

УДК 519.245; 519.674

Б. Г. Кухаренко, канд. физ.-мат. наук, ст. науч. сотр., вед. науч. сотр., e-mail: kukharenkobg@gmail.com,
Институт машиноведения РАН, г. Москва,
М. О. Солнцева-Чалей, аспирант, e-mail: solnceva.chalei@gmail.com,
Московский физико-технический институт (ГУ)

Использование методов выделения ярких областей при сегментировании телеметрических изображений для идентификации групп объектов

Для ускорения идентификации групп объектов и слежения в режиме реального времени необходимо удалить неинформативную часть фона и выделить наиболее яркие области изображения. Яркие объекты выделяются сегментированием изображения по методу сдвига среднего в пространстве характеристик пикселей и построением карты яркости. Численные эксперименты показывают, что процесс обнаружения ярких объектов не может быть только низкоуровневым. Поэтому используется унифицированная модель, интегрирующая характеристики пикселей изображения низкого уровня и знание высокого уровня. Эффективность подхода демонстрируется на примере изображения с яркими объектами.

Ключевые слова: анализ изображений, сегментирование, метод сдвига среднего, карта яркости

Введение

Для сокращения времени обработки при решении задачи наблюдения и идентификации групп объектов в режиме реального времени сегментирование телеметрических изображений низкого качества посредством алгоритма сдвига среднего выполняется после удаления неинформативной части (фона) с помощью метода швов [1]. После сегментирования определение ярких (*salient*) областей позволяет быстро и точно выделить яркие объекты из окружающего фона, что способствует значительному сокращению времени обработки изображений [2–5]. Таким образом, после сегментирования изображений методы определения ярких областей выявляют группы объектов [6], обеспечивают слежение за объектами (*object tracking*) [7], редактирование изображений [8], целевой поиск в базах данных изображений (*image retrieval*) [9].

Наблюдатель выделяет особенности отдельных областей изображения в связи с решением конкретной задачи [10]. Процесс выявления ярких областей определяется стадиями зрительного восприятия: сначала выполняется быстрый обзор изображения "снизу-вверх" (*bottom-up*) ([11–13]), затем более медленное рассмотрение кадра "сверху-вниз" (*top-down*) [14, 15]. При рассмотрении изображения "снизу-вверх" визуальное восприятие области изображения связано с ее цветом, градиентом цвета и

границами этой области на фоне остального изображения [16, 17]. После сегментирования изображения этот процесс моделируется на основе методов декомпозиции матричного представления изображения на матрицу низкого ранга, представляющую фон, и разреженную матрицу, представляющую яркие области [17, 18].

Алгоритмы определения ярких областей обычно используют ограничительные контуры (*bounding boxes*) [19, 20], сегментирование бинарного переднего плана и фона (*binary foreground and background segmentation*) [15, 21] и карты яркости (*saliency maps*), определяющие принадлежность каждого пикселя к яркой области [13]. Существующие методы определения ярких областей недостаточно совершенны. Не всегда удается использовать полный спектр пространственных частот исходного изображения [22]. Низкое разрешение карты яркости приводит к неточному определению ярких областей [23]. Кроме того, в зависимости от способа определения яркой области некоторые карты могут давать недостаточно определенные границы объектов [24]. Также существуют методы, которые точно выделяют только малые яркие области, но способны выделять более крупные. В настоящей работе задача точного выделения ярких областей решается на основе методов, учитывающих спектр пространственных частот сегментированного изображения [22, 25].

1. Выделение ярких объектов на основе спектра пространственных частот сегментированного изображения

Детектор (фильтр) карты яркости для сегментированного изображения должен обеспечивать: 1) выделение наиболее крупных ярких объектов; 2) выделение ярких объектов как связанных однородных областей; 3) точное определение границ ярких объектов, не чувствительное к высоким частотам, обусловленным текстурой, шумом и артефактами; 4) создание карты яркости полного разрешения. Обозначим пороговые значения низких и высоких пространственных частот как ω_{lc} и ω_{hc} , соответственно. Для выделения крупных ярких объектов ω_{lc} должна быть достаточно низкой (первый критерий). Это также выделяет яркие объекты как однородные области (второй критерий). Для получения хорошо определенных границ ярких объектов ω_{hc} должна быть достаточно высокой (третий критерий). Но для удаления текстурных паттернов, шума, артефактов кодирования следует отбросить самые высокие пространственные частоты (четвертый критерий). Поэтому для создания карты яркости следует объединить несколько узкополосных фильтров, покрывающих требуемый интервал частот $[\omega_{lc}, \omega_{hc}]$.

