



**VisionLabs**  
MACHINES CAN SEE

## VISIONLABS FACE STREAM

Описание функциональных характеристик программного обеспечения и информация, необходимая для установки и эксплуатации ПО

ООО «ВижнЛабс»

123458, г. Москва, ул. Твардовского д. 8, стр. 1

 +7 (499) 399 3361

 [info@visionlabs.ru](mailto:info@visionlabs.ru)

 [www.visionlabs.ru](http://www.visionlabs.ru)

## Содержание

Глоссарий .....	3
Введение .....	4
Аппаратные и программные требования.....	5
Минимальные аппаратные требования.....	5
Программные требования.....	6
Описание работы Face Stream.....	7
Режимы работы .....	7
Общие рекомендации по настройке Face Stream.....	8
Рекомендации перед началом настройки.....	8
Настройка производительности Face Stream.....	9
Определение области с движением .....	12
Обработка пакетов кадров.....	12
Общая информация о настройке .....	13
Отправка фотоизображений в LUNA PLATFORM.....	16
Дополнительная информация .....	18
Форматы видео .....	18
Стандарты сжатия видео .....	18
Протоколы.....	18

## Глоссарий

Термин	Значение термина
Трек	Информация о положении объекта (лица) одного человека на последовательности кадров. Если объект покидает зону кадра, то трек прерывается не сразу. Некоторое время он ожидает возвращения объекта в кадр. Если объект вернулся, то трек продолжается.
Трекинг	Функция отслеживания объекта (лица) на последовательности кадров.
Детекция	Обнаружение лица в кадре.
Ракурс	Степень поворота головы (в градусах) по каждой из трех осей вращения (наклон вверх/вниз относительно горизонтальной оси; наклон влево/вправо относительно вертикальной оси; поворот относительно вертикальной оси).
Лучший кадр, лучшая детекция	Лучший кадр выбирается из всех кадров одного трека. Основными условиями выбора лучшего кадра являются приемлемое качество изображения и наличие на нём лица с наилучшим ракурсом. Условия выбора лучшего кадра задаются в настройках Face Stream.
Портрет	Фрагмент изображения лица с кадра, отобранного в соответствии с настройками алгоритма и максимально приближенный к требованиям ГОСТ 19794-5-2006 / ISO IEC 19794-5 2005(E).
Нормализованное изображение (warp)	Специальный формат изображения для работы с LUNA PLATFORM 5 с выровненным по горизонтали лицом. Такое изображение содержит всю необходимую информацию о лице и при этом быстро обрабатывается системой.
ROI	Зона интереса на фотоизображении
LUNA PLATFORM Enterprise, LPE	ПО VisionLabs LUNA PLATFORM Enterprise включено в Единый реестр российских программ для электронных вычислительных машин и баз данных Запись в реестре №5878 от 20.09.2019 Доступны версии 3 и 4 LPE.
LUNA PLATFORM 5, LP 5	ПО VisionLabs LUNA PLATFORM 5

## Введение

Данный документ описывает функциональные характеристики ПО, и информацию, необходимую для установки и эксплуатации **VisionLabs Face Stream**

Приложение Face Stream осуществляет несколько функций:

- **Чтение видеопотоков**

В качестве источников данных могут выступать веб-камеры, IP-камеры (посредством протокола RTSP) и видеофайлы.

- **Обработка видеопотоков**

Выполняется поиск и сопровождение лиц в видеопотоке, пока они не покинут кадр или не будут перекрыты.

- **Отсылка изображений лиц в виде HTTP-запросов во внешние сервисы**

В качестве внешних сервисов может выступать ПО компании VisionLabs (LUNA PLATFORM ENTERPRISE и LUNA PLATFORM 5).

# Аппаратные и программные требования

## Минимальные аппаратные требования

Дальнейшие минимальные требования приведены для использования одного экземпляра Face Stream.

Для корректной работы приложения аппаратное обеспечение должно отвечать следующим минимальным требованиям:

- CPU с частотой 2 ГГц и выше;
- 4 Гб оперативной памяти и выше;
- 10 Гб свободного места на жестком диске.
- Доступ к Интернету (для контейнеров и дополнительных загрузок ПО).

На аппаратные требования влияют несколько факторов:

- Количество обрабатываемых видеопотоков;
- Частота и разрешение кадров видеопотоков;
- Параметры настройки Face Stream. Настройки по умолчанию являются наиболее универсальными. В зависимости от условий эксплуатации приложения с помощью их значений можно повлиять на качество или производительность.

