

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»
МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОНИКИ И МАТЕМАТИКИ
ДЕПАРТАМЕНТ КОМПЬЮТЕРНОЙ ИНЖЕНЕРИИ

НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

на тему: “Анализ решений по определению схожих видеозаписей”

Выполнил:
Титова Екатерина Егоровна

Научный руководитель:
Королев Денис Александрович,
Доцент

г. Москва, 2019 г.

Аннотация

Видеоданные могут храниться и распространяться в больших количествах с различными форматами. Чем больше материала появляется, тем больше возникает необходимость в управлении данными и поиском нужных отрезков видео. В данном исследовании представлены некоторые из существующих методов поиска, а также идентификации копий видеозаписей, представленных в виде сигнатур или нечетких дубликатов видео. Данные этой работы могут помочь в выборе метода для дальнейшей реализации. Структурный анализ дает четкое представление об устройстве видеозаписи и помогает выявить наиболее подходящий метод, исходя из входных данных. Поскольку в России такие методы не распространены, основными источниками стали зарубежные публикации и исследования.

Теория

Понятие «сигнатура» в видеотехнологиях означает целочисленную или текстовую константу, используемую для однозначной идентификации ресурса или данных [1]. В сфере программирования ее также называют “магическим числом”. Применение поиска сигнатур объектов используют практически во всех методах для определения дубликатов видеозаписи.

Нечеткий дубликат видео (далее НДВ) — это неполное или частичное совпадение объекта с другим объектом подобного класса. Дубликаты классифицируются на естественные и искусственные. В естественных дубликатах объекты схожи при схожих условиях, а искусственные берут истоки на основе оригинального видео, полностью копируя его [2].

Поиски схожих видеозаписей могут быть полезны для систем видеонаблюдения (например, для быстрого поиска развития событий в ночное время суток), определения текущего местонахождения, а также для применения различных фильтров на видеозаписи. Оптическая навигация беспилотных летательных аппаратов также может использовать методы нахождения сигнатур видео. Кибербезопасность может позволить себе находить «пиратские» видеозаписи для защиты авторских прав с помощью рассматриваемых в данной работе методов.

В вопросе нахождения нечетких дубликатов стоит разделять способы на глобальный и локальный. В первом случае рассматривают сигнатуры уровня

кадра для определения информации по цветовым, пространственным и временным параметрам. Данная категория полезна для выявления незначительных правок, тем самым, определять не копии, а схожие видеозаписи. Такие характеристики обобщают статистику в целом для низкоуровневых признаков. Схожие видеозаписи находят по соответствию сигнатур последовательно, однако данные способы поиска не сработают при работе с искусственными дубликатами. Здесь эффективнее методы, которые используют низкоуровневые характеристики кадра. Подходы на уровне сегмента способны к выявлению копий, подвергшихся редактированию, однако данный метод медленнее, так как требователен по памяти.

Основная часть

Теперь рассмотрим метод, основанный на локальных характеристиках. Здесь видеозапись трактуется как объединение изображений, которые сравниваются. Таким образом, фокус внимания переключен именно на них. Для начала определяют особые точки (далее ОТ) на изображении с помощью детектора Харриса, определяющие уникальность, после чего выделяют окрестности этих точек. Далее строят вектора признаков, после чего можно выделить дескрипторы. Сравнению в дальнейшем подлежат именно они. Траектории точек можно отслеживать в течении всего видеоролика. Сравнение с образцом основывается на нечетком поиске, что облегчает положение НДВ в фрагментации. Поскольку траектории точек чувствительны, данный метод применим только в области поиска точных дубликатов видео. В работе «Robust voting algorithm based on labels of behavior for video copy detection» [3] данный метод описан более подробно с математическими обоснованиями.

Следующий метод оценивается на основе SIFT - Scale Invariant Feature Transform. В данной интерпретации подобие определяется средним арифметическим от количества совпавших ОТ. Далее, чтобы выявить сходство видеоролика, требуется полная оценка соответствия (ПОС). Она высчитывается как среднее значение подобия ключевых кадров для всего видеоролика. Поскольку выбираются особые кадры, а не все сразу, данный метод более эффективен по времени. В этой работе привлекает внимание также особый вид выделения временной шкалы вдоль временной оси, что позволяет применить SIFT и вычислить ПОС (рисунок 1). Данный метод сложно применим для видеороликов с частой резкой сменой кадра, а также

для роликов с большим разрешением. Это делает нахождение ОТ более затруднительным [4].