Объединение разностей Гауссовых фильтров. В качестве узкополосного фильтра выбираем разность Гауссовых фильтров (*Difference of Gaussians — DoG*)

$$DoG(x, y) = G(x, y, \sigma_1) - G(x, y, \sigma_2) = \frac{1}{2\pi} \left[\frac{1}{\sigma_1^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_1^2}\right) - \frac{1}{\sigma_2^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_2^2}\right) \right], \quad (1)$$

поскольку она хорошо аппроксимирует широко применяемый при определении локальных особенностей изображения лапласиан Гауссового фильтра (*Laplacian of Gaussian — LoG*), когда стандартные отклонения σ_1 и σ_2 двух Гауссовых фильтров, контролирующие ширину полосы, имеют отношение 1:1,6 [26–29].

Если ввести обозначение $\sigma_1 = \rho\sigma$ и $\sigma_2 = \sigma$, то результат суммирования L узкополосных *DoG*-фильтров

$$\sum_{l=0}^{L-1} G(x, y, \rho^{l+1}\sigma) - G(x, y, \rho^l\sigma) \equiv G(x, y, \sigma\rho^L) - G(x, y, \sigma), \quad (2)$$

равно разности двух Гауссовых фильтров с отношением стандартных отклонений $P = \rho^L$, поскольку все члены за исключением первого и последнего при сложении дают ноль. Если $P \gg 1$, то применение *DoG*-фильтра (2) суммирует результат детекторов краев (фильтров) с $\rho = 1,6$ в нескольких масштабах изображения (поэтому яркие области открыты полностью).

Для полосы частот $[\omega_{lc}, \omega_{hc}]$ *DoG*-фильтра (1) с $\sigma_1 > \sigma_2$ нижняя граница ω_{lc} определяется σ_1 , а верх-

няя граница $\omega_{hc} = \sigma_2$. Однако на практике, чтобы задать большое значение отношения стандартных отклонений, σ_1 полагается бесконечно большой. Это позволяет калибровать пороговые значения частот при сохранении всех остальных частот. Чтобы устранить высокие частоты шума и текстуры используется малое Гауссово ядро. В дискретном случае малое Гауссово ядро хорошо аппроксимирует биномиальный фильтр. У биномиального фильтра с элементами $(1/6) [1, 4, 6, 4, 1]$ верхняя граница частот $\omega_{hc} = \pi/2,75$.

Определение карты яркости на основе характеристик низкого уровня для пикселей сегментированного изображения. Карта яркости $S[i, j]$ для сегментированного изображения $L[i, j]$ размерности $M \times N$ вычисляется по формуле

$$S[i, j] = |I_\mu - I_{\omega_{hc}}[i, j]|, \quad (3)$$

где I_μ — среднее значение пикселей изображения и $I_{\omega_{hc}}[i, j]$ — результат свертки изображения $L[i, j]$ с Гауссовым фильтром. Так как используется исходное изображение, карта яркости имеет полное разрешение (четвертый критерий в разд. 1). Формула (3) обобщается с учетом других характеристик низкого уровня, например интенсивности и цветов в пикселе:

$$S[i, j] = \|I_\mu - I_{\omega_{hc}}[i, j]\|_2, \quad (4)$$

где I_μ — вектор усредненных характеристик изображения, $I_{\omega_{hc}}[i, j]$ — вектор характеристик пикселя изображения в пространстве интенсивности и цветов Luv , свернутого с Гауссовым ядром (используется аппроксимация разделимым биномиальным ядром размерности 5×5), и $\|\cdot\|_2$ — Евклидова норма.

