

УДК 519.674

Применение байесовского подхода для оценки непрозрачности пикселей и создания композитных изображений для целей рекламы

Э.М. Гиоргис¹

¹Московский физико-технический институт (государственный университет)

Введение

В данной работе развивается подход к анализу и созданию рекламных изображений, описанный в выпускной квалификационной работе «Анализ искусственных и естественных изображений в рекламе посредством непараметрической оценки функции плотности распределения пикселей».

Широкое применение во многих областях современной жизни находят композитные (составные) изображения. Среди таких областей визуализация различных архитектурных проектов, реклама, киноиндустрия, телевидение и прочие области.

Для создания композитного изображения необходим передний план исходного изображения и некоторый фон. Сложность выделения переднего плана изображения заключается в том, что края объекта могут состоять из очень мелких частей или быть нечеткими (волосы, перья, прозрачные материалы). Методы сегментации изображений, направленные на выделение объекта переднего плана, не дают детализации и не дают информации о прозрачности объекта [1].

Рекламные изображения в большинстве своем содержат в себе рекламируемый объект и фон. Поэтому композитные изображения один из наиболее используемых методов их создания. Основным направлением в рекламе сейчас является естественность и ненавязчивость созданных рекламных продуктов [2]. Композитное рекламное изображение с передним планом, полученным методами, основанными на предварительной сегментации изображения, не обладает в полной мере характеристиками естественного изображения, в то время как изображение, содержащее в себе информацию о прозрачности переднего плана и детализации на границах, воспринимается как более естественное.

В данной работе описывается процесс создания композитного изображения путем сложения фона и переднего плана изображения (альфа-канала), полученного при применении к исходному изображению MAP-оценивания.

Описание метода

Любое изображение, состоящее из фона и некоторого целевого объекта, можно разделить на три области: объект, фон и спорную область на границе объекта. Каждый пиксель изображения, соответственно, относится к одной из этих областей. Изображение разбивки на области носит название тримап. Объект на тримапе обозначается белым цветом, фон - чёрным, а спорная область - серым.

Задачей метода является выделение из спорной области той её части, которая относится к объекту. Для этого каждому пикселю спорной области назначается показатель его отношения к объекту. С визуальной точки зрения этот показатель является компонентой прозрачности пикселя (и обозначается α).



Рис. 1. Пример изображения с тримапом

Далее будет представлен общий вид оператора композиции. Пусть I , F и B – цвета композита, переднего плана и фона пикселя соответственно. Тогда I может быть представлен как сумма цветов F и B , каждый из которых имеет свой показатель прозрачности. Эти показатели зависимы, их сумма равна единице. Таким образом, происходит линейная смесь переднего плана и фона. Описанное выше выражается уравнением композиции (1).

$$I = \alpha * F + (1 - \alpha) * B \quad (1)$$

При МАР-оценивании (maximum a posteriori) производится поиск наиболее правдоподобных оценок для F , B и α при заданном изображении I . Другими словами, производится максимизация распределения вероятностей. Затем, с использованием правила Байеса, результат представляется как максимизация суммы логарифмов правдоподобия [3,4]:

$$\begin{aligned} \underset{F, B, \alpha}{\operatorname{argmax}} P(F, B, \alpha | I) &= \\ &= \underset{F, B, \alpha}{\operatorname{argmax}} P(I | F, B, \alpha) P(F) P(B) P(\alpha) / P(I) = \\ &= \underset{F, B, \alpha}{\operatorname{argmax}} (L(I | F, B, \alpha) + L(F) + L(B) + L(\alpha)) \end{aligned} \quad (2)$$

где $L(.) = \log P(.)$ – логарифм правдоподобия, а член $P(I)$ опущен, потому что он постоянен относительно параметров оптимизации.

Таким образом, задача сводится к определению логарифмов правдоподобий $L(I|F, B, \alpha)$, $L(F)$ и $L(B)$:

I. $L(I|F, B, \alpha)$ моделируется измерением разности между наблюдаемым цветом I и цветом, предсказываемым с помощью оцененных F , B и α (3).

$$L(I|F, B, \alpha) = -\|I - \alpha * F - (1 - \alpha) * B\|^2 / \sigma_i^2 \quad (3)$$

Этот логарифм правдоподобия моделирует ошибку в измерении цвета композита и соответствует Гауссову распределению с центром $\bar{I} = \alpha * F + (1 - \alpha) * B$ и стандартным отклонением σ_i .

