

АНАЛИЗ ТЕКУЩИХ АЛГОРИТМОВ ВЫЧИСЛЕНИЯ ОБЛАСТИ РЕГИОНОВ ИНТЕРЕСА ПОЛЬЗОВАТЕЛЕЙ ПРИ ПОТОКОВОЙ ПЕРЕДАЧИ ВИДЕОКОНТЕНТА

Степанов Никита,

Московский технический университет связи и информатики,
Москва, Россия, stepanovns2022@gmail.com

Матуа Дмитрий,

Московский технический университет связи и информатики,
Москва, Россия, mrskorp.ms@gmail.com

Мазин Владимир,

Московский технический университет связи и информатики,
Москва, Россия, vladimir.a.mazin@gmail.com

Вотьяков Семён,

Московский технический университет связи и информатики,
Москва, Россия, wsemion@gmail.com

Винецкий Всеволод,

Московский технический университет связи и информатики,
Москва, Россия, sevavinetski@gmail.com

Власюк Игорь,

Московский технический университет связи и информатики,
Москва, Россия, i.v.vlasiuk@mtuci.ru

DOI: 10.36724/2072-8735-2023-17-2-27-32

Manuscript received 11 January 2023;

Accepted 07 February 2023

*Исследование выполнено за счет гранта
Российского научного фонда № 23-29-00302
и гранта Российского научного фонда
№ 23-29-00299*

Ключевые слова: регион интереса,
оценка качества видео, машинное обучение,
субъективные данные

В настоящее время существует множество алгоритмов оценки качества видео, которые используют для корректной работы алгоритм поиска региона интереса. Для того, чтобы разработчики алгоритмов могли улучшить работу методов поиска региона интереса, в данной работе мы проводим сравнение алгоритмов поиска регионов интереса с помощью данных, полученных в ходе субъективных тестов. Субъективные оценки региона интереса видео требуют много времени для экспериментов и неосуществимы в приложениях реального времени. Интерес к данной технологии растет с каждым годом. В данной статье мы сравниваем несколько алгоритмов вычисления регионов интереса: нейронную сеть, алгоритм вычисления порога бинаризации для полутонового изображения, метод обнаружения выделяющейся области с помощью преобразования цвета с высокой размерностью и алгоритм сегментации по водоразделу. Эта работа будет полезна исследователям сжатия видео, а также при создании дополнительных тестовых материалов, планировании будущих экспериментов и улучшении существующих алгоритмов поиска регионов интереса.

Для цитирования:

Степанов Н.С., Матуа Д.Д., Мазин В.А., Вотьяков С.Ю., Винецкий В.Ю., Власюк И.В. Анализ текущих алгоритмов вычисления области регионов интереса пользователей при потоковой передаче видеоконтента // Т-Comm: Телекоммуникации и транспорт. 2023. Том 17. №2. С. 27-32.

For citation:

Stepanov N.S., Matua D.D., Mazin V.A., Votyakov S.Yu., Vinetskiy V.Yu., Vlasyuk I.V. (2023) Analysis of the current algorithms for calculating regions area of users interest when streaming video content. *T-Comm*, vol. 17, no.2, pp. 27-32. (in Russian)

Введение

Применение алгоритмов поиска регионов интереса (ROI) широко распространено в наше время. Использование данной технологии можно увидеть, как в медицине, так и в IP-камерах. Методы поиска регионов интереса можно разделить на две категории:

Субъективный. Данный подход определения ROI является довольно точным, так как получателем в большинстве приложений для обработки цифровых изображений или видео является человек, а регионы интереса определяются людьми при помощи устройств отслеживания взгляда, в следствие чего модель, полученная таким методом, может являться эталонной [1, 2]. Однако необходимо отметить, что данный подход является непрактичным ввиду своей дороговизны [3]. Более того, проведение субъективных тестов осложняется различными второстепенными факторами, такими как: зрение, освещение, здоровье, настроение испытуемых и устройство отображения [4]. Результат работы данного подхода показан на рисунке 1.



Рис. 1. Пример работы субъективного метода

Объективный. В данном случае поиск региона интереса осуществляется алгоритмически [5]. Разработка алгоритмов для поиска ROI – цель объективного подхода. Однако, чтобы результаты, полученные при помощи данного метода, были полезны, они должны быть соотнесены с субъективными, или другими словами, результаты должны коррелировать со зрительным восприятием пользователей [6, 7].