Чтобы выделить яркие объекты, на основе карты яркости сегментированного изображения строится бинарная карта, которая получается путем приписывания единиц пикселям ярких объектов и нулей остальным пикселям фона [22]. Для этого вводится адаптивный порог, который определяется как удвоенная средняя яркость $S[i, j]$ (4) сегментированного изображения:

$$T_a = \frac{2}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} S[i, j].$$

В настоящей работе для сегментирования изображения применяется алгоритм сдвига среднего (*mean-shift*), который улучшает границы объектов (в пространстве интенсивности и цветов Luv) [30]. Полный вектор характеристик пикселей изображения в формуле (4) может состоять из нескольких десятков компонент, например, дискретные координаты (i, j) пикселя на плоскости изображения, цветовые характеристики (три цветовых канала RGB или другие каналы цветов, яркость (*hue*) и насыщенность), регулируемые пирамиды (*steerable*

pyramids), дающие 12 откликов фильтров в каждом пикселе (четыре направления по трем различным масштабам), и фильтры Габора (*Gabor filters*) (12 ориентаций и три масштаба, ширина полосы наименьшего фильтра выбирается равной восьми, а фактор масштабирования — двум) (в общей сложности размерность пространства характеристик пикселей изображения равна 55). При сегментировании изображения посредством алгоритма сдвига среднего (*mean-shift* [30]) выбираются полосы пространственных частот и характеристик (*spatial and features bandwidth*), и в результате даже для визуально однородного изображения формируются многочисленные сегменты $\{C_i, i = \overline{1, K}\}$, где $K \gg 1$ — число сегментов. Пусть $I_p, i = \overline{1, K}$ — вектор характеристик центра сегмента $C_p, i = \overline{1, K}$. Поэтому процесс обнаружения ярких объектов не может быть в чистом виде низкоуровневым.

Для обнаружения ярких объектов требуется унифицированная модель, способная интегрировать характеристики изображения низкого уровня и знание высокого уровня [17].

2. Унифицированная модель карты яркости, интегрирующая характеристики низкого уровня и знания высокого уровня

Матричное представление характеристик низкого уровня сегментированного изображения для определения карты яркости. Матричное представление сегментированного изображения имеет вид

$$F = [I_1, \dots, I_K]. F \in \mathbb{R}^{D \times N}, \quad (5)$$

где D — размерность вектора характеристик (например, $D = 55$). Как отмечается во введении (см. [17]), яркие объекты выделяют на относительно однородном фоне сегментированного изображения, используя декомпозицию матричного представления F (5) сегментированного изображения

$$F = L + S, \quad (6)$$

где L — матрица низкого ранга, представляющая фон в пространстве меньшей размерности; S — разреженная матрица, представляющая яркие области. Задача оценки матриц L и S имеет вид

$$(L^*, S^*) = \underset{L, S}{\operatorname{argmin}} (\operatorname{rank}(L) + \lambda \|S\|_0) \quad (7)$$

при условии (6), где $\|\cdot\|_0$ — псевдонорма L_0 , т.е. число ненулевых компонент матрицы. Поскольку (6) — NP -сложная задача оптимизации, то в качестве альтернативы решается упрощенная выпуклая задача

$$(L^*, S^*) = \underset{L, S}{\operatorname{argmin}} (\|L\|_* + \lambda \|S\|_1), \quad (8)$$

при условии (6), где $\|L\|_* = \sum_i \sigma_i(L)$, т.е. сумма сингулярных собственных чисел, — ядерная норма (*nuclear norm*) матрицы и $\|\cdot\|_1$ обозначает L_1 -норму.

В работе [31] показано, что L и S могут быть достаточно точно оценены как решение (8) посредством устойчивого анализа главных компонент (*Robust PCA*). После того как определена матрица S , L_1 -норма столбца $s_i, i = \overline{1, K}$, в $S = [s_1, \dots, s_K]$ используется как оценка яркости сегмента $C_i, i = \overline{1, K}$.

Чем больше $\|s_i\|_1, i = \overline{1, K}$, тем большее значение яркости назначается сегменту $C_i, i = \overline{1, K}$, изображения. Карта яркости, генерируемая и нормированная таким образом, должна быть серо-шкальным изображением. Чтобы гарантировать законность модели (6), необходимо выполнить обучение линейного преобразования пространства характеристик пикселей сегментированного изображения.

