

Секция «Вычислительная математика и кибернетика»

Автоматический выбор метода построения карт салиентности для видеопоследовательностей

Харенко Максим Викторович

Студент

Московский государственный университет имени М.В. Ломоносова, Факультет вычислительной математики и кибернетики, Королев, Россия

E-mail: extreemal88@yandex.ru

Салиентность – это термин обозначающий свойство объекта, человека, пикселя и т.д. выделяться на фоне группы других, соседних объектов того же типа. Карты салиентности (Saliency maps) – графическое отображение вероятности того, что при первом взгляде человек обратит внимание на конкретные пиксели (Рис. 1).

Saliency maps стоятся для изображений и видеопоследовательностей. Области применения saliency maps: сегментация, распознавание объектов, удаление объектов из видео, зрение роботов, сжатие, создание коллажей, создание миниатюр, уменьшение разрешения с учетом важной окрестности, веб дизайн, а так же Auto Focus и Art effects.

Существует множество факторов того, на что человек обращает свое внимание. На их основе строятся зрительные модели салиентности. Например, для изображений существуют алгоритмы, строящие saliency maps по цвету, яркости, контрастности, частотным характеристикам, наличию лиц и др. В случае видеопоследовательностей дополнительно используются такие параметры как движение, глубина сцены и фокус камеры.

Однако saliency maps, построенные с помощью разных зрительных моделей для одной видеопоследовательности, часто сильно отличаются (Рис.2). Цель данной работы – автоматически оценивать и определять, какой алгоритм построения saliency maps для данной видеопоследовательности больше соответствует поведению человеческих глаз.

Для сравнения результатов алгоритмов между собой, в качестве эталона (ground truth), были взяты данные, полученные с помощью систем слежения за глазами (eye-tracking) (Рис. 3.). Использовались HD-видеопоследовательности из eye-tracking баз TUD(1) и DIEM(2), предоставляющие исходное видео и saliency maps, полученные из результатов слежения за глазами испытуемых (от 25 до 100 человек). Характер видеопоследовательностей: фильмы, трейлеры, новости, спорт, реклама, анимация, документалистика.

Следующая стадия: сравнение карт салиентности, сгенерированных алгоритмами с результатами eye-tracking. В тестировании приняли участие алгоритмы построения saliency maps по движению (motion), частотным характеристикам (frequency) и контрасту (contrast). Для автоматического сравнения двух соответствующих кадров видеопоследовательности был произведен поиск и тестирование метрики, наиболее отвечающей человеческому восприятию. Из более чем 200 сгенерированных роликов была создана база, состоящая из 60ти исходных кадров из 20 видеопоследовательностей (для каждого исходного видео были сгенерированы saliency maps для трех алгоритмов и eye-tracking ground truth).

В тестировании приняли участие метрики: SAD, NCC, PSNR, Similarity, а так же 2 метрики предложенные автором (LMM и LMA), каждая метрика просчитывалась для

всей выбранной базы кадров (Рис 4). Так же на этих данных было проведено субъективное тестирование: 10 человек попросили расположить результаты работы алгоритмов в порядке их «похожести» на ground truth. Если для каждой последовательности метрика располагала алгоритмы так, как люди, то она получала больше баллов. В итоге тестирования победили метрики Similarity и LMA(Рис. 5,6).

RVM-классификатор был обучен выдавать результат работы метрик для пары: исходное изображение и saliency maps текущего алгоритма. Таким образом, это позволяет автоматически выбирать saliency map наиболее соответствующую поведению глаз человека, даже не имея для данной последовательности результатов eye-tracking.

Литература

1. Hani Aleresa, Judith A. Redia, and Ingrid Heynderickxa, “Examining the effect of task on viewing behavior in videos using saliency maps,” Human Vision and Electronic Imaging XVII, 2012.
2. Henderson, J.M.: Diem video and eye tracking database. (<http://thediemproject.wordpress.com>)
3. S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-aware saliency detection,” CVPR, 2010.
4. X. Hou and L. Zhang, “Saliency detection: A spectral residual Approach”, CVPR, 2007.
5. 4.Tilke Judd, Fredo Durand and Antonio Torralba, “A Benchmark of Computational Models of Saliency to Predict Human Fixations,” PAMI, 2012.