Следует подбирать аппаратное обеспечение на основе вышеперечисленных факторов.

Face Stream также может работать в режиме ускорения вычислений за счет:

### Использование ресурсов видеокарты

Вычисления с использованием видеокарты поддерживаются только для детектора FaceDetV3. См. параметр “defaultDetectorType” в “faceengine.conf”.

Требуется минимум 6Гб оперативной или выделенной видеопамяти. Рекомендуется 8 Гб VRAM или более.

Поддерживаются архитектуры Pascal, Volta, Turing.

Требуется Compute Capability 6.1 или выше.

CUDA версии 11.2.1 уже установлена в Docker контейнере Face Stream. Рекомендуемые драйверы NVIDIA - r450, r455.

В данный момент для одного экземпляра Face Stream поддерживается только одна видеокарта.

### Использование AVX2 инструкций

Требуется CPU с поддержкой AVX2.

Система автоматически определяет наличие инструкций и запускается в оптимальном режиме.

## Программные требования

Для установки Face Stream должны выполняться следующие программные требования:

- Для запуска Face Stream может использоваться RedOS (РЕД ОС) версии 7.3 и выше, CentOS версии 7.8 и выше.

## Описание работы Face Stream

### Режимы работы

Работа Face Stream возможна в двух режимах: обычном и серверном. Основные отличия режимов приведены в таблице ниже.

*Основные отличия обычного и серверного режимов*

	<b>Обычный режим</b>	<b>Серверный режим</b>
Источники	Видеопотоки с IP и USB камер, видеофайлы, фотоизображения	Видеопотоки с IP камер
Использование запросов в API для обработки видеопотоков (см. FaceStreamApi.html)	нет	да
Использование LUNA Configurator для хранения настроек Face Stream	нет	да
Динамическое создание, изменение, удаление источников видеопотоков посредством запросов API	нет	да
Просмотр видеопотоков в браузере с указанными параметрами в режиме реального времени	нет	да
Получение метрик по видеопотоку (количество потоков, количество ошибок, количество необработанных кадров, FPS)	нет	да

Подробную информацию об API Face Stream и возвращаемых данных можно найти в документе "FaceStreamApi.html".

## Общие рекомендации по настройке Face Stream

Данный раздел содержит общие рекомендации по настройке Face Stream.

В разделе указаны названия конфигурационных файлов, в которых настраиваются описываемые параметры.

### Рекомендации перед началом настройки

Настройку Face Stream следует выполнять для каждой используемой камеры отдельно. Face Stream должен работать с видеопотоком камеры, находящейся в стандартных рабочих условиях. Данные требования обусловлены следующими причинами:

- Кадры с разных видеокамер могут отличаться по:
  - уровню шума,
  - размеру кадра,
  - освещённости,
  - размытости и т. п.
- Настройки Face Stream зависят от условий освещения, поэтому будут отличаться для камер, расположенных в тёмном и светлом помещении;
- Производительность Face Stream зависит от количества лиц, одновременно находящихся в кадре. Поэтому настройки для камеры, которая детектирует одно лицо в 10 секунд, будут отличаться от настроек для камеры, которая детектирует 10 лиц в секунду;
- Количество детекций лиц и качество этих детекций зависят от правильного расположения камеры. Если камера расположена под неправильным углом, то лица не будут детектироваться на кадрах или не могут быть использованы для дальнейшей обработки из-за слишком больших углов поворота головы на кадрах;
- Лица в зоне видимости камеры могут быть частично или полностью перекрыты посторонними предметами или на фоне могут располагаться объекты, мешающие алгоритмам распознавания.

Камера может быть расположена так, что освещение или условия съёмки изменяются в течение дня. Рекомендуется протестировать работу Face Stream при разных условиях и выбрать оптимальный режим, обеспечивающий надёжную работу Face Stream при любых условиях.

Если невозможно настраивать Face Stream непосредственно на камере клиента во время её работы, то следует записать видео с этой камеры и использовать это видео для настройки Face Stream.

Если в зону видимости камеры редко попадают лица, то следует записать небольшое видео, самостоятельно имитируя ожидаемые перемещения людей перед камерой. Это видео следует использовать для настройки Face Stream.

Скорость обработки кадров для видео можно задать с помощью параметра “Real\_time\_mode\_fps”.