Обучение линейного преобразования характеристик сегментированного изображения. Рассмотрим линейные преобразования T исходного пространства характеристик пикселей изображения

$$J_i = T I_p, \text{ где } T \in \mathbb{R}^{D \times D}. G = [J_1, \dots, J_K] = T F \quad (9)$$

и уравнение (6) принимает вид

$$T F = L + S.$$

В правильном пространстве характеристик большая часть векторов фона изображения должна лежать в подпространстве меньшей размерности, которое представляется матрицей низкого ранга. Для обнаружения такого подпространства характеристик можно использовать обучающие изображения, в которых яркие объекты отмечены вручную и выделены в прямоугольную рамку. При заданной выделяющей яркой объект рамке q_i указывает принадлежность вектора I_i в выражении (5) к яркому объекту ($q_i = 0$, если соответствующая область является яркой, $q_i = 1$ — в противном случае). Такая информация может быть представлена в виде диагональной матрицы $Q = \operatorname{diag}(q_1, \dots, q_K)$. Отличие матрицы $T F Q \in \mathbb{R}^{D \times D}$ от $T F$ состоит в том, что векторы характеристик, соответствующие ярким областям, в $T F Q$ становятся равными нулю, так как они умножаются на $q_i = 0$. Поэтому матрица $T F Q$ содержит только информацию о фоне и должна иметь низкий ранг при заданном правильном преобразовании T (9). Следовательно, задача оценки T имеет вид

$$T^* = \underset{T}{\operatorname{argmin}} \left(\frac{1}{I} \sum_{k=1}^I \|T F_k Q_k\|_* - \gamma \|T\|_* \right), \quad (10)$$

при условии $\|T\|_2 = c$.

Здесь F_k и Q_k — представление характеристик и индикатор яркости для k -го обучающего изображения соответственно (I — число обучающих изображений), c — константа. Член $-\gamma \|T\|_*$ позволяет избежать тривиального решения, когда ранг матрицы T становится малым настолько, что ранг матрицы $T F Q$ сильно превышает ранг T , и γ — весовой параметр. Ограничение $\|T\|_2 = c$ позволяет избежать того, чтобы

норма матрицы \mathbf{T} становилось слишком большой или малой.

Локально-оптимальное решение (10) обеспечивается методом спуска по градиенту. Градиент

$$O(\mathbf{T}) = \frac{1}{I} \sum_{k=1}^I \|\mathbf{T}\mathbf{F}_k\mathbf{Q}_k\|_* - \gamma\|\mathbf{T}\|_*$$

относительно \mathbf{T} имеет вид

$$\begin{aligned} \frac{\partial O(\mathbf{T})}{\partial \mathbf{T}} &= \frac{1}{I} \sum_{k=1}^I \frac{\partial \|\mathbf{T}\mathbf{F}_k\mathbf{Q}_k\|_*}{\partial \mathbf{T}} - \gamma \frac{\partial \|\mathbf{T}\|_*}{\partial \mathbf{T}} = \\ &= \frac{1}{I} \sum_{k=1}^I \frac{\partial \|\mathbf{T}\mathbf{F}_k\mathbf{Q}_k\|_*}{\partial (\mathbf{T}\mathbf{F}_k\mathbf{Q}_k)} (\mathbf{F}_k\mathbf{Q}_k)^T - \gamma \frac{\partial \|\mathbf{T}\|_*}{\partial \mathbf{T}}. \end{aligned} \quad (11)$$

Частная производная ядерной нормы произвольной матрицы \mathbf{X} может быть получена с помощью разложения сингулярных чисел (*singular value decomposition* — *SVD*) этой матрицы $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ в виде

$$\frac{\partial \|\mathbf{X}\|_*}{\partial \mathbf{X}} = \mathbf{U}\mathbf{V}^T + \mathbf{W}, \quad (12)$$

где \mathbf{W} — матрица такая, что $\mathbf{U}^T\mathbf{W} = 0$, $\mathbf{V}\mathbf{W} = 0$ и $\|\mathbf{W}\| \leq 1$. При использовании метода спуска по градиенту для (10) начальное значение $\mathbf{T} = \mathbf{I}$, где \mathbf{I} — единичная матрица. Используя (11), итерации

$\mathbf{T}_{t+1} = \mathbf{T}_t - \alpha \frac{\partial O(\mathbf{T})}{\partial \mathbf{T}}$, где α — шаг (на каждой итерации нормируется $\|\mathbf{T}_{t+1}\|_2 = c$), выполняются до достижения локального минимума, который дает оценку правильного преобразования \mathbf{T} (9).