Иллюстрации

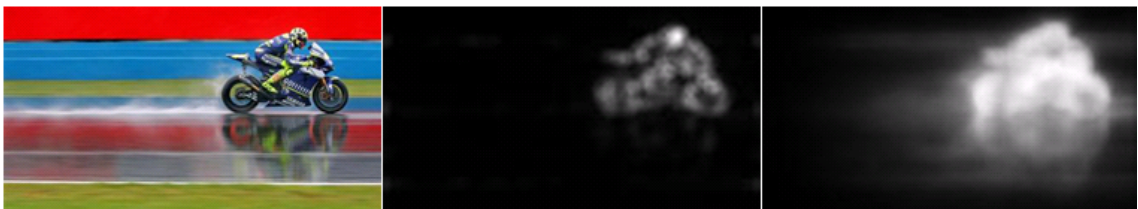


Рис. 1: Исходное изображение и 2 карты салиентности, полученные разными алгоритмами

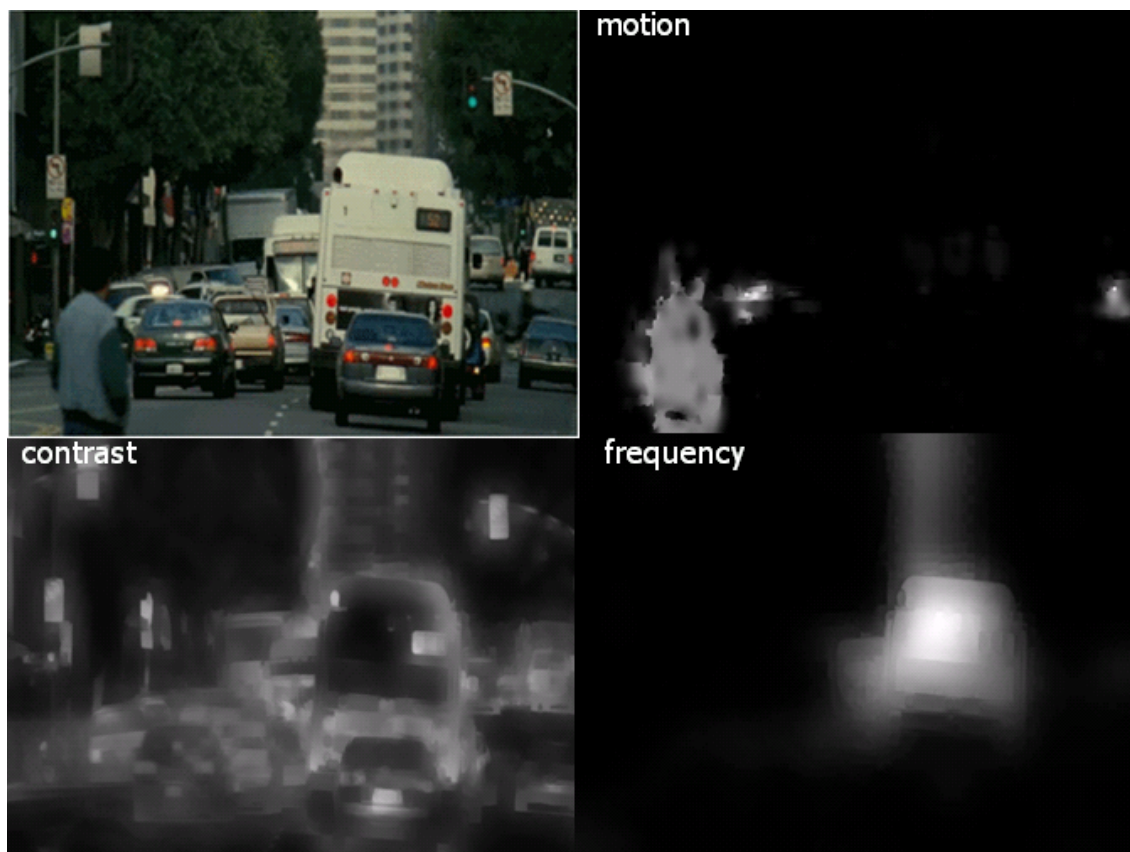
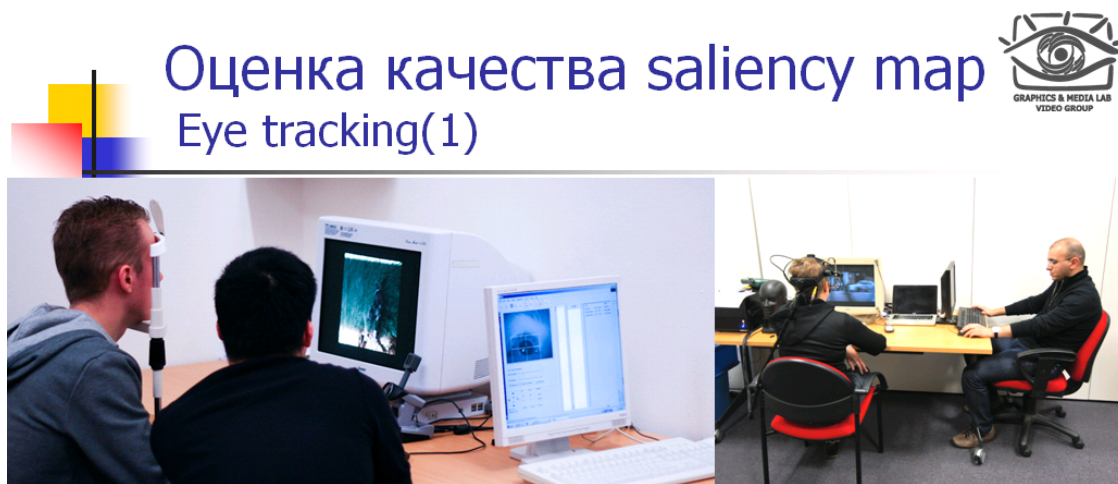