В данной работе мы предоставляем анализ часто используемых в настоящее время алгоритмов, которые используются для поиска регионов интереса. В анализе будет использоваться значение корреляции между рассчитанными значениями и данными эталонной модели, которые получены в ходе субъективного тестирования. Эта работа будет полезна для разработчиков алгоритмов поиска ROI, для увеличения производительности алгоритмов поисков регионов интереса, а также позволит оценить работу уже существующих алгоритмов.

Текущая работа

В компьютерном зрении и оптическом распознавании символов область интереса определяет границы рассматриваемого объекта. В данной работе мы анализируем наиболее часто используемые алгоритмы поиска регионов интереса: нейронную сеть (YOLOv4) [8], алгоритм сегментации по водоразделам (Watershed) [9], алгоритм вычисления порога бинаризации для полутонового изображения (метод Оцу) [10] и

метод обнаружения выделяющейся области с помощью преобразования цвета с высокой размерностью (HDCT) [11].

Нейронная сеть (YOLOv4). YOLO — современный алгоритм глубокого обучения, который широко используется для обнаружения объектов. Цель данной нейронной сети - анализировать кадр один раз (одно пропускание кадра через одну нейронную сеть), при этом отображать все необходимые определения объектов для данного кадра. В данном методе все кадры разбиваются на ячейки размером $S \times S$, после чего каждая ячейка отвечает за предсказание нескольких параметров: нескольких содержащих рамок и показателя достоверности для каждой из них, что является вероятностью того, что данная рамка содержит объект поиска. Если в какой-то ячейке сетки объектов нет, то очень важно, чтобы показатель достоверности для этой ячейки был очень малым. Количество выводов считается по формуле:

$$S^2[B(4+1+C)], \quad (1)$$

где B – количество содержащих рамок, предсказанных каждой ячейкой, C – количество классов, 4 – количество содержащих рамок, а 1 – предсказание объектности.

Все предсказания визуализируются. Получается карта всех объектов и набор содержащих рамок, ранжированных по их достоверности. Последнее обновление нейронной сети было представлено в апреле 2020 года. Результаты работы алгоритма можно увидеть на рисунке 2.

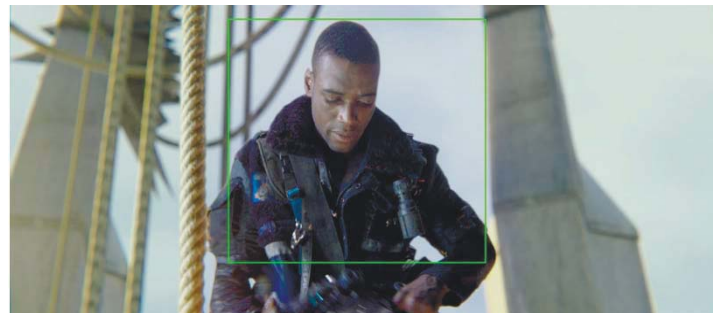


Рис. 2. Работа нейронной сети YOLOv4

Алгоритм сегментации по водоразделам (Watershed). Сегментация кадра — это разбиение кадра на множество покрывающих зон. Данный метод работает с кадром, как с функцией от двух переменных:

$$f = I(x, y), \quad (2)$$

где x, y – координаты пикселя.

Когда по оси OZ откладывается абсолютное значение градиента, можно заметить, что в однородных регионах интенсивности образуются равнины, а в местах перепада – хребты. После нахождения минимумов функции, идет процесс заполнения «водой», начинающийся с глобального минимума. В момент, когда уровень «воды» достигает значения очередного локального минимума, начинается его заполнение. Если два региона начинают сливаться, то строится перегородка, для того чтобы предотвратить объединение областей. «Вода» поднимается до тех пор, пока регионы не будут отделяться

только искусственно построенными перегородками. В результате работы алгоритма получается маска с сегментированным кадром, в котором пиксели одного сегмента имеют одинаковую метку и образуют связную область. Основным недостатком данного алгоритма – необходимость использования процедуры предварительной обработки для кадров с большим количеством локальных минимумов (кадр со сложной текстурой и с обилием различных цветов). Последнее обновление алгоритма было представлено в июле 1994 года. На Рис.3 изображен результат работы алгоритма.



Рис. 3. Результат работы Watershed

Метод Оцу. Данный метод – алгоритм вычисления порога бинаризации для полутонового кадра, используемый в области обработки кадров для получения чёрно-белых изображений и компьютерного распознавания образов.