Для использования видео в качестве источника видеопотока для Face Stream в конфигурационном файле input.json следует указать источник “video-sources” и задать “url” к видео.

Список протестированных с Face Stream камер можно найти в разделе “Совместимость с камерами”.

## Настройка производительности Face Stream

Описанные ниже настройки оказывают наибольшее влияние на производительность Face Stream.

### *Уменьшение области поиска лиц*

Не все области кадра содержат лица. Кроме того, не все лица в кадре имеют необходимый размер и качество. Например, размер лиц на заднем фоне может быть меньше допустимого, а лица по краям кадра могут иметь слишком большие углы поворота.

Параметр “roi” (конфигурационный файл input.json секция “input”), позволяет задать прямоугольную область, в которой будет выполняться поиск лиц.



*Исходный кадр с зоной ROI*

Указанная прямоугольная область вырезается из кадра и именно это изображение в дальнейшем обрабатывает Face Stream.



### *Вырезанная область кадра, используемая для обработки в Face Stream*

Чем меньше зона поиска, тем меньше ресурсов потребляется на обработку каждого кадра.

Правильное использование параметра “roi” существенно повышает производительность работы Face Stream.

### *Масштабирование кадра*

Опция “detector-scaling” (конфигурационный файл “trackengine.conf”) позволяет масштабировать кадр перед обработкой.

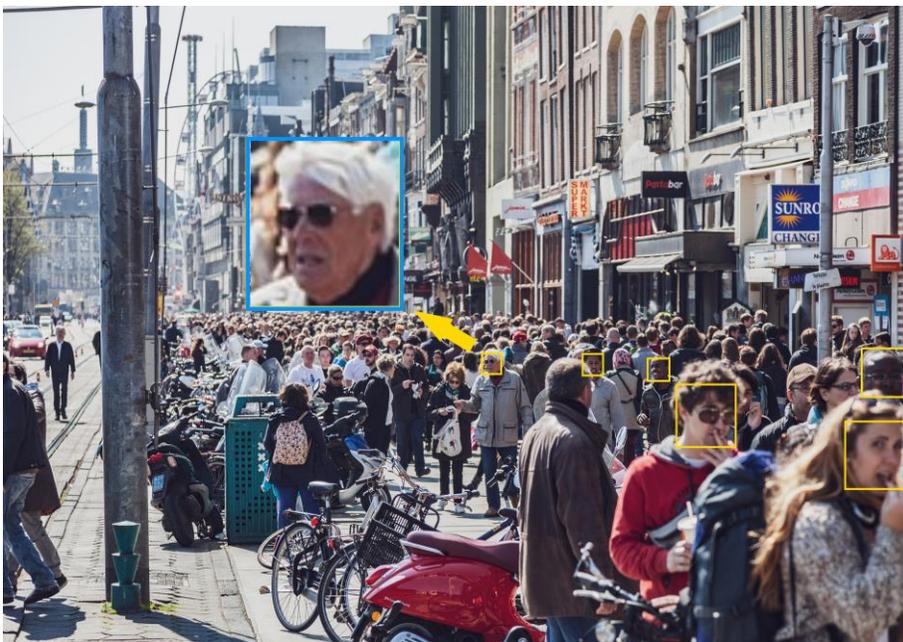
Подходящий для работы размер кадра следует подобрать с помощью параметра “scale-result-size” (конфигурационный файл trackengine.conf). Параметр задаёт максимальный размер кадра после масштабирования по наибольшей из сторон кадра. Если исходный кадр имел размер 1920 на 1080 и значение “scale-result-size” равно 640, то Face Stream будет обрабатывать кадр размером 640 на 360.

Если кадр был обрезан с помощью параметра “roi”, то масштабирование будет применено к этому уменьшенному кадру. В этом случае значение “scale-result-size” следует задавать в зависимости от наибольшей стороны области ROI.

Для подбора оптимального значения “scale-result-size” следует постепенно уменьшать масштаб кадра и проверять, происходит ли на кадре детектирование лиц. Следует выставить минимальный размер изображения, при котором все лица в интересующей области детектируются.