Рис. 2: Различие результатов работы разных алгоритмов.



Оценка качества saliency map Eye tracking(1)

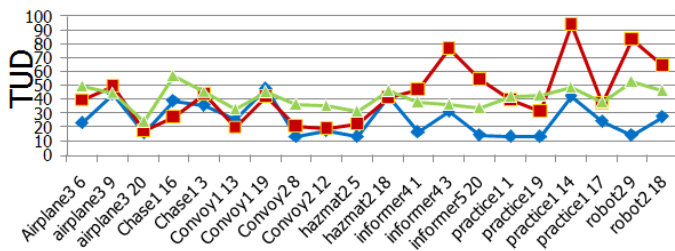
Система слежения за глазами

- В ходе эксперимента испытуемый смотрит в экран
- Плавное последовательное движение глаза, когда наблюдатель следит за происходящим на экране, регистрируется аппаратом как фиксация
- Фиксации всех зрителей объединяются в карты saliency

Рис. 3: Система слежения за глазами



Сравнение метрикой SAD



$$SAD = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |T_{i,j}^1 - T_{i,j}^2|$$

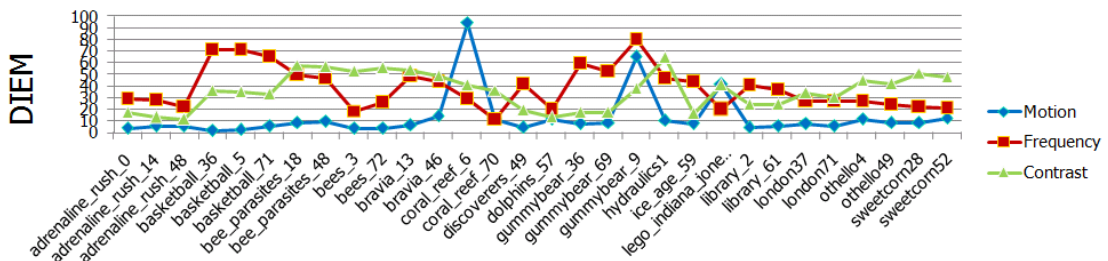
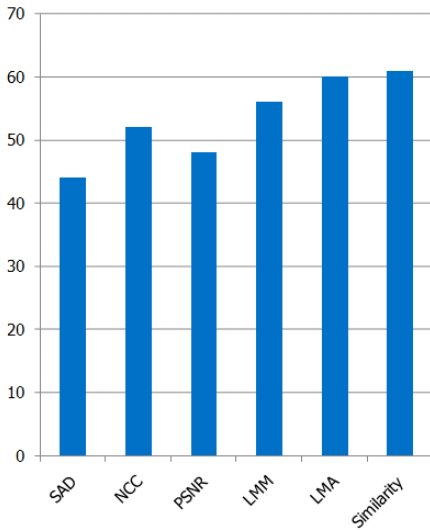


