

Первый метод является более распространенным, так как устройства, которые больше всего распространены в повседневной жизни (телефоны, планшеты, веб-камеры) не обладают достаточно мощными сенсорами, для того, чтобы захватить широкий диапазон цветов. Последний метод не получил распространения, потому что задача перевода LDR или SDR изображения в HDR возможна только при помощи применения алгоритмов реконструкции, завязанных на нейронных сетях, появившихся достаточно недавно.

### 1.2.1 Конвейер получения HDR изображения на камере со слабым сенсором

- Выравнивание кадров,
- Реконструкция и удаление движущихся объектов(артефактов),
- Слияние кадров с разной экспозицией,
- Преобразование HDR к LDR(tonemapping) если это требуется

Наглядно процесс получения HDR кадра с возможностью перевода его в LDR показан на схеме 1.2

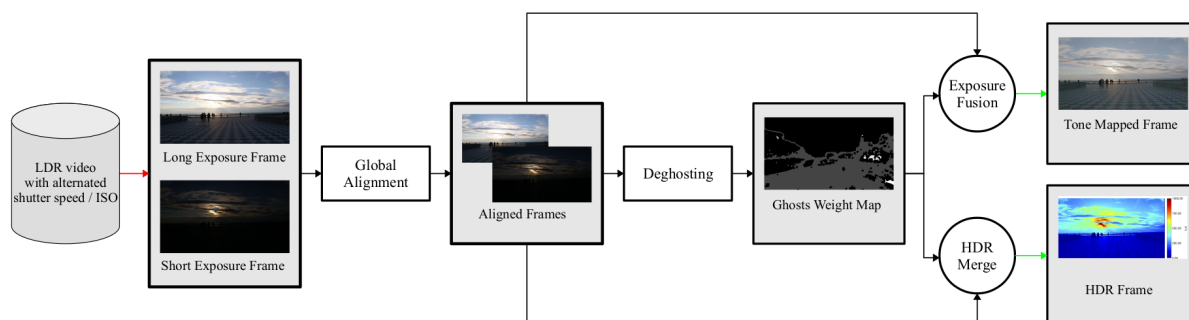


Рисунок 1.2 — Процесс получения HDR изображения путем слияния кадров с разной экспозицией

### 1.2.2 Получение HDR видео на камере со слабым сенсором в режиме реального времени

Данная задача подразумевает собой быстрое создание HDR изображений из 2-3 кадров с изменяющейся экспозицией, которые получаются, путем изменения экспозиции через каждый последующий кадр. Таким образом, можно быстро получить два кадра, которые будут содержать всю информацию о самых темных и самых светлых участках сцены. Так, как эти кадры следуют друг за другом, различия в них, предполагается, будут незначительными.

При записи видео в таком формате очень важна скорость алгоритмов, использующихся при получении HDR кадра. Операции выравнивания, удаления артефактов, слияния кадров и преобразование полученного кадра в LDR должны протекать быстро.

#### 1.2.2.1 Способы получения разной экспозиции в видео

Получение разной экспозиции кадров можно добиться несколькими способами:

- а) Использовать несколько камер, каждая из которых будет вести запись с соответствующей экспозицией;

б) Использовать камеру с несколькими сенсорами, каждый из которых будет вести запись с соответствующей экспозицией;

в) Использовать одну камеру, меняя параметр выдержки или ISO, которые влияют на полученную экспозицию;

Первые два способа позволяют сохранить количество кадров в секунду и улучшить качество получаемой картинки, но они требуют предварительной калибровки. Так же несколько камер и камеры с дополнительными сенсорами не самые распространенные решения в телефонах и других гаджетах.

Третий способ позволяет получать кадры с разной экспозицией на любом устройстве, в котором есть камера, что делает его более универсальным, но это снижает количество кадров в секунду почти в два раза.