Алгоритм позволяет различать пиксели двух классов («полезные» и «фоновые»), рассчитывая порог таким образом, чтобы дисперсия внутри класса достигала минимального значения.

Метод Оцу ищет порог, уменьшающий дисперсию внутри класса, определяемую как взвешенная сумма дисперсий двух классов:

$$\sigma_{\omega}^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \quad (3)$$

где веса ω_i – это вероятности двух классов, разделённых порогом t , σ_i^2 – дисперсия этих классов.

Оцу показал, что минимизация дисперсии внутри класса равносильна максимизации дисперсии между классами:

$$\sigma_b^2(t) = \sigma^2 - \sigma_{\omega}^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2, \quad (4)$$

которая выражается в терминах вероятности ω_i и среднего арифметического класса μ_i , которое, в свою очередь, может обновляться итеративно. Последнее обновление метода было в 2016. Результат работы показан на рисунке 4.

HDCT – метод автоматического определения заметных областей кадра с помощью многомерного преобразования цвета. Основная идея состоит в представлении карты значимости кадра в виде линейной комбинации многомерного цветового пространства, где можно четко разделить выделяющиеся области и фон. Это основано на том, что выделяющиеся области часто имеют отличительные цвета по сравнению с фоном в человеческом восприятии, однако стоит отметить, что человеческое восприятие часто бывает достаточно сложным и с нелинейным.



Рис. 4. Результат работы метода Оцу

Глобальный контраст определяется по формуле:

$$D_{G_i} = \sum_{j=1}^N d(c_i, c_j), \quad i = 1, 2, \dots, N, \quad (5)$$

где $d(c_i, c_j)$ обозначает евклидово расстояние между значениями цвета i -го и j -го суперпикселей c_i и c_j .

В то время, как локальный контраст определяется по формуле:

$$D_{L_i} = \sum_{j=1}^N \omega_{i,j}^p d(c_i, c_j), \quad i = 1, 2, \dots, N, \quad (6)$$

где

$$\omega_{i,j}^p = \frac{1}{Z_i} \exp\left(-\frac{1}{2\sigma_p^2} \|p_i - p_j\|_2^2\right), \quad (7)$$

где p_i обозначает положение i -го суперпикселя, Z_i – нормализующее выражение.

После чего необходимо построить карту значимости, которая обозначает линейную комбинацию цветового коэффициента многомерного пространства преобразования цвета по формуле:

$$S_{L,S}(X_i) = \sum_{j=1}^l K_{i,j} \alpha_j, \quad i = 1, 2, \dots, N \quad (8)$$

Достаточно трех итераций для сходимости к стабильной точной карте значимости. Также необходимо сделать пространственное уточнение. Идея состоит в том, что пикселям, которые находятся ближе к определенной области переднего плана, придается больший вес, и наоборот:

$$S_S(X_i) = \exp\left(-k \frac{\min_{j \in F}(d(P_i, P_j))}{\min_{j \in B}(d(P_i, P_j))}\right) \quad (9)$$

Чтобы ускорить процессы уточнения, выполняется уточнение карты значимости. На уровне суперпикселей используется средний цвет суперпикселя в качестве цвета пикселя и центральное расположение суперпикселя в качестве местоположения пикселя. Последнее обновление алгоритма было в 2014 году. На рисунке 5 можно увидеть работу алгоритма.



Рис. 5. Результат работы HDCT

Сбор субъективных данных

В данной работе использовалось устройство отслеживания взгляда – технология, которая отслеживает положения глаз, её также называют технологией отслеживания линий взгляда или точек взгляда.

Eye tracker – это сенсорная технология, обнаруживающая присутствие человека и следящая в режиме реального времени за тем, куда человек смотрит. При помощи данной технологии движения глаз преобразуются в поток данных, в котором содержатся: положение зрачка, вектор взгляда для каждого глаза и точка взгляда. Данная технология расшифровывает движения глаз и переводит их в информацию, которую можно использовать в различных приложениях или для дополнительного метода ввода. При сборе данных использовалось устройство отслеживания взгляда на основе экрана, при этом участник эксперимента подвергался воздействию стимулов на экране.