Рис. 4: Метрика SAD для 3х алгоритмов(по оси X название видеопоследовательности)



Оценка метрик для DIEM

Max = 30*3=90



Max = 30*9=270

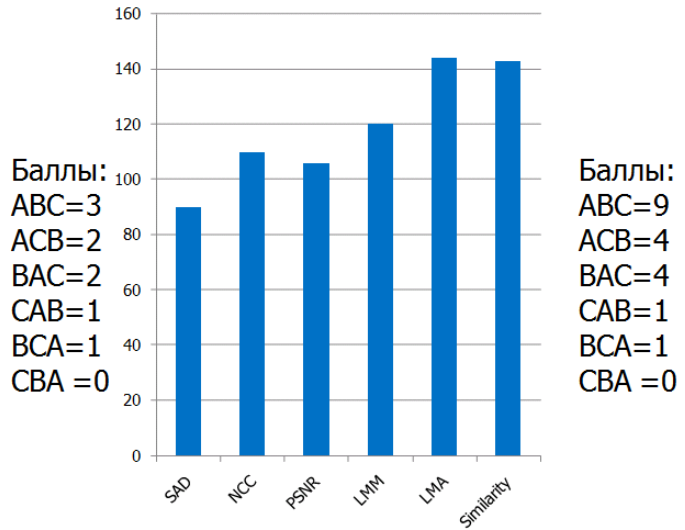
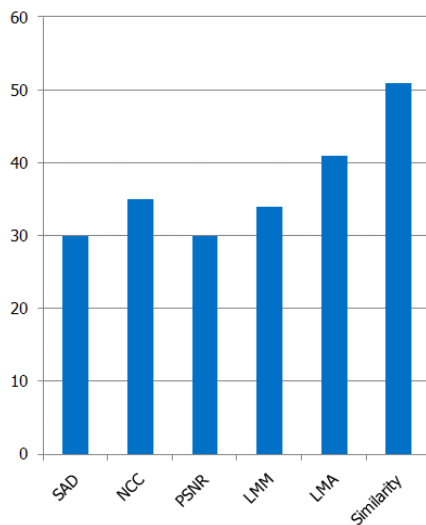


Рис. 5: DIEM



Оценка метрик для TUD

Max = 20*3=60



Max = 20*9=180

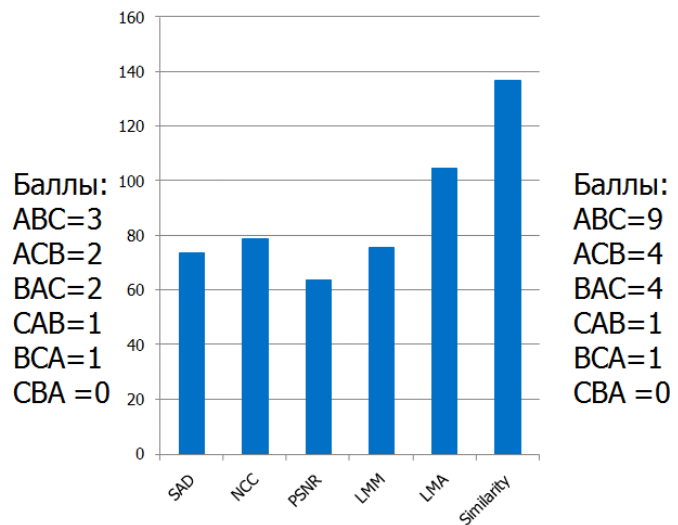


Рис. 6: TUD