

Восстановление изображений социальных сетей, подвергшихся обработке, на примере социальной сети Instagram

М.И. Хотилин¹, Н.С. Кравцова¹, А.В. Куприянов¹, Р.А. Парингер¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. Данная работа посвящена нахождению исходных изображений для обработанных изображений социальных сетей. Рассматриваются методы анализа цветового распределения. В работе исследован и намечен к доработке алгоритм восстановления изображений.

1. Введение

Изображения всегда занимали важное место в жизни людей: будь то наскальные рисунки в первобытном обществе, картины великих художников и особенно фотографии. С появлением социальных сетей, фотографии стали переходить из категории личного в категорию общественного достояния и стали не только новым способом самопрезентации, но и породили иные пути мышления и видения[1]. Люди буквально транслируют свое «Я»: высказывают свое мнение, загружают так называемые «сэлфи», устраивают дискуссии, делятся впечатлениями и эмоциями в комментариях к фотографиям или путем цветовой обработки фотографий при помощи разнообразных редакторов или фильтров.

Цветовая характеристика является одной из наиболее значимых характеристик, описывающих изображение. При помощи нее можно извлечь различные данные, которые могут быть полезны во многих сферах деятельности человека, например в психологии и социологии. В 2015 году Лаборатория изучения больших данных, под руководством Льва Мановича собрала и проанализировала 2,3 миллиона фотографий из социальной сети Instagram таких крупных городов, как Нью-Йорк, Токио, Бангкок, Сан-Франциско и еще одиннадцати глобальных городов. С помощью проведенного лабораторией анализа удалось увидеть «биологические ритмы», по которым живет мегаполис, а также отличается жизнь жителя одного города от жителя другого, как отличаются временные рамки и многое другое [2].

Также извлеченные данные могут быть полезны например в медицине, при выделении и распознавании объектов и во многих других областях.

Восстановление изображений является одной из наиболее значимых задач компьютерного зрения и обработки изображений. Она находит свое решение, например, при нахождении различного рода опухолей в медицине, при слежении за объектами в системах безопасности, при восстановлении изображений и во многих других. Для восстановления изображений используются различные методы, как классические, так и современные, использующие работы с нейронными сетями и инструментами обработки больших данных, или BigData.

В данной работе описан метод нахождения исходного изображения посредством определения характеристики фильтра с помощью которого он был обработан.

2. Получение исходных данных для восстановления

В ходе работы, для анализа, было принято решение использовать социальную сеть Instagram. Пользователи данной социальной сети зачастую выкладывают фотографии и изображения, предварительно обработанные теми или иными фильтрами.

Для получения исходных данных было принято решение о написании программного обеспечения, позволяющего осуществлять доступ к социальной сети и получению из нее необходимых данных, а именно непосредственно самого изображения и дополнительной служебной информации, содержащей описание и, так называемые, тэги, указывающие на используемые фильтры.

Для получения доступа к изображениям необходимо пройти аутентификацию, что требует наличия специальных ключей доступа. Аутентификация происходит посредством протокола OAuth версии 2.0. Данный протокол является открытым протоколом авторизации, который позволяет получить ограниченный доступ к защищенным ресурсам третьей стороне (нашему программному комплексу), без необходимости передачи третьей стороне (комплексу) логина и пароля пользователя. После прохождения авторизации и аутентификации появляется возможность доступа к изображениям и их описаниям и служебной информации.

3. Восстановление исходных изображений

Каждое изображение представляет собой набор пикселей, характеризующимися своими значениями цветовой яркости и все изображение можно рассматривать как трехслойную таблицу, каждый слой которой это таблица, состоящая из пикселей с определенными значениями яркости красного, зеленого и синего цветов. Изображения обработанные с применением специальных фильтров имеют цвето-яркостное распределение отличное от исходного изображения.

В ходе работы была подготовлена обучающая выборка изображений, содержащая 100 пар изображений – исходного и обработанного определенным видом фильтра, а также его название. Пример данных изображений представлен на рисунке 1.

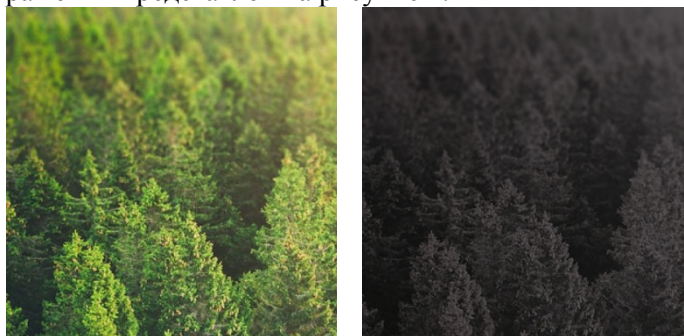


Рисунок 1. Пример используемых изображений.