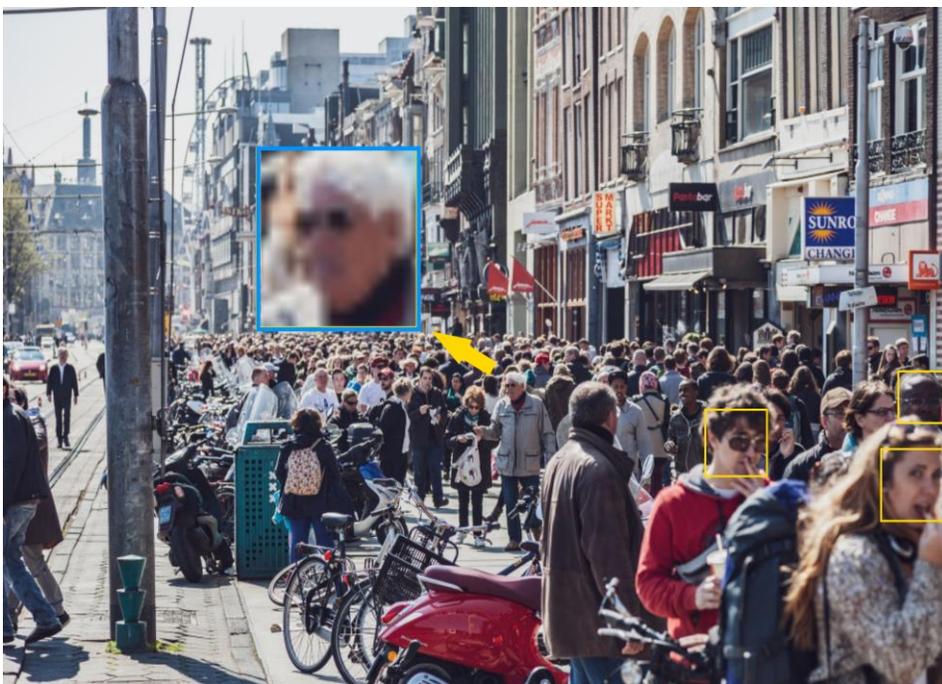
На изображениях ниже приведены примеры детекций для кадра разрешением 1920x1080 и для того же кадра после его масштабирования до разрешения 960x640.

Шесть лиц найдено на изображении с разрешением 1920x1080.



*Детекции на кадре с разрешением 1960X1080*

Три лица найдено после масштабирования изображения до разрешения 960x640. Лица на заднем плане меньше по размеру и имеют низкое качество.



*Детекции на кадре с разрешением 960X640*

Чем меньше разрешение кадра, тем меньше потребляется ресурсов.

## Определение области с движением

	frg-subtractor	frg-regions-alignment	frg-regions-square-alignment
Рекомендованное значение при использовании CPU	1	0	0
Рекомендованное значение при использовании GPU	1	360	0

При включении параметра “frg-subtractor” (конфигурационный файл trackengine.conf) включается режим учёта перемещений в кадре. Последующая детекция лица будет выполняться не на всём кадре, а только в областях с движением.

Определение областей с движением выполняется после масштабирования кадра.

Включение параметра “frg-subtractor” увеличивает производительность работы Face Stream.

Параметр “frg-regions-alignment” (конфигурационный файл trackengine.conf) позволяет задать выравнивание для областей с движением.

При включённом параметре “frg-regions-square-alignment” (конфигурационный файл trackengine.conf) ширина и высота области с движением всегда будут равны.

## Обработка пакетов кадров

Следующие параметры отвечают за пакетную обработку кадров. Параметры настраиваются в конфигурационном файле trackengine.conf.

Параметр “batched-processing” включает пакетную обработку кадров.

При работе с несколькими камерами, с каждой камеры собирается по кадру. После чего происходит обработка данного пакета кадров.

Если параметр отключён, то кадры обрабатываются один за другим.

При использовании режима пакетной обработки увеличивается задержка перед обработкой, но при этом сама обработка выполняется быстрее.

Рекомендуется включать параметр и при использовании GPU, и при использовании CPU.

Параметр “min-frames-batch-size” задаёт минимальное количество кадров, которое следует накопить со всех камер перед обработкой.

Рекомендуется выставлять значение параметра “min-frames-batch-size” равным количеству подключённых потоков при использовании GPU.

Рекомендуется выставлять значение параметра “min-frames-batch-size” равным “2” при использовании CPU.

Параметр “max-frames-batch-gather-timeout” задаёт время между обработкой пакетов кадров.

Если обработка одного кадра укладывается в указанное время и остается дополнительный запас, то Face Stream ожидает дополнительные кадры, чтобы повысить утилизацию GPU.

