

## БЫСТРЫЙ МЕТОД СРАВНЕНИЯ КАРТ ВИЗУАЛЬНОГО ВНИМАНИЯ, УСТОЙЧИВЫЙ К ОСНОВНЫМ ИСКАЖЕНИЯМ

*Людвиченко Виталий Андреевич*

*Студент*

*Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия*

*E-mail: vlyudvichenko@graphics.cs.msu.ru*

Карта визуального внимания (карта салиентности) для изображения — это функция вероятности, определенная на множестве пикселей изображения, показывающая вероятность того, что наблюдатель при первом просмотре изображения посмотрит в соответствующий пиксель (рис. 1).

В последнее время публикуется все больше новых методов построения карт внимания для изображений и видео. Также наблюдается рост качества и количества тестовых баз для алгоритмов построения карт салиентности, обусловленный удешевлением и упрощением процесса окулографии. Однако, до сих пор остается ряд вопросов, связанных с методологией сравнения карт визуального внимания.

Обычно для определения меры похожести карт салиентности используются метрики Similarity Score [1] или PSNR. В работе [2] была показана неустойчивость этих метрик к простым преобразованиям, и предложена метрика инвариантная к ряду преобразований. Авторы предлагают рассматривать карту визуального внимания не как функцию вероятности, а как обычное мннхромное изображение, и перед сравнением её с эталонной картой по метрике SS произвести над ней три простых преобразования: гамма-коррекция, коррекция контраста, смешивание с пространственной гауссианой. С помощью такого подхода авторы смогли даже от самых простых автоматических методов построения карт салиентности добиться качества, сопоставимого с передовыми методами. Таким образом встает вопрос проведения специальной процедуры нормализации карт визуального внимания перед их сравнением.

Для решения этой проблемы авторы предложили следующую метрику сравнения карт салиентности:

$$\text{RobustMetric}_{\text{SS}} = \min_{t \in \mathbb{R}^6} \text{SS}(\mathbf{T}(\mathbf{SM}, t), \mathbf{GT}) \quad (1)$$

где  $t$  — вещественные параметры трех преобразований, описанных

выше;  $T$  – функция, выполняющая эти преобразования над картой  $\mathbf{SM}$ ;  $SS$  – метрика Similarity Score, сравнивающая карту  $\mathbf{GT}$  и преобразованную карту  $\mathbf{SM}$ . Таким образом сначала проводится нормализация карты внимания с оптимальным набором параметров, и только после этого вычисляется итоговая оценка. Основным недостатком данной метрики является сложный функционал, имеющий высокую вычислительную сложность, для оптимизации которого приходится задействовать общие алгоритмы оптимизации, не гарантирующие нахождения глобального минимума.

В этой работе представлена новая метрика сравнения карт визуального внимания, которая в отличие от метрики [2] учитывает более широкий класс преобразований яркости, гарантирует нахождение оптимальных параметров нормализации, и имеет на порядки более низкую вычислительную сложность. В предложенном методе процедура нормализации параметризуется произвольным монотонным преобразованием яркости  $f : \mathbb{R} \rightarrow \mathbb{R}$ , и коэффициентом смешивания  $\alpha \in \mathbb{R}$  карты saliентности с предпросчитанным изображением пространственной гауссианы  $\mathbf{CP}$ . Процедура нормализации  $N$  и итоговая метрика имеют следующий вид:

$$N_p(\mathbf{SM}, f, \alpha) = f(\mathbf{SM}_p) + \alpha \mathbf{CP}_p, p \in \mathbb{R}^2 \quad (2)$$

$$\text{ProposedMetric}_{\text{MSE}} = \min_{f \in \mathbb{M}, \alpha \in \mathbb{R}} \sum_p (N_p(\mathbf{SM}, f, \alpha) - \mathbf{GT}_p)^2 \quad (3)$$

Перейдем к задаче оптимизации с конечным числом скалярных переменных. Проведем дискретизацию изображения  $\mathbf{SM}$  на  $n$  целочисленных градаций яркости  $0 \dots n-1$ . Тогда  $f \in \mathbb{R}^n$  с ограничениями на монотонность  $f_{i+1} > f_i$ . Приняв  $t = (f, \alpha) \in \mathbb{R}^{n+1}$ , мы сможем переписать уравнение 3 в матричном виде:

$$\text{ProposedMetric}_{\text{MSE}} = \min_{\substack{t \in \mathbb{R}^{n+1} \\ t_{i+1} > t_i, i=1 \dots n-1}} \frac{1}{2} t^T H t + K t + C \quad (4)$$

где  $H \in \mathbb{R}^{(n+1) \times (n+1)}$  – положительно определенная матрица с ненулевыми элементами на диагонали, в последнем столбце и последней строке;  $K \in \mathbb{R}^{n+1}$ ;  $C \in \mathbb{R}$ . Таким образом исходная задача оптимизации параметров нормализации может быть сведена к задаче квадратичного программирования с выпуклой целевой функцией, для которой можно довольно быстро вычислить её единственное решение с помощью уже известных методов, например метода внутренней точ-

ки.

Предложенная метрика была протестирована на тестовой последовательности из 20000 тыс. кадров для 10 различных методов построения карт салиентности. Для всех методов удалось добиться нормализации карт внимания, превосходящей по метрике PSNR нормализации, предложенной в методе [2] (при использовании в нем целевой функции PSNR). Также был получен выигрыш в скорости более чем 100 раз т. к. предложенный метод требует всего один проход по видеоследовательности, а в методе [2] количество проходов равно количеству итераций, требуемых оптимизационной схемой из функции `fmincon` пакета Matlab.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта 15-01-08632 а.

### Иллюстрации

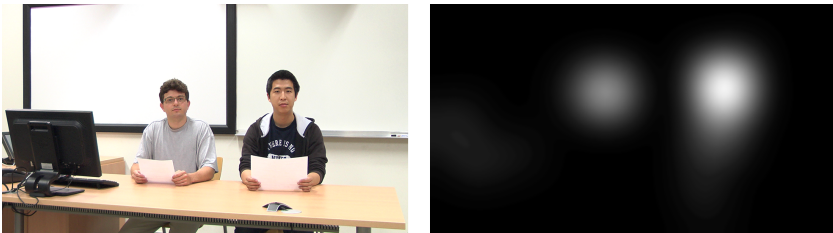


Рис. 1. Слева: Исходное изображение. Справа: Карта визуального внимания.

### Литература

1. Tilke Judd, Krista Ehinger, Fredo Durand, and Antonio Torralba, “Learning to predict where humans look,” *International Conference on Computer Vision (ICCV)*, 2009, pp. 2106–2113.
2. Y. Gitman, M. Erofeev, D. Vatolin, A. Bolshakov, A. Fedorov. “Semiautomatic Visual-Attention Modeling and Its Application to Video Compression”. *2014 IEEE International Conference on Image Processing (ICIP)*. Paris, France, pp. 1105-1109.