

1. Синхронизация

Описание:

Синхронизация аудио и видео может быть количественно измерена через вычисление кросс-корреляции, позволяющей определить временной сдвиг, при котором сигналы наиболее тесно соответствуют друг другу.

Алгоритм:

1. Преобразование аудио и видео сигналов во временные ряды. Для аудио, это может быть амплитуда в зависимости от времени; для видео — изменение интенсивности или движения в каждом кадре.
2. Нормализация временных рядов для обеспечения сопоставимости.
3. Вычисление кросс-корреляции между этими временными рядами.
4. Определение максимального значения кросс-корреляции и соответствующего сдвига, при котором достигается наилучшая синхронизация.

2. Эмоциональное соответствие

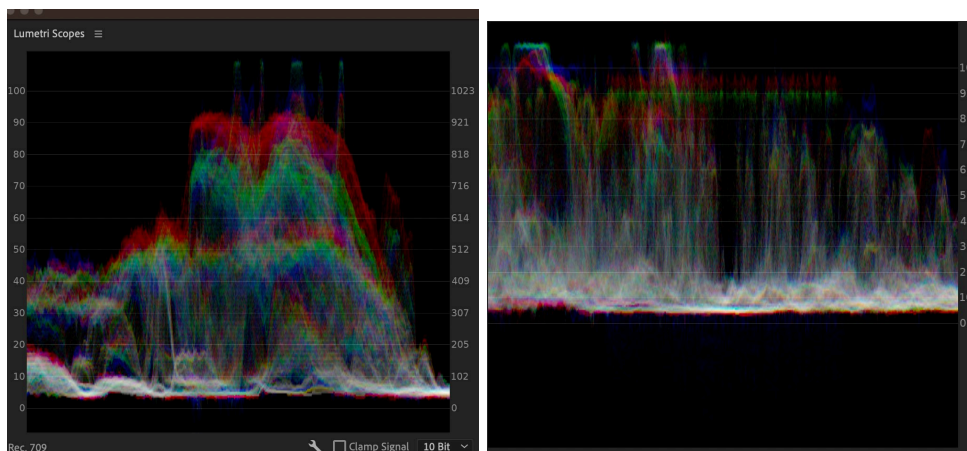
Описание:

Эмоциональное соответствие между аудио и видео может быть оценено путем анализа эмоциональной окраски обеих модальностей и их сравнения.

Алгоритм:

1. Использование аудио анализа для выделения характеристик, связанных с эмоциями, таких как темп, громкость и тональность.
2. Анализ видео на предмет эмоциональных индикаторов, например, цветовой палитры, динамики движения и освещения.

Например можно извлекать у видео данные об альфа-канале:



На рисунке пример альфа канала из видео - два соседних кадра. Видно что их структура сильно различается, на что можно опираться как при оценке темпа так и получить информацию о палитре.

3. Применение алгоритмов машинного обучения для классификации аудио и видео по эмоциональным категориям.

4. Сравнение классификаций аудио и видео для определения степени их эмоционального соответствия.

Может быть выполнено на основе предобученных моделей (например, BERT, VADER) для получения оценок эмоционального тона, которые затем сопоставляются на основе статистического анализа схожести.

3. Оценка темпа музыки и динамики видео

Описание:

Темп музыки и динамика видео могут быть сравнены для оценки их синхронизации, что указывает на ритмическое соответствие (смена сцен, движение).

Алгоритм:

1. Анализ аудио для определения BPM (ударов в минуту).
2. Измерение динамики видео, возможно через анализ изменений между последовательными кадрами, чтобы определить среднюю скорость визуальных изменений.
3. Сравнение BPM аудио с измеренной динамикой видео, например, используя простое отношение или более сложные статистические методы для оценки степени синхронизации.

На сколько представляется мне, BPM может определяться через анализ спектральных характеристик аудио сигнала и его частотных пиков. Возможно CNN или RNN для анализа видеопотока на предмет выявления ритмических паттернов, которые затем синхронизируются с схожими паттернами в музыке.

4. Как не самый быстрый и не самый простой вариант также можно (а возможно даже следует) использовать людей для оценки. Всё таки это то что делается для потребления людьми и есть множество случаев когда технически идеальные в плане музыки ил видео произведения не были по достоинству оценены аудиторией.