Интегрирование априорных распределений высокого уровня в модель карты яркости. Особенности человеческого восприятия высокого уровня интегрируются в модель карты яркости. Поскольку наибольшее внимание привлекают объекты, близкие к центру изображения, при построении карты яркости используется априорное Гауссово распределение

$$p_c(\mathbf{x}) = \exp(-d(\mathbf{x}, \mathbf{c})/\sigma_1^2),$$

на основе расстояния $d(\mathbf{x}, \mathbf{c})$ от пикселей до центра изображения. Для учета особенностей восприятия цветов (например, красного и желтого) строится $2D$ -распределение гистограммы $H(S)$ в $nR-nG$ пространстве цветов

$$nR = \frac{R}{R+G+B}, \quad nG = \frac{G}{R+G+B},$$

для отмеченных ярких объектов. Аналогично строится распределение гистограммы цветов фона $H(B)$. Для каждого дискретного цвета (*quantized color*) значения, полученные из двух гистограмм $H(S)$ и $H(B)$, обозначаются как h_S и h_B . Затем устанавливается априорное распределение цветов для карты яркости

$$p_c(\mathbf{x}) = \exp((h_S(c_x) - h_B(c_x))/\sigma_3^2),$$

где c_x показывает цвет в положении \mathbf{x} . Далее эти априорные распределения перемножаются и дают окончательную априорную карту яркости.

Априорная вероятность для сегмента быть ярким определяется положением центра сегмента и обозначается p_i . Для всех сегментов изображения эта априорная вероятность может быть представлена диагональной матрицей $\mathbf{P} = \text{diag}(p_1, \dots, p_K)$, и оптимизация

$$(\mathbf{L}^*, \mathbf{S}^*) = \underset{\mathbf{L}, \mathbf{S}}{\text{argmin}} (\|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1) \quad (13)$$

выполняется при условии (6) в виде

$$\mathbf{TFP} = \mathbf{L} + \mathbf{S}.$$

Отметим, что в матрице \mathbf{P} большинство компонент p_i малы. Следовательно, векторы характеристик, умноженные на p_i , в матрице низкого порядка \mathbf{L} являются посторонними (*outliers*), и наиболее вероятно они будут включены в разреженную матрицу \mathbf{S} . Элементы матрицы \mathbf{S} увеличиваются при увеличении p_i так, что на карте яркости области с большими p_i имеют большую яркость. Как показывают численные эксперименты, при включении в модель (6) априорных распределений высокого уровня вероятность точного определения карты яркости при решении (13) повышается в 2 раза.

3. Численный эксперимент

Обработка изображения (рис. 1, *a* [32], см. четвертую сторону обложки) посредством сегментирования изображения методом сдвига среднего (*mean-shift* [30]) (параметры: $\sigma_S = 7$, $\sigma_R = 10$ и $\text{min-Region} = 20$) (рис. 1, *б*) и построения карты яркости сегментированного изображения (рис. 1, *в*) выделяет группу трех ярких объектов (рис. 1, *з*).

Заключение

Показано, что комбинация метода сегментирования на основе алгоритма сдвига среднего и методов выделения ярких областей, интегрирующие характеристики низкого уровня для пикселей изображения и знания высокого уровня, обеспечивают точное выделение ярких объектов в изображении.

Список литературы

1. Кухаренко Б. Г., Солнцева М. О. Использование методов сокращения фона при сегментировании телеметрических изображений для идентификации групп объектов // Информационные технологии. 2014. № 2. С. 3–8.
2. Borji A., Sihite D. N., Itti L. Salient object detection: A benchmark / Computer Vision—ECCV 2012. Lecture Notes in Computer Science. V. 7573. Berlin, Heidelberg: Springer, 2012. P. 414–429.
3. Ren T., Liu Y., Ju R., Wu G. How important is location information in saliency detection of natural images // Multimedia Tools and Applications. 2015. V. 74, Is. 20. P. 1–22.
4. Lang C., Liu G., Yu J., Yan S. Saliency detection by multitask sparsity pursuit // IEEE Transactions on Image Processing. 2012. Vol. 21, N. 3. P. 1327–1338.
5. Goferman S., Zelnik-Manor L., Tal A. Context-aware saliency detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2012. Vol. 34, N. 10. P. 1915–1926.