Если “max-frames-batch-gather-timeout” равен 20 мс, то за это время обрабатывается предыдущий пакет и собирается новый пакет. Через 20 мс начинается следующая обработка, даже если не удалось собрать число кадров равное “min-frames-batch-size”. Следующая обработка кадров не может начаться до окончания обработки предыдущего пакета кадров.

Если он равен “0”, то нет задержки на наполнение пакета кадрами, а обработка происходит непрерывно и “min-frames-batch-size” игнорируется.

Рекомендуется выставлять значение параметра “max-frames-batch-gather-timeout” равным “0” и при использовании GPU, и при использовании CPU.

### *Минимальный размер лица*

В файле “faceengine.conf” следует настроить параметр “minFaceSize”, задающий минимальный размер лица для выполнения детектирования.

Следует задавать наибольший возможный минимальный размер лица, т. к. чем больше лицо, тем меньше ресурсов требуется на выполнение детекции.

Важно, что размер лица будет зависеть от фактического размера кадра, заданного параметром “scale-result-size” (конфигурационный файл trackengine.conf). Лицо, которое имело размер 100 пикселей на кадре 1280x960, будет иметь размер 50 пикселей на кадре размером 640x480.

## **Общая информация о настройке**

### *Работа с треком*

Для каждого найденного лица создаётся новый трек. В рамках этого трека определяются и отправляются лучшие кадры.

В общем случае трек прерывается, когда лицо больше не может быть найдено на кадре.

Если трек был прерван, а человек вновь появился в кадре, то для него будет создан новый трек.

Возможна ситуация, когда два лица пересекаются в кадре (один человек проходит за спиной другого). В этом случае треки для обоих людей обрываются и создаются новые треки.

Возможна ситуация, когда человек отворачивается или его лицо временно перекрывается. В этом случае трек можно не обрывать сразу, а указать параметр “skip-frames” (“trackengine.conf”). Он задаёт количество кадров, в течение которых система будет ожидать повторного появления лица в области, где оно исчезло.

Параметр “detector-step” в “trackengine.conf” позволяет указать количество кадров, на которых будет выполняться редетекция лица в заданной области до выполнения детекции лица. Редетекция требует меньше ресурсов, но лицо может быть потеряно при задании большого количества кадров для редетекции.

### *Отправка лучшего кадра*

Группа параметров “sending” (input.json) позволяет задать параметры для отправки лучшего кадра. Face Stream отправляет полученные лучшие снимки в LUNA PLATFORM (см. “Отправка фотоизображений в LUNA PLATFORM”).

Для увеличения точности распознавания можно отправлять для одного лица не один, а сразу несколько лучших снимков. Для этого нужно включить параметр “number-of-bestshots-to-send” (input.json).

LUNA PLATFORM позволяет агрегировать данные лучшие снимки и создавать на их основе дескриптор лица более высокого качества.

Если за указанный период или до окончания трека не набралось необходимое количество лучших снимков, то отправляются только те лучшие снимки, которые были собраны.

Значения параметров “time-period-of-searching” и “silent-period” могут быть заданы в секундах или кадрах. Для выбора единиц измерения используется параметр “type”.

Далее перечислены основные варианты настроек параметров “time-period-of-searching” и “silent-period” группы “sending”.

1. Лучший снимок отправляется только после того, как трек закончился и человек вышел из зоны видимости камеры.

Таким образом анализируются все кадры с лицом человека и выбирается лучший кадр.

```
time-period-of-searching = -1  
silent-period = 0
```

2. Требуется получить лучший кадр как можно быстрее и отправлять лучшие снимки с указанной периодичностью.

Например, если злоумышленник зашёл в супермаркет, требуется сразу отправить лучший кадр для сравнения с чёрным списком.

Данный режим так же используется для демонстрации работы Face Stream в реальном времени.

В конце трека лучший кадр будет отправлен, даже если заданный период не закончился.

```
time-period-of-searching = 3  
silent-period = 0
```

3. Требуется быстро отправить лучший кадр, а потом отправить заново, только если человек долго находится в кадре.

```
time-period-of-searching = 3  
silent-period = 20
```

4. Требуется быстро отправить лучший кадр и больше никогда не отправлять с этого трека лучший кадр.

```
time-period-of-searching = 3  
silent-period = -1
```

### *Фильтрация кадров*

Фильтрация кадров выполняется по трём основным критериям (задаются в конфигурационном файле “input.json”):

- Углом наклона головы (“detection-yaw-threshold”, “detection-pitch-threshold”, “detection-roll-threshold”).