Пятьдесят наблюдателей в возрасте от 20 до 40 лет без патологий зрения были набраны через Московский технический университет связи и информатики. Участникам был предоставлен монитор с типичным пользовательским контентом и устройством, отслеживающим глаза наблюдателя. Нормальная среда потребления медиа контента имитировалась максимально точно. Следовательно, мы не используем держатель для головы или какие-либо вспомогательные средства для просмотра, такие как линзы. Дисплей представляет собой плоский экран, в данной работе диагональ монитора в опытах 23.6 дюйма, соотношение сторон 16:9, разрешение экрана 1920x1080. Проведена калибровка яркости и цвета дисплея с использованием профессионального экспонетра. Калибровка дисплея соответствует параметрам, определенным в соответствующей рекомендации для проводимого испытания. Цветовая температура белого D65, устройство отслеживания взгляда установлено на мониторе согласно рекомендациям по эксплуатации используемой модели.

Всего в эксперименте было получено более 500 оценок. В работе использовался 95% доверительный интервал.

Для того, чтобы собрать данные, сначала производилась калибровка, при которой использовалось как темное, так и яркое отслеживание зрачка. Для фактической записи выбирался метод, который обеспечивал наибольшую точность. Во время записи устройства отслеживания глаз могут переключаться между ярким и темным отслеживанием зрачка, когда условия изменяются таким образом, что это оказывает сильное негативное влияние на возможность отслеживания.

Если это произойдет, устройство отслеживания взгляда продолжит переключение между двумя методами до тех пор, пока снова не будут получены достоверные данные. После калибровки запускалось видео и проходил сам тест, в результате которого, с помощью программного обеспечения исследователи определяли положение глаза и фокус. Эта технология позволяет изучать мелкие движения глаз и зрительное поведение. Так же с помощью программного обеспечения, используемого для отслеживания взгляда, была создана визуальная карта того, как их испытуемые рассматривали элементы сцены. В процессе работы создавался документ с разрешением csv.

Анализ

В данном разделе представлено сравнение работы алгоритмов поиска региона интереса с эталонным значением. Сравнение проводилось с использованием базы данных качества видео LIVE-NFLX [12]. В данной базе данных содержится 420 искаженных видеопоследовательностей, которые

охватывают широкий набор жанров контента, а также имеют большое разнообразие характеристик, включая сцены с различным освещением, текстурой, ускоренного и замедленного движения. Videopоследовательности данной базы данных были сгенерированы из 15 эталонных видео, которые транслировались в 7 различных сетевых условиях, при использовании 4 стратегий адаптации клиентов. Данная база данных выбрана потому, что это очень реалистичный материал с хорошими откликами на различные аспекты измерений. Для сравнения были выбраны последовательности AirShow, Skateboarding, Sparks, Soccer, TearsOfSteelRobot с использованием стратегии адаптации клиентов OracleVMAFViterbiQualityBasedAdaptor и сетевыми условиями №5, с содержанием сцен, максимально приближенных к реальному пользовательскому контенту. Также мы получили видеопоследовательность с высокими искажениями, после сжатия исходной последовательности с помощью VLC, используя стандарт сжатия MPEG-4 и уменьшив битрейт до 16 кбит/с. Сравнение проводилось по коэффициенту корреляции Пирсона [13] между центром региона интереса, найденным при помощи алгоритма, и точкой взгляда пользователя при субъективных тестах. Результат данного сравнения представлен в таблице 1.

При сравнении полученных показателей для четырех различных алгоритмов, можно сделать вывод, что нейронная сеть YOLOv4 обладает крайне низким коэффициентом корреляции. Наилучший результат продемонстрировал метод обнаружения выделяющейся области с помощью преобразования цвета с высокой размерностью HDCT. Если сравнивать результаты при обработке видеопоследовательности с высокими искажениями, заметно, что метод Оцу показал наилучший результат, в то время как нейронная сеть показала наихудший. Результаты можно увидеть на рисунке 6.

При сравнении полученных показателей для четырех различных алгоритмов, можно сделать вывод, что нейронная сеть YOLOv4 обладает крайне низким коэффициентом корреляции. Наилучший результат продемонстрировал метод обнаружения выделяющейся области с помощью преобразования цвета с высокой размерностью HDCT. Если сравнивать результаты при обработке видеопоследовательности с высокими искажениями, заметно, что метод Оцу показал наилучший результат, в то время как нейронная сеть показала наихудший. На Рис.6 мы можем увидеть сравнение результатов, полученных субъективным методом и алгоритмическим. Как видно из Рис. 6, регионы интереса пользователей, полученные на основе субъективных тестов, расположены в стороне от объекта. Данное расхождение может объясняться тем, что когнитивная составляющая зрительной системы человека имеет задержку восприятия, проявляющуюся в момент смены сцены. На рисунке 7 представлена одна сцена на 18 секунде и на 19,5 секунде после смены сцены в видеопоследовательности.