6. **Cheng M. M., Mitra N. J., Huang X., Hu S. M.** Salient-shape: Group saliency in image collections // The Visual Computer. Springer. 2014. Vol. 30, Is. 4. P. 443–453.
7. **Ren T., Qui Z., Yu T., Liu Y., Bei J.** Soft-assigned bag of features for object tracking // Multimedia Systems. 2014. Vol. 21, N. 2. P. 189–205.
8. **Cheng M. M., Mitra N. J., Huang X., Hu S. M.** RepFinder: Finding approximately repeated scene elements for image editing // ACM Transactions on Graphics (TOG). Proc. of ACM SIGGRAPH 2010. 2010. Vol. 29, N. 4. Article No. 83.
9. **Gao Y., Wang M., Zha Z. G., Shen J., Li X., Wu X.** Visual-textual joint relevance learning for tag-based social image search // IEEE Transactions on Image Processing, 2013. Vol. 22, N. 1. P. 363–376.
10. **Borji A., Itti L.** State-of-the-art in visual attention modeling // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013. Vol. 35, N. 1. P. 185–207.
11. **Cheng M., Mitra N. J., Huang X., Torr P. H.S., Hu S.** Global contrast based salient region detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. Vol. 37, Is. 3. P. 569–582.
12. **Qiu Y., Sun X., She M. F.** Saliency detection using hierarchical manifold learning // Neurocomputing. 2015. Vol. 168. P. 538–549.
13. **Jiang P., Vasconcelos N., Peng J.** Generic promotion of diffusion-based salient object detection // Proc. of the IEEE International Conference on Computer Vision. 2015. P. 217–225.
14. **Yang J., Yang M. H.** Top-down visual saliency via joint CRF and dictionary learning // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2012), IEEE. 2012. P. 2296–2303.
15. **Liu T., Yuan Z., Sun J., Wang J., Zheng N., Tang X., Shum H. Y.** Learning to detect a salient object // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011. Vol. 33, Is. 2. P. 353–367.
16. **Borji A., Itti L.** State-of-the-art in visual attention modeling // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013. Vol. 35, Is. 1. P. 185–207.
17. **Shen X., Wu Y.** A unified approach to salient object detection via low rank matrix recovery // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2012). IEEE, 2012. P. 853–860.
18. **Yang C., Zhang L., Lu H., Ruan X., Yang M. H.** Saliency detection via graph-based manifold ranking // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2013). IEEE, 2013. P. 3166–3173.
19. **Chang K. Y., Liu T. L., Chen H. T., Lai S. H.** Fusing generic objectness and visual saliency for salient object detection // Proc. of the IEEE International Conference on Computer Vision (ICCV'2011). IEEE, 2011. P. 914–921.
20. **Feng J., Wie Y., Tao L., Zhang C., Sun J.** Salient object detection by composition // Proc. of the IEEE International Conference on Computer Vision (ICCV'2011). IEEE. 2011. P. 1028–1035.
21. **Lu Y., Zhang W., Lu H., Xue X.** Salient object detection using concavity context // Proc. of the IEEE International Conference on Computer Vision (ICCV'2011). IEEE, 2011. P. 233–240.
22. **Achanta R., Estrada F., Wils P., Süssstrunk S.** Frequency-tuned salient region detection // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2009). IEEE, 2009. P. 1597–1604.
23. **Hou X., Zhang L.** Saliency detection: A spectral residual approach // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07). IEEE, 2007. P. 1–8.
24. **Itti L., Koch C.** A saliency-based search mechanism for overt and covert shifts of visual attention // Vision Research. 2000. Vol. 40, N. 10. P. 1489–1506.
25. **Achanta R., Estrada F., Wils P., Süssstrunk S.** Salient region detection and segmentation / Proceedings of the International Conference on Computer Vision Systems (ICVS '08). Springer Lecture Notes in Computer Science. V. 5008. Berlin, Heidelberg: Springer. 2008. P. 66–75.
26. **Marr D.** Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, CA: W. H. Freeman and Company. 1982.
27. **Lowe D. G.** Distinctive image features from scale-invariant keypoints // International Journal of Computer Vision. 2004. Vol. 60, N. 2. P. 91–110.
28. **Harel J., Koch C., Perona P.** Graph-based visual saliency. Advances in Neural Information Processing Systems. Vol. 19 / Ed. Schölkopf B., Platt J. C., Hoffman T. Cambridge, MA: The MIT Press. 2006. P. 545–552.
29. **Itti L., Koch C., Niebur E.** A model of saliency-based visual attention for rapid scene analysis // IEEE Transactions on Pattern Analysis & Machine Intelligence. 1998. Is. 11. P. 1254–1259.
30. **Comaniciu D., Meer P.** Mean shift: A robust approach toward feature space analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. Vol. 24, Issue 5. P. 603–619.
31. **Wright J., Ganesh A., Rao S., Peng Y., Ma Y.** Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization / Bengio Y., Schuurmans D., Lafferty J. D., eds. Advances in Neural Information Processing Systems. V. 22. Cambridge, MA: The MIT Press. 2009. P. 2080–2088.
32. **Chambers G.** Oxford's WWII Harvard Airplanes. URL: <http://www.whatravelwriterssay.com/oxfordcounty4feb09gene.html>