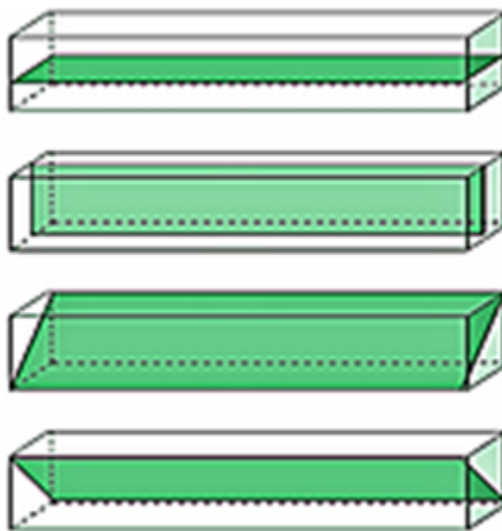


Рисунок 1. Продольный разрез временной шкалы кадра.

Далее появляется более модернизированный метод со сравнением через ОТ. Он изучает визуальные слова, которые строятся на квантовании ОТ. Таким образом, кадры сравнимы друг с другом с применением специальных словарей, идентифицируя операции. Производительность метода увеличена, однако данный метод опирается на предыдущий, используя в основе SIFT-особенности. Минусом данного метода является необходимостью наличия специальной базы – частотного словаря [5], который необходимо составить перед обработкой.

Следующий метод содержит в себе локально-чувствительное хэширование (ЛЧХ), что означает понижение размерности для многомерных данных - разновидности реляционной модели, которая использует многомерные структуры для организации данных и выражают отношения между ними [6]. Основная идея ЛЧХ заключается в подборе хэш-функций для высокой вероятности попадания похожих объектов при некоторых изменениях в одну корзину [7]. Данный метод применяется для отображения цветовой гистограммы каждого ключевого кадра на бинарный вектор. Эти диаграммы в последующем сравниваются для похожих кадров (рисунок 2). Таким образом, показатели каждого кадра представляют из себя набор дискретных величин. Данный метод оказался эффективным, однако он требует большого количества памяти. Подробное экспериментальное подтверждение имеется в статье «Large-scale near-duplicate Web video search: challenge and opportunity» [8].

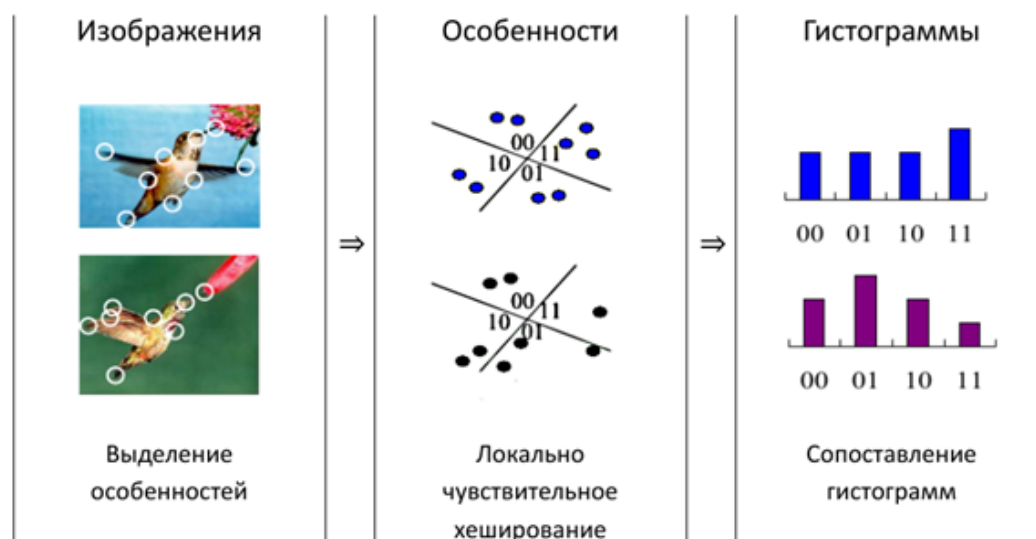


Рисунок 2. Локально-чувствительное хеширование.