Таблица 1

Результат сравнения по коэффициенту корреляции Пирсона

	Метод Оцу	HDCT	YOLOv4	Watershed
AirShow	0,8278	0,7815	0,8636	0,6058
Skate	0,8293	0,8976	0,6077	0,7189
Sparks	0,8208	0,7789	0,363	0,6122
Soccer	0,809	0,9309	0,643	0,76
TearsOfSteelRobot	0,8132	0,8001	0,4104	0,6288
Mean	0,82002	0,83768	0,57754	0,66514
SC	0.83	0.74	0.52	0.80

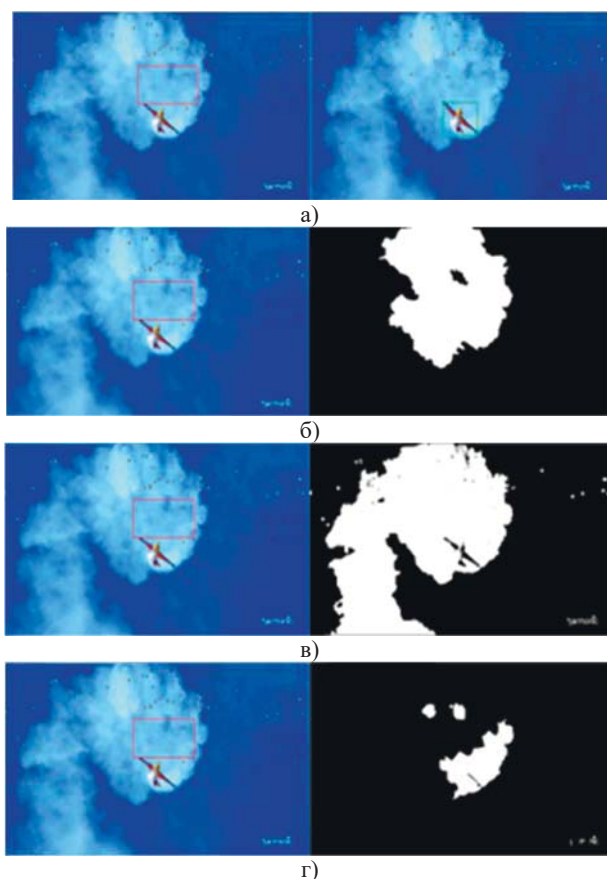


Рис. 6. Результаты обработки видео с высокими искажениями при помощи алгоритмов:
а) YOLOv4, б) Watershed, в) метод Оцу, г) HDCT

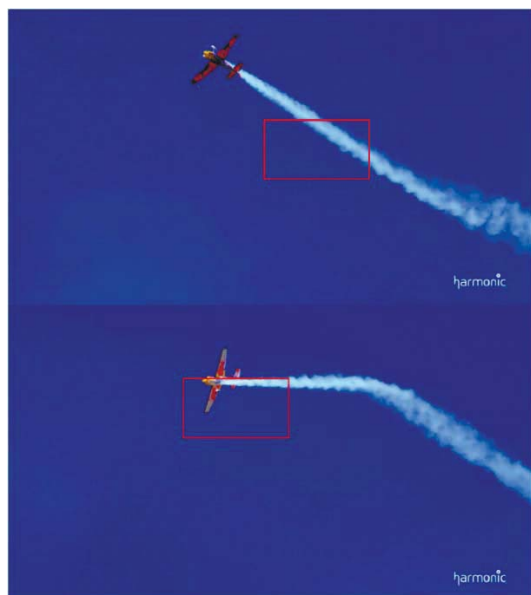


Рис. 7. Точка взгляда в момент смены сцены и через 1,5 секунды после смены сцены

Заключение

В этой статье мы представили анализ часто используемых алгоритмов поиска региона интереса. Сравнение проводилось на видео с использованием сжатия с коэффициентом 2х

из базы данных LIVE-NFLX. Показано, что существуют алгоритмы, которые сильно коррелируют со зрением человека: метод Оцу и HDCT, однако есть и те, у которых корреляция при определении региона интереса довольно слаба: YOLOv4, Watershed. Также при помощи алгоритмов мы обработали видеопоследовательность с высокими искажениями, из чего сделали вывод, что самый высокий результат показал метод Оцу, а самый низкий YOLOv4. Эта работа будет полезна исследователям сжатия видео, а также при создании дополнительных тестовых материалов, планировании будущих экспериментов, и улучшении существующих алгоритмов поиска регионов интереса.