II. Для оценки логарифма правдоподобия переднего плана $L(F)$ строится распределение вероятности цвета F . При этом используются известные и прежде оцененные цвета в пределах окрестности N пикселя. Чтобы надежней моделировать распределение цвета переднего плана, вклад каждого пикселя ближайшего к $i \in N$ взвешивается в соответствии с двумя отдельными факторами:

- вклад пикселя взвешивается посредством $(\alpha[i])^2$, который дает цветам непрозрачных пикселей более высокое доверие;
- используется Гауссово уменьшение $g[i]$ с $\sigma = 8$, чтобы подчеркнуть вклад ближайших пикселей над удаленными.

Из этих двух факторов складывается комбинированный вес $w[i] = (\alpha[i])^2 g[i]$. При заданных цветах переднего плана и соответствующих им весах, сначала производится разбиение цветов пикселей на несколько кластеров. Для каждого полученного кластера вычисляется взвешенный средний вектор \bar{F} RGB-цветов размерности 3×1 и взвешенная ковариационная матрица Σ_F размерности 3×3 :

$$\bar{F} = \frac{1}{W} \sum_{i \in N} w[i] F[i] \quad (4)$$

$$\Sigma_F = \frac{1}{W} \sum_{i \in N} w[i] (F[i] - \bar{F})(F[i] - \bar{F})^T \quad (5)$$

где $W = \sum_{i \in N} w[i]$. Затем логарифм правдоподобия для переднего плана $L(F)$ может

моделироваться, как выводимый из ориентированного эллиптического Гауссова распределения, используя взвешенную ковариационную матрицу

$$L(F) = -(F - \bar{F})^T \Sigma_F^{-1} (F - \bar{F}) / 2 \quad (6)$$

III. Определение логарифма правдоподобия для фона $L(B)$ зависит от того, какая задача решается. Для оценки области непрозрачности (альфа канал) естественного изображения используется выражение аналогичное используемому для переднего плана изображения, устанавливая $w[i] = (1 - \alpha[i])^2 g[i]$ и подставляя B вместо F в каждый член формул (6), (7) и (8).

Далее, предполагается, что логарифм правдоподобия для непрозрачности $L(\alpha)$ постоянный. Из-за того, что α умножается на F и B в логарифме правдоподобия $L(I|F, B, \alpha)$ (5), функция, которая максимизируется в (4), не квадратичная относительно неизвестных. Для эффективного решения этой задачи она разбивается на две квадратичные подзадачи.

В первой подзадаче, предполагается, что α постоянная. При этом предположении, взятие частных производных (4) относительно F и B и приравнение их нулю дает

$$\begin{aligned} & \begin{bmatrix} \sum_F^{-1} + l\alpha^2/\sigma_I^2 & l\alpha(1-\alpha)/\sigma_I^2 \\ l\alpha(1-\alpha)/\sigma_I^2 & \sum_B^{-1} + l(1-\alpha)^2/\sigma_I^2 \end{bmatrix} \begin{bmatrix} F \\ B \end{bmatrix} \\ & = \begin{bmatrix} \sum_F^{-1} \bar{F} + I\alpha/\sigma_I^2 \\ \sum_B^{-1} \bar{B} + I(1-\alpha)/\sigma_I^2 \end{bmatrix} \end{aligned} \quad (7)$$

где l – тождественная матрица размерности 3×3 . Следовательно, для постоянной α , можно найти наилучшие параметры F и B , решая 6×6 линейное уравнение (7).