Yandan Wang в работе «Near-duplicate Video Retrieval Based on Clustering by Multiple Sequence Alignment» [9] представляет подход, сравнимый с методом построения ДНК. Он предлагает начать поиск сигнатур с классификации видео, используя MSA – Measurement System Analysis. Таким образом, этот биоинформатический метод переносится на видеоряд. Рассмотрим подробнее: для видео выстраивается представление “в виде ДНК” (рисунок 3) и сравнивается с другими видео из ранее существующей базы данных. На основе вычислений строится матрица расстояний, по которой далее методом Neighbor Joining, рисуется направляющий граф – Guide tree. Взяв за основу получившееся дерево, по критериям яркости выполним выравнивание видео, после чего из результатов формируются кластеры видео. Если при сравнении с запросом из базы данных с центром кластера превышен порог, заданный нами, то все ролики внутри кластера можно называть НДВ. По сравнению с предыдущими методами этот является самым точным, однако временная шкала не рассматривается.

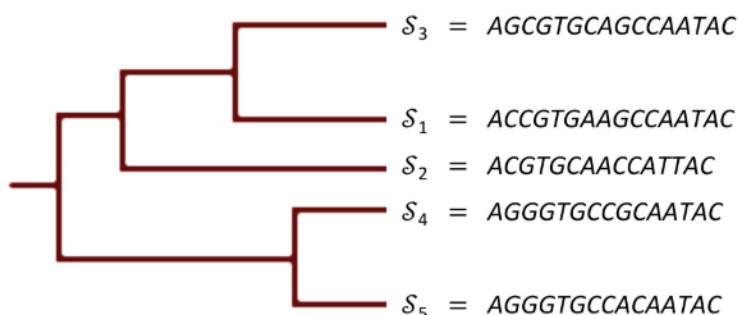


Рисунок 3. Пример графа в ДНК методе

Ранее мы рассмотрели один из методов, использующий подписи. Порядковые подписи можно применять для моделирования относительного распределения интенсивности в кадре. При этом расстояние между фрагментами измеримо временным сходством этих подписей. Так в работе «Robust video signature based on ordinal measure» [10] описан подход, для которого частота кадров, разрешение, а также некоторые пространственные изменения не будут помехами. Данный метод подойдет для поиска только искусственных копий. Наглядно подход изображен на рисунке 4.

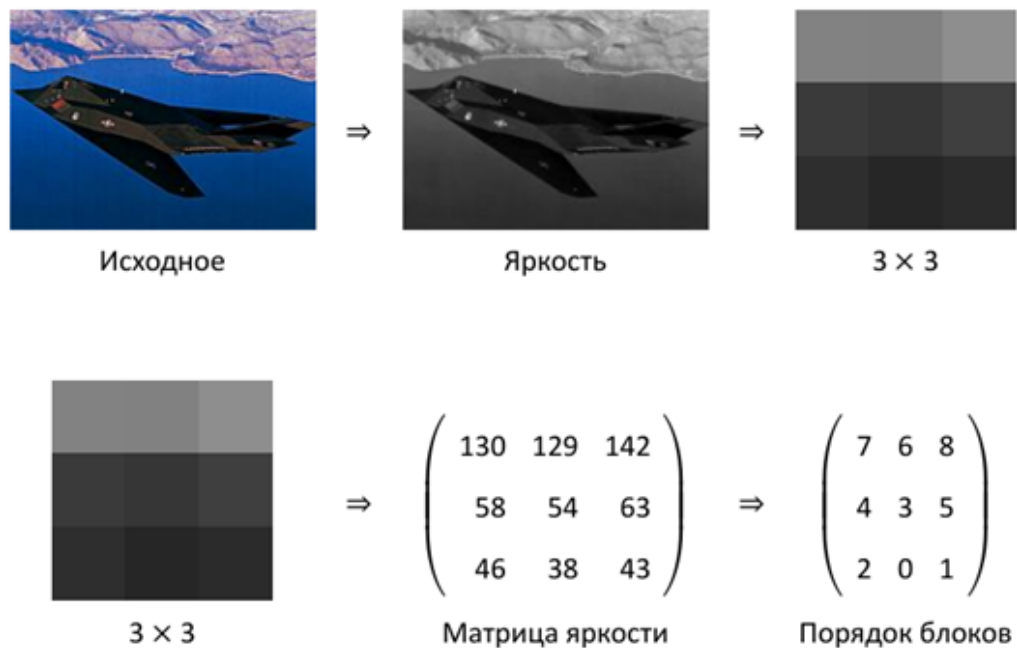


Рисунок 4. Порядковые подписи.

Данный ролик будет содержать особые события с набором ключевых кадров. Таким образом измеряются длины между ключевыми кадрами с помощью гистограмм яркости и записываются в одномерную последовательность. Ее и называют подписью или сигнатурой видео. Для сравнения подписей в данном исследовании используют суффиксный массив, который сортирует лексикографически по всем суффиксам строки [11]. Минус этого метода состоит в поиске общих подстрок для таких массивов. Метод работает неэффективно, если динамика видео выходит за пределы средних значений (переходы между сценами слишком плавные или в сцене присутствует большое количество быстрых перемещений).