Литературы

1. P. Mohammadu, A. Ebrahimi-Moghadam, S. Shirani. Subjective and Objective Quality Assessment of Image: A Survey // *Majlesi Journal of Electrical Engineering*, 2015. vol.9(1), pp. 55-83.
2. A. Mozhaeva, A. Potashnikov, I. Vlasuyk and L. Streeter. Constant Subjective Quality Database: The Research and Device of Generating Video Sequences of Constant Quality // *2021 International Conference on Engineering Management of Communication and Technology (EMCTECH)*, 2021, pp. 1-5.
3. A. Mozhaeva, I. Vlasuyk, A. Potashnikov, L. Streeter. Full Reference Video Quality Assessment Metric on Base Human Visual System Consistent with PSNR // *2021 28th Conference of Open Innovations Association (FRUCT)*, 2021, pp. 309-315.
4. A. Mozhaeva, I. Vlasuyk, A. Potashnikov, M. J. Cree, L. Streeter. The Method and Devices for Research the Parameters of The Human Visual System to Video Quality Assessment // *2021 Systems of signals generating and processing in the field of onboard communications*, 2021, pp. 1-5.
5. K. Nezhivleva, A. Davydova, A. Drebuszhan, A. Mozhaeva, A. Balobanov. Comparing of Modern Methods Used to Assess the Quality of Video Sequences During Signal Streaming with and Without Human Perception // *2022 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, 2022, pp. 1-6.
6. A. M. Potashnikov, V. A. Mazin, N. S. Stepanov, A. P. Smirnov and A. I. Mozhaeva. Analysis of Modern Methods Used to Assess the Quality of Video Sequences During Signal Streaming // *2022 Systems of Signals Generating and Processing in the Field of on Board Communications*, 2022, pp. 1-4.
7. A. Mozhaeva, V. Mazin, M.J. Cree, L. Streeter. Video quality assessment considering the features of the human visual system // *IVCNZ 2022*.
8. Analysis and Adaptation of YOLOv4 for Object Detection in Aerial Images Aryaman Singh Samyal, Akshatha K R2, Soham Hans, Karunakar A K and Satish Shenoy B arXiv:2203.10194v1 [cs.CV] 18 Mar 2022.
9. Jos B.T.M. Roerdink, Arnold Meijster. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies // *Fundamenta Informaticae* 41, 2001, pp. 187-228 I IOS Press.
10. Noduyuki Otsu. A Threshold Selection Method from Gray-Level Histograms, IEEE, 1979. 0018-9472/79/0100-0062\$00.75
11. Jiwhan Kim, Dongyoon Han, Yu-Wing Tai, Junmo Kim. Salient Region Detection via High-Dimensional Color Transform.
12. Towards Perceptually Optimized End-to-end Adaptive Video Streaming Christos G. Bampis, Zhi Li, Ioannis Katsavounidis, Te-Yuan Huang, Chaitanya Ekanadham and Alan C. Bovik arXiv:1808.03898v1 [eess.IV] 12 Aug 2018.
13. PKD: General Distillation Framework for Object Detectors via Pearson Correlation Coefficient Weihang Cao, Yifan Zhang, Jianfei Gao Anda Cheng, Ke Cheng, Jian Cheng arXiv:2207.02039v2 [cs.CV] 30 Nov 2022.

ANALYSIS OF THE CURRENT ALGORITHMS FOR CALCULATING REGIONS AREA OF USERS INTEREST WHEN STREAMING VIDEO CONTENT

Nikita S. Stepanov, Moscow Technical University of Communications and Informatics, Moscow, Russia, stepanovns2022@gmail.com

Dmitry D. Matua, Moscow Technical University of Communications and Informatics, Moscow, Russia, mrskorp.ms@gmail.com

Vladimir A. Mazin, Moscow Technical University of Communications and Informatics, Moscow, Russia, vladimir.a.mazin@gmail.com

Semion Yu. Votyakov, Moscow Technical University of Communications and Informatics, Moscow, Russia, wsemion@gmail.com