B. G. Kukhareno, Leading Research Scientist, e-mail: kukharenskobg@gmail.com,
Blagonravov Mechanical Engineering Research Institute of the RAS,
M. O. Solntseva-Chalei, e-mail: solnceva.chalei@gmail.com,
Post-Graduate Student, Moscow Institute of Physics and Technology (SU)

Use of Salient Region Detecting Methods after Telemetric Image Segmentation for Object Group Identification

In this paper we describe how to accelerate object group identification and tracking in real time. It is in need to remove non-informational part of image backward and to detect highly salient objects. The salient objects are detected by mean-shift segmentation in image pixel feature high dimensional space and by determining the segmented image saliency map. The frequency-tuned salient region detection method is in use, which produces segmented image full resolution saliency map displaying salient object boundaries. The method preserves the object boundaries by retaining main frequency content from the segmented image and exploits features of color and luminance. Since salient object detection is not a pure low-level process, next a unified model is in use to incorporate the image low-level features with higher-level knowledge. In the model, a special image matrix in the feature space is decomposed as a low-rank matrix representing background plus sparse matrix indicating the salient regions. To ensure the model validity, a linear transform of the image feature space has to be learned. Next, the higher-level knowledge is fused to compose a prior map, and is treated as a prior term in the objective function. The approach efficiency is demonstrated on sample image with salient objects