### 1.3 Выравнивание изображений

Первым шагом в получении HDR изображения является глобальное выравнивание последовательности кадров с разной экспозицией. Предполагается, что 2 соседних кадра не меняются слишком быстро.

Но при съемке с рук или при съемке в движении, кадры все равно меняются, положение записывающего устройства тоже может быть нестабильным. В дальнейших шагах получения HDR снимка очень важно, чтобы кадры с разной экспозицией были максимально выровнены по отношению друг к другу.

Задача выравнивания усложняется тем, что экспозиции изображений сильно разнятся. Для решения этой задачи используется алгоритм MTB (Median Threshold Bitmap) пороговое отсечение по медианному значению.

#### 1.3.1 Алгоритм MTB (Median Threshold Bitmap)

На вход алгоритму подается N 8-битных монохромных изображений, которые можно получить путем использования зеленого канала изображения, либо путем перевода каждого пикселя 24-битного sRGB изображения в серый цвет с помощью выражения 1.1

$$grey = \frac{(54 * red + 183 * green + 19 * blue)}{256} \quad (1.1)$$

Одно из изображений выбирается базовым. Остальные N-1 при выравнивании будут опираться на базовое.

Далее требуется для каждого из изображений:

а) Найти медианное значение (8-битное значение) из гистограммы низкого разрешения пикселей монохромного изображения.

б) Создать растровое изображения, в котором 0 будут отмечаться пиксели, которые меньше или равны медианному значению и 1 будут отмечаться пиксели, которые строго больше этого медианного значения.

Результаты этих шагов представлены на примере двух изображений с разной экспозицией 1.3

После этого каждое N-1 изображение подгоняется под базовое. Вначале они сравниваются. Сравнение происходит путем побитовой операции XOR – каждый пиксель изображения высчитывается с каждым соответствующим пикселем базового изображения операцией XOR. Результаты сохраняются и на выходе получается очередное растровое изображение, в котором 0 – пиксели, информация которых совпала с базовым изображением и 1 – пиксели, информация которых не совпала.

Для устранения шумов, которые могут получиться при съемке в темных, плохо освещенных помещениях, используется *исключающее растровое изображение*. Т.к. шумы возникают близко к медианному значению, пиксели, лежащие на расстоянии ближе чем 4 пикселя к медианному значению не учитываются.

Так генерируется исключающее растровое изображение, в котором 0 – отмечается пиксель, лежащий близко к медианному значению, иначе 1.

Для выравнивания находится значение сдвига (x, y): x – смещение изображения по горизонтали, y – сдвиг изображения по вертикали. Самый оптимальный сдвиг будет тот, у которого при сравнении изображения с базовым, количество пикселей 1 будет минимальное. Искать его можно несколькими способами. Самый очевидный – простой перебор всевозможных сдвигов (x, y), но используется более оптимальный метод. С помощью пирамидальной сегментации изображения можно добиться наименьшего количества сравнений двух растровых побитово размеченных изображений.

Каждое монохромное изображение переводится на следующий уровень путем уменьшения его разрешения в два раза. Таким образом получаем несколько уровней, которые содержат монохромные изображения. Начиная с самого последнего уровня (изображения которых содержат наименьшее разрешение) нужно найти для них исключающее растровое изображение по медианному значению, найти для них растровое изображение по медианному значению и с помощью операции AND получить растровое изображение, которое должно будет содержать минимальное число 1. Далее найти сдвиг перебором по значениям  $\pm 1$  пиксель в каждом направлении (по горизонтали и вертикали)

На следующем уровне "пирамиды" нужно домножить полученный сдвиг на 2 (значение, на которое менялось разрешение на каждом уровне пирамиды) и снова вычислить наилучший сдвиг на расстоянии  $\pm 1$  пиксель от текущего сдвига в разные направления x и y.