Во второй подзадаче предполагается, что F и B постоянные, что дает квадратичное уравнение для α . Решение этого уравнения получается проекцией наблюдаемого цвета I на линейный сегмент FB в пространстве цвета

$$\alpha = \frac{(I - B) * (F - B)}{\|F - B\|^2}, \quad (8)$$

где числитель содержит покомпонентное умножение двух векторов разности цветов. Чтобы максимизировать логарифм правдоподобия (2), попеременно предполагается, что α фиксированная, чтобы получить решение для F и B , используя (7), и предполагается, что F и B фиксированные, чтобы получить решение для α , используя (8). Оптимизация (2) начинается с инициализации α средним значением α по окрестности ближайших пикселей и затем решается уравнение (7) с постоянной α [3,4].

Использование метода и полученные результаты

В работе было исследовались изображения, состоящие из фона и объекта. Предварительно для каждого изображения был составлен тримап для последующего анализа. В данной работе для применения MAP-оценивания используется программа Matlab [5] (рис. 2).



Рис. 2. Интерфейс программы Matlab

Интуитивно понятно, что чем точнее указана принадлежность пикселя к фону или объекту (чем ближе тримап к истинному представлению о переднем плане и фоне изображения), тем лучше будет результат применения метода. Было показано, что вид альфа-канала зависит от точности подаваемого на вход программы тримапа изображения. На рис. 3 показана такая зависимость. Изображения во второй строке расположены по возрастанию детализации тримапа.



Рис. 3. Зависимость альфа-канала от степени детализации тримапа

В работе было проведено сравнение с методом интерактивной сегментации. Для этого был использован метод MSRM [1,6], который использует предварительную сегментацию изображения методом сдвига среднего, сделанную с помощью [7].

На рисунке 5-а представлена предварительная сегментация изображения и разметка изображения линиями-маркерами на передний план и фон. Видно, что уже на этапе выделения объекта MSRM “захватывает” часть изображения, принадлежащую фону, в то время как MAP-оценивание дает достаточную для естественного изображения детализацию.



Рис. 4. Применение MAP-оценивания к исходному изображению (а) для выделения переднего плана (б) и создания композитного изображения (в)

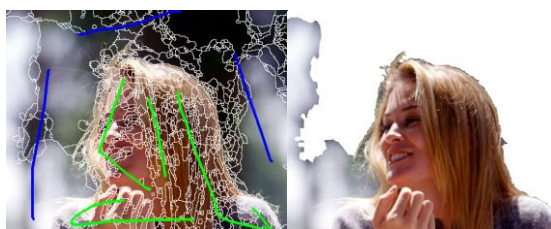


Рис. 5. Применение метода MSRM для выделения переднего плана

Выводы

Использование байесовского подхода к созданию композитных изображений показывает лучшие результаты, чем точное выделение объекта (интерактивная сегментация). При таком подходе вместе с объектом копируются и его нечеткие границы, что дает изображению большую естественность.

Литература

1. *Ning J., Zhang L., Zhang D., Wu C.* Interactive Image Segmentation by Maximal Similarity Based Region Merging // *Pattern Recognition*. – 2010. – Vol. 43. - № 2. – pp. 445-456.
2. *Tutaj K., van Reijmersdal E.A.* Effects of online advertising format and persuasion knowledge on audience reactions // *Journal of Marketing Communications*. - 2012. – Vol. 18. - №1. – P. 5-18.
3. *Chuang Y-Y., Curless B., Salesin D.H., Szeliski R.* A Bayesian Approach to Digital Matting // *Proceedings of IEEE CVPR*. – 2001 . – Vol. 2. – P. 264–271.

4. *Chuang Y-Y., Agarwala A., Curless B., Salesin D.H., Szeliski R.* Video Matting of Complex Scenes // ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2002) . – 2002. – Vol. 21. – № 3. – P. 243–248.
5. *Rubinstein M.* Bayesian Matting // Computational Photography, The Interdisciplinary Center Hertzelia. – 2009.
6. *Zhang L.* MSRM Matlab code [archived files] Hong Kong. – 2008-2015. URL: http://www4.comp.polyu.edu.hk/~cslzhang/MSRM/Matlab%20Code/MSRM_1.0.rar (accessed 13.10.2015).
7. EDISON (EDge detection and Image SegmentatiON) [html] // Robust Image Understanding Lab. New Jersey.– 2002-2015.
URL: <http://coewww.rutgers.edu/riul/research/code/EDISON/doc/ref.html> (accessed 13.10.2015).