

Секция «Вычислительная математика и кибернетика»

Автоматическое построение карты распределения внимания зрителя по площади кадра

Гитман Юрий Александрович

Студент

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

E-mail: oratuomi@gmail.com

Карта внимания (saliency map) для выбранного изображения — это изображение в градациях серого, на котором представлена вероятность (более светлые области означают большую вероятность) для наблюдателя посмотреть на некоторую точку при первом взгляде на изображение.

Приложения алгоритма построения карт внимания обширны. Ниже приведены некоторые из них:

1. Автофокусировка
2. Изменение размера изображения с учетом его структуры
3. Кадрирование изображения с учетом его структуры
4. Сжатие видео и изображений
5. Автоматическая классификация видеопоследовательностей или изображений
6. Web дизайн

В рамках данной работы были реализованы подходы для построения карт внимания на основе следующих признаков:

1. Анализ движения в видео
2. Анализ карты глубины (для стереовидео)
3. Анализ распределения частот [1,2]
4. Анализ самоподобия изображения [3]
5. Анализ семантической информации о кадре (в рамках данной работы был реализован только детектор лиц) [4]

Первые два метода разработаны в течении данного исследования.

Для повышения качества получаемых карт внимания было предложено несколько стратегий постобработки на основании кросс-билатеральной фильтрации и протягивания карт глубины.

Так же была поставлена задача автоматического определения по сцене наиболее подходящего метода. Был обучен бинарный классификатор применяемый к эвристическим признакам, извлеченным из бинарной маски, построенной по карте внимания. Результатом классификации для входной карты внимания является бинарный ответ, является ли данная карта правдоподобной.

Для обучения использовались такие техники машинного обучения, как «метод опорных векторов» (support vector machine [5]) и «метод релевантных векторов» (relevant vector

machine [6]). На тестовой выборке была получена точность классификации — 0.83. Для окончательной верификации результатов были обработаны 8000 кадров выбранных из хорошо известных фильмов.

Литература

1. C. Guo, Q. Ma, and L. Zhang, “Spatio-Temporal Saliency Detection using Phase Spectrum of Quaternion Fourier Transform,” CVPR, 2008, pp. 1–8.
2. X. Hou and L. Zhang, “Saliency detection: A Spectral Residual Approach,” CVPR, 2007, pp. 1–8.
3. S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-Aware Saliency Detection,” CVPR, 2010, pp. 2376–2383.
4. P. Viola and M.J. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” in CVPR, 2001, vol. 1, pp. 511–518.
5. B.E. Boser, I.M. Guyon, V.N. Vapnik, “A training algorithm for optimal margin classifiers,” COLT, 1992, pp. 144–152.
6. M.E. Tipping, “Sparse Bayesian Learning and the Relevance Vector Machine,” The Journal of Machine Learning Research, 2001, pp. 211–244.

Иллюстрации



Рис. 1: Карта внимания построенная на основе информации о глубине

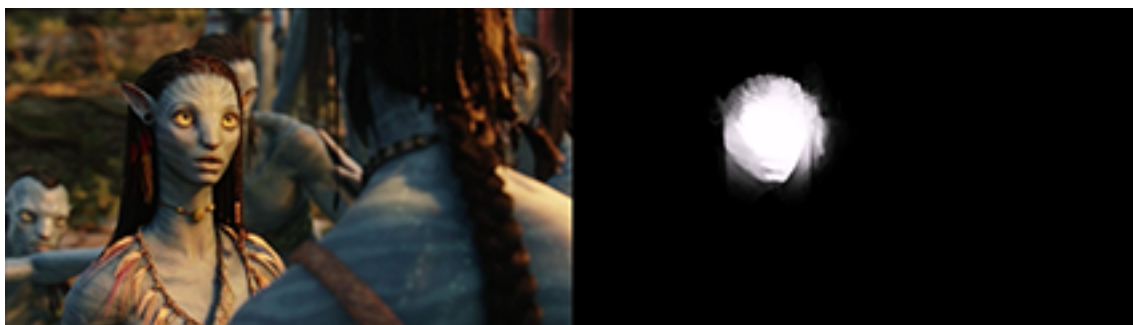


Рис. 2: Карта внимания построенная на основе детектирования лиц

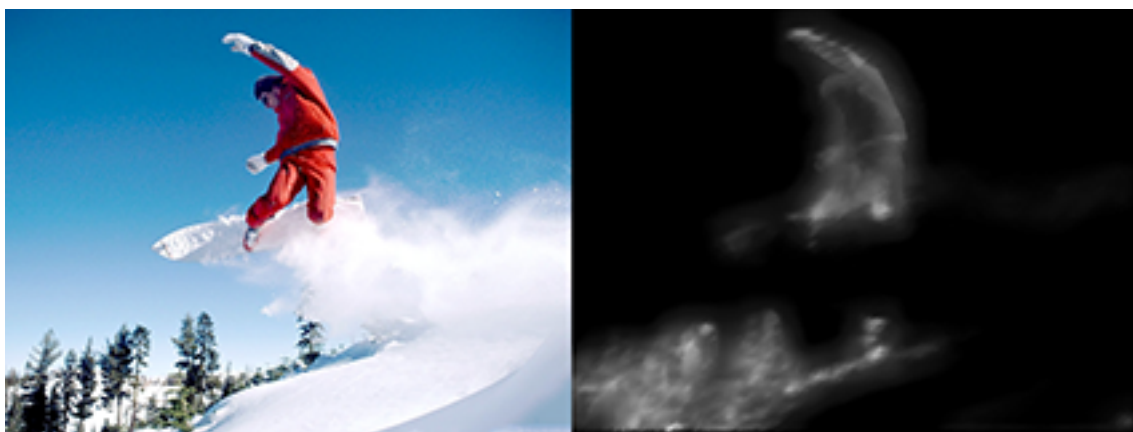


Рис. 3: Карта внимания построенная на основе анализа самоподобий