В последнем рассмотренном исследовании предлагают на основе подхода сравнения видео по сценам, сопоставлять двоичные деревья сцен, где

поддерево – фрагмент видео. Так, сравнивая координаты корневых вершин подграфов, а также их положение в главных сценах, можно найти дубликаты. Рекурсивно выполняя эти операции для остальных сцен, можно найти искусственные дубликаты. Метод основан на временной информации, однако минусом является отсутствие характеристик самих сцен [12].

Заключение

В результате исследования существующих методов по поиску схожих видеозаписей, наиболее эффективным методом для поиска копии видеозаписи был выбран ДНК-метод. По сравнению с другими подходами он выделяется как самый точный. В нем сравнение между представлениями реализуется только внутри скользящего окна. Предположим, его размерность n . Тогда видеоряды будут сравниваться на основе n -граммного окна с заданными параметрами n и размера шага – сдвига окна после сравнения. Узлом дерева становится профиль видео, а его листьями – видео ДНК. Помимо окон матрицы расстояний, можно сравнивать подобным способом профили, после чего реализовать дерево принятых решений.

В данной научно-исследовательской работе мы рассмотрели 8 различных методов по поиску разных видеокопий - как искусственных, так и естественных. В продолжении планируется выделить один из них и реализовать программно, основываясь на англоязычной литературе. Прежде всего, перед выбором метода важно уметь классифицировать видеозапись, поступающую для сравнения. В зависимости от результата может пригодиться иной метод для обработки. Принципы работы переведены на русский язык, используя словари, объясняющие некоторые математические и функциональные методы, которые используют исследователи. Данное исследование помогло произвести сравнительный анализ существующих способов.

Список литературы

1. Википедия. Магическое число.
[https://ru.wikipedia.org/wiki/Магическое_число_\(программирование\)](https://ru.wikipedia.org/wiki/Магическое_число_(программирование))
2. Никитин Илья Константинович. Элементы поиска нечетких дубликатов видео. <https://www.slideshare.net/w-495/nkp-2015>
3. Robust voting algorithm based on labels of behavior for video copy detection
https://www.researchgate.net/publication/221573843_Robust_voting_algorithm_based_on_labels_of_behavior_for_video_copy_detection
4. Non-identical duplicate video detection using the SIFT method
https://www.researchgate.net/publication/4267374_Non-identical_duplicate_video_detection_using_the_SIFT_method
5. Matthijs Douze, Adrien Gaidon, Herve Jegou, Marcin Marszałek and Cordelia Schmid. INRIA-LEAR'S VIDEO COPY DETECTION SYSTEM.
https://lear.inrialpes.fr/pubs/2008/DGJMS08a/trecvid_douze_jegou_schmid.pdf
6. Национальная библиотека им. Н. Э. Баумана. Многомерные данные.
https://ru.bmstu.wiki/%D0%9C%D0%BD%D0%BE%D0%B3%D0%BE%D0%BC%D0%B5%D1%80%D0%BD%D1%8B%D0%B5_%D0%B1%D0%B0%D0%B7%D1%8B_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85
7. Википедия. Locality Sensitive Hashing.
https://ru.wikipedia.org/wiki/Locality-sensitive_hashing
8. Wan-Lei Zhao, Song Tan and Chong-Wah Ngo. LARGE-SCALE NEAR-DUPLICATE WEB VIDEO SEARCH: CHALLENGE AND OPPORTUNITY. <https://vireo.cs.cityu.edu.hk/papers/icme09-wanlei.pdf>
9. Yandan Wang, Mohammed Belkhatir, Bashar Tahayna. Near-Duplicate Video Retrieval Based on Clustering by Multiple Sequence Alignment.
https://www.academia.edu/2031926/Near-Duplicate_Video_Retrieval_Based_on_Clustering_by_Multiple_Sequence_Alignment
10. Robust video signature based on ordinal measure.
https://www.researchgate.net/publication/224612969_Robust_video_signature_based_on_ordinal_measure
11. Википедия. Суффиксный массив.
https://ru.wikipedia.org/wiki/Суффиксный_массив

12. Алгоритм поиска дубликатов в базе видеопоследовательностей на основе сопоставления иерархии смен сцен.
<https://msu-issled.ru/index.php/kibernetika/20-kibernetika/866-duplikate>