Keywords: image analysis, segmentation, mean-shift, salient map

References

1. **Kukharenskiy B. G., Solntseva-Chalei M. O.** Ispolzovanie metodov sokraschenia fona pri segmentirovanii telemekhicheskikh izobrazheniy dlya identifikatsii grupp objectov, *Informatsionnye tehnologii*, 2014, no. 2, pp. 3–8.
2. **Borji A., Sihite D. N., Itti L.** Salient object detection: A benchmark, *Computer Vision — ECCV 2012. Lecture Notes in Computer Science*, vol. 7573, Berlin, Heidelberg: Springer, 2012. P. 414–429.
3. **Ren T., Liu Y., Ju R., Wu G.** How important is location information in saliency detection of natural images, *Multimedia Tools and Applications*, 2015, vol. 74, issue 20, pp. 1–22.
4. **Lang C., Liu G., Yu J., Yan S.** Saliency detection by multitask sparsity pursuit, *IEEE Transactions on Image Processing*, 2012, vol. 21, no. 3, pp. 1327–1338.
5. **Goferman S., Zelnik-Manor L., Tal A.** Context-aware saliency detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, vol. 34, no. 10, pp. 1915–1926.
6. **Cheng M. M., Mitra N. J., Huang X., Hu S. M.** Salienshape: Group saliency in image collections, *The Visual Computer*, Springer, 2014, vol. 30, issue 4, pp. 443–453.
7. **Ren T., Qui Z., Yu T., Liu Y., Bei J.** Soft-assigned bag of features for object tracking, *Multimedia Systems*, 2014, vol. 21, no. 2, pp. 189–205.
8. **Cheng M. M., Mitra N. J., Huang X., Hu S. M.** RepFinder: Finding approximately repeated scene elements for image editing, *ACM Transactions on Graphics (TOG), Proceedings of ACM SIGGRAPH — 2010*, 2010, vol. 29, no. 4, article No. 83.
9. **Gao Y., Wang M., Zha Z. G., Shen J., Li X., Wu X.** Visual-textual joint relevance learning for tag-based social image search, *IEEE Transactions on Image Processing*, 2013, vol. 22, no. 1, pp. 363–376.
10. **Borji A., Itti L.** State-of-the-art in visual attention modeling, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, vol. 35, no. 1, pp. 185–207.
11. **Cheng M., Mitra N. J., Huang X., Torr P. H.S., Hu S.** Global contrast based salient region detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, vol. 37, issue 3, pp. 569–582.
12. **Qiu Y., Sun X., She M. F.** Saliency detection using hierarchical manifold learning, *Neurocomputing*, 2015, vol. 168, pp. 538–549.
13. **Jiang P., Vasconcelos N., Peng J.** Generic promotion of diffusion-based salient object detection, *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 217–225.
14. **Yang J., Yang M. H.** Top-down visual saliency via joint CRF and dictionary learning, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2012)*, IEEE, 2012, pp. 2296–2303.
15. **Liu T., Yuan Z., Sun J., Wang J., Zheng N., Tang X., Shum H. Y.** Learning to detect a salient object, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, vol. 33, issue 2, pp. 353–367.
16. **Borji A., Itti L.** State-of-the-art in visual attention modeling, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, vol. 35, issue 1, pp. 185–207.
17. **Shen X., Wu Y.** A unified approach to salient object detection via low rank matrix recovery, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2012)*, IEEE, 2012, pp. 853–860.
18. **Yang C., Zhang L., Lu H., Ruan X., Yang M. H.** Saliency detection via graph-based manifold ranking, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2013)*, IEEE, 2013, pp. 3166–3173.
19. **Chang K. Y., Liu T. L., Chen H. T., Lai S. H.** Fusing generic objectness and visual saliency for salient object detection, *Proceedings of the IEEE International Conference on Computer Vision (ICCV'2011)*, IEEE, 2011, pp. 914–921.
20. **Feng J., Wic Y., Tao L., Zhang C., Sun J.** Salient object detection by composition, *Proceedings of the IEEE International Conference on Computer Vision (ICCV'2011)*, IEEE, 2011, pp. 1028–1035.
21. **Lu Y., Zhang W., Lu H., Xue X.** Salient object detection using concavity context, *Proceedings of the IEEE International Conference on Computer Vision (ICCV'2011)*, IEEE, 2011, pp. 233–240.
22. **Achanta R., Estrada F., Wils P., Süsstrunk S.** Frequency-tuned salient region detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2009)*, IEEE, 2009, pp. 1597–1604.
23. **Hou X., Zhang L.** Saliency detection: A spectral residual approach, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, IEEE, 2007, pp. 1–8.
24. **Itti L., Koch C.** A saliency-based search mechanism for overt and covert shifts of visual attention, *Vision Research*, 2000, vol. 40, no. 10, pp. 1489–1506.
25. **Achanta R., Estrada F., Wils P., Süsstrunk S.** Salient region detection and segmentation, *Proceedings of the International Conference on Computer Vision Systems (ICVS '08)*. Springer Lecture Notes in Computer Science, V. 5008. Berlin, Heidelberg: Springer. 2008. P. 66–75.
26. **Marr D.** *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco, CA: W. H. Freeman and Company, 1982.
27. **Lowe D. G.** Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 2004, vol. 60, no. 2, pp. 91–110.
28. **Harel J., Koch C., Perona P.** Graph-based visual saliency / Schölkopf B., Platt J. C., Hoffman T., *Advances in Neural Information Processing Systems*, vol. 19, Cambridge, MA: The MIT Press, 2006, pp. 545–552.
29. **Itti L., Koch C., Niebur E.** A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1998, issue 11, pp. 1254–1259.
30. **Comaniciu D., Meer P.** Mean shift: A robust approach toward feature space analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, vol. 24, issue 5, pp. 603–619.
31. **Wright J., Ganesh A., Rao S., Peng Y., Ma Y.** Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization, Bengio Y., Schuurmans D., Lafferty J. D., eds., *Advances in Neural Information Processing Systems*, vol. 22, Cambridge, MA: The MIT Press, 2009, pp. 2080–2088.
32. **Chambers G.** Oxford's WWII Harvard Airplanes, URL: <http://www.whatstravelwriterssay.com/oxfordcounty4feb09gene.html>