Далее, используя написанное программное обеспечение для анализа, был произведен цветовой анализ изображений: вычислены распределения цветов по яркостям и построены гистограммы цветового распределения каждого из используемых фильтров. Пример исходной гистограммы цветового распределения по слоям одного изображения представлен на рисунке 2.

Аналогичным образом строится распределение и для всего класса изображений, обработанных определенным фильтром.

Далее, так как фильтр использует одни и те же преобразования яркостей, то путем вычитания из матрицы изображения, обработанного фильтром, матрицы исходного изображения, можно получить матрицу фильтра. Путем сложения всех результатов вычислений матриц фильтров для пар изображений обработанных одними и теми же фильтрами, а также путем нормировки, получаем общую матрицу для фильтра определенного типа.

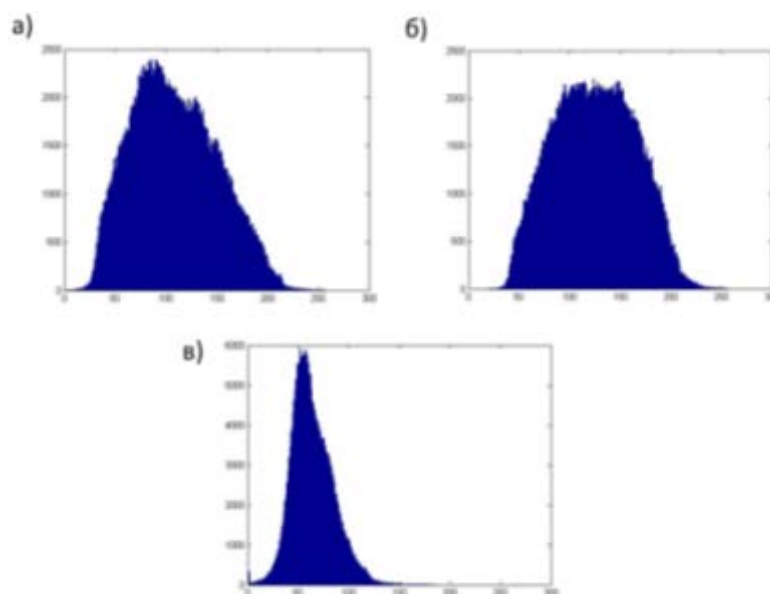


Рисунок 2. Пример цветового распределения одного изображения (а) – красный цвет, б) – зеленый, в) - синий).

Отсюда, поскольку служебная информация, полученная из социальной сети на предыдущем шаге, содержит данные о примененном фильтре, то не составляет труда восстановить исходное изображение, используя матрицу фильтра.

Для проверки данной гипотезы была организована тестовая выборка, содержащая порядка 1000 изображений, полученных из сети Instagram и проведен анализ и восстановление исходных изображений. Пример результата работы можно видеть на рисунке 3.



Рисунок 3. Пример обработанного и исходного восстановленного изображения (а) – обработанное, б) – исходное восстановленное).

4. Заключение

Определения исходных изображений является собой одну из наиболее значимых задач компьютерного зрения и обработки изображений. Она находит свое решение в различных сферах деятельности человека. Зачастую, существующие методы обработки изображений отлично работают с условно небольшими объемами исходных данных. Обработка же больших

массивов исходных изображений занимает значительное время, что в ряде задач является абсолютно неприемлемым.

В настоящее время ведется работа над адаптацией указанных выше методов для работы с данными большого объема или BigData, а также ведется разработка программного средства, позволяющего осуществлять соответствующую обработку массивов изображений и распознаванию объектов на них.

5. Литература

- [1] Манович, Л. Инстаграм – это окно в мысли и воображение человека [Электронный ресурс]. – Режим доступа: http://mediaprofi.org/community/interview/item/2492-lev_manovich_instagram (04.03.2018).
- [2] Хайкин, С. Нейронные сети: полный курс. – Издательский дом Вильямс, 2008. – 1131 с.
- [3] Сергеев, В.В. Применение методологии распознавания образов в задачах цифровой обработки изображений // Автометрия. – 1998. – Т. 2. – С.63-76.
- [4] Верхаген, К. Распознавание образов: состояние и перспективы / К. Верхаген, Р. Дёйн, Ф. Грун – М.: РиС, 1985. – 104 с.
- [5] Самаль, Д.И. Алгоритмы идентификации человека по фотопортрету на основе геометрических преобразований // Диссертация на соискание ученой степени кандидата наук. Ин-т техн. киберн. НАН Беларуси. – Минск, 2002. – 170 с.

Restoration of images of social networks that have undergone processing, on the example of the social network Instagram

M.I. Khotilin¹, N.S. Kravtsova¹, A.V. Kupriyanov¹, R.A. Paringer¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. This paper is devoted to finding the source images for the processed images of social networks. Existing methods and approaches that take place in the performance of this task are considered. The algorithms of image recovery is investigated and planned for finalization.