Для угла поворота (yaw) задаются два дополнительных параметра “yaw-number” и “yaw-collection-mode”, которые позволяют уменьшить вероятность возникновения ошибки, когда вместо большого угла наклона возвращается угол “0”;

- Качеству кадра для его дальнейшей обработки (“min-score”);
- Перекрытию области рта (“mouth-occlusion-threshold”).

Если кадр не прошёл фильтрацию хотя бы по одному из указанных фильтров, то он не может быть выбран в качестве лучшего кадра.

Если задан параметр “number-of-bestshots-to-send”, то кадр, прошедший фильтрацию, добавляется в массив лучших снимков для отправки. Если уже набрано необходимое количество лучших снимков для отправки, то один из этих снимков, имеющий худшее качество кадра (score), будет заменён новым лучшим кадром, если его качество выше.

## Работа со СКУД

При работе со СКУД используются настройки “primary-track-policy”. Они позволяют включить режим работы только с одним лицом, размер которого является самым большим. Предполагается, что интересующее лицо находится близко к камере.

Трек самого большого лица в кадре становится главным треком. Все остальные лица в кадре детектируются, но не обрабатываются. Лучшие снимки для них не отправляются.

Как только другое лицо достигнет большего размера, чем лицо главного трека, трек этого лица станет новым главным треком и обработка будет выполняться для него.

Включение режима выполняется настройкой “use\_primary\_track\_policy”.

Определение лучшего кадра начинается только когда лицо достигает размера (по вертикали), заданного в параметре “best\_shot\_min\_size”. Кадры с лицами меньшего размера не могут стать лучшими снимками.

Когда размер детекции лица по вертикали достигает значения установленного в “best\_shot\_proper\_size”, сразу же выполняется отправка лучшего кадра.

Значения “best\_shot\_min\_size” и “best\_shot\_proper\_size” задаются в зависимости от используемой видеокамеры и её расположения.

Далее приведены примеры настройки параметров группы “sending” для работы со СКУД.

1. Турникет откроется только один раз. Для повторного открытия турникета потребуется оборвать трек (отойти из зоны видимости видеокамер).

```
time-period-of-searching = -1  
silent-period = 0
```

2. Турникет будет открываться с некоторой периодичностью (в данном случае каждые три секунды), если человек встал прямо перед ним.

```
time-period-of-searching = 3  
silent-period = 0
```

При включённом параметре “use\_primary\_track\_policy” лучший кадр никогда не отправляется при окончании трека.

## Отправка фотоизображений в LUNA PLATFORM

Основные параметры, которые должны быть указаны для отправки фотоизображений лиц в соответствующий сервис, указаны ниже.

### Параметры Face Stream

Параметр	Описание
luna_api	Позволяет задать версию API сервиса API для отправки запросов в LP - "6"
request_type	Тип запроса для отправки изображений в LP - "jpeg"
portrait_type	Формат передачи изображения - "warp"
aggregate_attr_requests	Позволяет включать агрегацию лучших кадров для получения одного биометрического шаблона в LUNA PLATFORM - "true" или "false"

### Параметры источников видеопотока (input.json)

Параметр	Описание
name	Значение "name" (20 символов) и автоматически генерируемый идентификатор трека (16 символов) используются для формирования поля external_id в LP - "строка"
luna_account_id	"Luna_account_id", к которому относится выполняемый запрос - "строка в формате UUID"
url	URL, формирующий запрос "generate events" в LP 5 - "http://<luna_api_adress>:5000/6/handlers/{handler_id}/events". В примере указан порт "5000" для сервиса API, который используется по умолчанию

Подробную информацию о ресурсах "/handlers" и "/handlers/{handler\_id}/events" можно найти в документации APIReferenceManual.html в комплекте поставки LUNA PLATFORM 5.

## Дополнительная информация

### Форматы видео

Основные форматы видео, обрабатываемые Face Stream:

- AVI,
- MP4,
- MOV,
- MKV,
- FLV.

### Стандарты сжатия видео

Основные стандарты сжатия видео, с которыми работает Face Stream:

- MPEG4,
- MS MPEG4,
- MS MPEG4v2,
- MJPEG,
- H.264,
- H.265.

### Протоколы

Основные протоколы, используемые Face Stream для получения данных:

- HTTP,
- RTP,
- RTSP,
- TCP,
- UDP.