Vsevolod Y. Vinetsky, Moscow Technical University of Communications and Informatics, Moscow, Russia, sevavinetski@gmail.com

Igor V. Vlasyuk, Moscow Technical University of Communications and Informatics, Moscow, Russia, i.v.vlasiuk@mtuci.ru

Abstract

Currently, there are many algorithms for evaluating video quality, which use the region of interest search algorithm for correct operation. In order for algorithm developers to improve the performance of region-of-interest search methods, in this paper we compare algorithms for searching regions of interest using data obtained during subjective tests. Subjective estimates of the region of video interest require a lot of time for experimentation and are not feasible in real-time applications. Interest in this technology is growing every year. In this article, we compare several algorithms for calculating regions of interest: a neural network, an algorithm for calculating a binarization threshold for a grayscale image, a method for detecting a prominent area using a high-dimensional color transformation, and a watershed segmentation algorithm. This work will be useful for video compression researchers, as well as for creating additional test materials, planning future experiments and improving existing algorithms for searching regions of interest.

Keywords: region of interest, video quality assessment, machine learning, subjective data.

References

1. P. Mohammadu, A. Ebrahimi-Moghadam, S. Shirani (2015) Subjective and Objective Quality Assessment of Image: A Survey. *Majlesi Journal of Electrical Engineering*, vol.9(1), pp. 55-83.
2. A. Mozhaeva, A. Potashnikov, I. Vlasuyk and L. Streeter (2021) Constant Subjective Quality Database: The Research and Device of Generating Video Sequences of Constant Quality. *2021 International Conference on Engineering Management of Communication and Technology (EMCTECH)*, pp. 1-5.
3. A. Mozhaeva, I. Vlasuyk, A. Potashnikov, L. Streeter (2021) Full Reference Video Quality Assessment Metric on Base Human Visual System Consistent with PSNR. *2021 28th Conference of Open Innovations Association (FRUCT)*, pp. 309-315.
4. A. Mozhaeva, I. Vlasuyk, A. Potashnikov, M. J. Cree, L. Streeter (2021) The Method and Devices for Research the Parameters of The Human Visual System to Video Quality Assessment. *2021 Systems of signals generating and processing in the field of onboard communications*, pp. 1-5.
5. K. Nezhivleva, A. Davydova, A. Drebuszhan, A. Mozhaeva, A. Balobanov (2022) Comparing of Modern Methods Used to Assess the Quality of Video Sequences During Signal Streaming with and Without Human Perception. *2022 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, pp. 1-6.
6. A. M. Potashnikov, V. A. Mazin, N. S. Stepanov, A. P. Smirnov and A. I. Mozhaeva (2022) Analysis of Modern Methods Used to Assess the Quality of Video Sequences During Signal Streaming. *2022 Systems of Signals Generating and Processing in the Field of on Board Communications*, pp. 1-4.
7. A. Mozhaeva, V. Mazin, M.J. Cree, L. Streeter (2022) Video quality assessment considering the features of the human visual system. *IVCNZ 2022*.
8. Analysis and Adaptation of YOLOv4 for Object Detection in Aerial Images Aryaman Singh Samyal, Akshatha K R2, Soham Hans, Karunakar A K and Satish Shenoy B arXiv:2203.10194v1 [cs.CV] 18 Mar 2022.
9. Jos B.T.M. Roerdink, Arnold Meijster (2001) The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. *Fundamenta Informaticae 41*, pp. 187-228 I IOS Press.
10. Noduyuki Otsu (1979) A Threshold Selection Method from Gray-Level Histograms. 0018-9472/79/0100-0062\$00.75. IEEE.
11. Jiwhan Kim, Dongyoon Han, Yu-Wing Tai, Junmo Kim. Salient Region Detection via High-Dimensional Color Transform.
12. Towards Perceptually Optimized End-to-end Adaptive Video Streaming Christos G. Bampis, Zhi Li, Ioannis Katsavounidis, Te-Yuan Huang, Chaitanya Ekanadham and Alan C. Bovik arXiv:1808.03898v1 [eess.IV] 12 Aug 2018
13. PKD: General Distillation Framework for Object Detectors via Pearson Correlation Coefficient Weihao Cao, Yifan Zhang, Jianfei Gao Anda Cheng, Ke Cheng, Jian Cheng arXiv:2207.02039v2 [cs.CV] 30 Nov 2022.