Рисунок 1.3 — Получение растровых изображений, использующихся для дальнейшего сравнения(справа), оригиналы изображений(слева)

Эти действия продолжаются до последнего(изначального) уровня пирамиды, на котором получается окончательный требуемый сдвиг  $(x, y)$  для изображения, чтобы быть выравненным по отношению к базовому.

## 1.4 Удаление артефактов изображения

При записи видео с камеры, объект, который ведет съемку, или коружение вокруг него могут находиться в движении. По этой причине могут появляться так называемые "призраки". На одном из этапов создания HDR изображения полученные кадры сливаются. Но если предварительно не позаботиться о движущихся объектах, на результирующем изображении можно наблюдать подобные артефакты, пример на изображении 1.4



Рисунок 1.4 — Изображение с удаленными артефактами движущегося объекта(справа), артефакты на необработанном изображении(слева)

Во время съемки видео такие артефакты могут значительно повлиять на отображаемую информацию исходного продукта. Чтобы этого не допустить требуется позаботиться об удалении избыточной информации перед слиянием кадров с разной экспозицией.

### 1.4.1 Алгоритм BMD(Bitmap Movement Detection)

Алгоритм BMD основывается на использовании МТВ растрового изображения. Использование побитовых операций позволяет добиться быстрой скорости и точности вычислений. В качестве результата алгоритм предоставляет *Размеченную матрицу движений*. С помощью этой матрицы при склеивании кадров с разной экспозиции можно будет удалить артефакты.

Первым делом требуется найти для каждого изображения МТВ растровое изображение. Пусть  $B_i$  – массив этих растровых изображений. На неподвижной, статичной сцене предполагается, что каждый пиксель сохраняет свое значение на всех растровых изображениях, хранящихся в  $B_i$ . Очевидно, что если значение пикселя поменяется, в области пикселя произошло движение.

Суммируя все растровые изображения из  $B_i$  между собой, получается  $M^*$ , в которой любые значения кроме 0 или N(количество исходных изображений) рассматриваются как совершенное движение в районе пикселя.

$M^*$  может содержать некоторое количество шумов, которые могут повлечь за собой неприятные последствия. Используя sequence of morphological dilation and erosion,  $M^*$  переводится в  $M$  – матрицу движений 1.7.

Затем матрица движений переводится в кластерное отображение, которое получается с помощью Connected Component labeling, где у каждого кластера есть своя метка. Отсюда получается размеченная карта движений  $L_M$  с размеченными

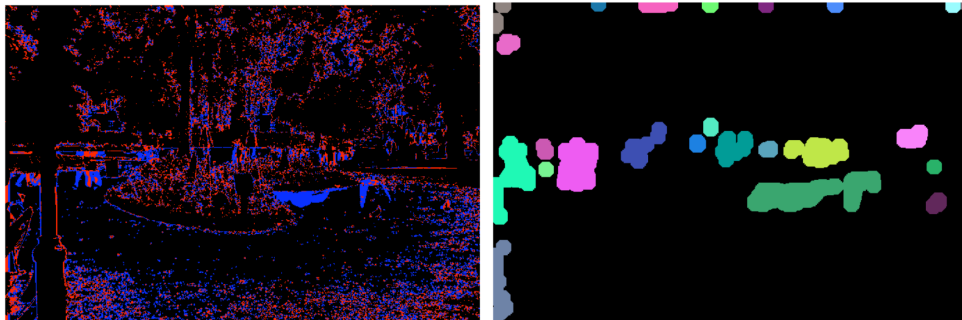


Рисунок 1.5 — Представление матрицы движений(справа), и матрицы движений до преобразований(из-за большого количества шума сложно понять где и что передвигается).



Рисунок 1.6 —  
Исходная  
последовательно  
изображений с  
разными  
экспозициями.

Рисунок 1.7 — Пример нахождения движений на последовательности изображений с разными экспозициями.

кластерными зонами  $\Omega_i$ , которые содержат пиксели, влияющие на появление артефактов.

## 2 Конструкторский раздел

check