

# Совмещение каналов изображения

Александр Сергеев, Влад Шахуро



## Обзор задания

Первым цветным фотографом России является Михаил Сергеевич Прокудин-Горский, сделавший единственный цветной портрет Льва Толстого. Каждый его снимок представляет из себя три изображения в градациях серого, соответствующие красному, зеленому и синему цветовым каналам. Сейчас коллекция его снимков находится в американской библиотеке конгресса, сканкопии фотопластинок доступны в интернете. В данном задании мы предлагаем вам создать программу, которая будет совмещать изображения, полученные с фотопластинок Прокудина-Горского.



## Описание задания

### Базовое совмещение (3 балла)

Реализация базовой части программы реализуется в несколько этапов:

1. Загрузка изображения и разделение изображения на три канала. Достаточно разделить изображение на три равные части по высоте.
2. Удаление рамок пленки. Каждый из каналов изображения нужно обрезать на 5% с каждой стороны.
3. Поиск наилучшего сдвига для совмещения каналов. Для того, чтобы совместить два изображения, будем сдвигать одно изображение относительно другого в некоторых пределах, например, от  $-15$  до  $15$  пикселей. Далее, для перекрывающихся областей изображений посчитаем метрику. Оптимальным будет тот сдвиг, при котором метрика принимает наибольшее или наименьшее значение (в зависимости от метрики). Предлагается реализовать две метрики и выбрать ту, которая позволяет получить более качественный результат при совмещении:

(а) Среднеквадратичное отклонение для изображений  $I_1$  и  $I_2$ :

$$MSE(I_1, I_2) = \frac{1}{width \cdot height} \sum_{x,y} (I_1(x, y) - I_2(x, y))^2,$$

где  $width, height$  — ширина и высота изображений соответственно. Для нахождения оптимального сдвига нужно взять минимум по всем сдвигам.

(b) Нормализованная кросс-корреляция для изображений  $I_1$  и  $I_2$ :

$$I_1 \star I_2 = \frac{\sum_{x,y} I_1(x,y) I_2(x,y)}{\sum_{x,y} I_1(x,y) \sum_{x,y} I_2(x,y)}.$$

Для нахождения оптимального сдвига нужно взять максимум по всем сдвигам.

## Совмещение с помощью пирамиды (2 балла)

Совмещение больших изображений при базовом подходе будет проходить очень медленно. Для ускорения совмещения предлагается реализовать пирамиду изображений. В пирамиде изображений исходное изображение последовательно уменьшается в 2 раза до некоторого размера (например, чтобы обе стороны были не больше 500 пикселей в длину). Поиск оптимального сдвига начинается с самого маленького изображения, а затем на пути к исходному изображению уточняется на уменьшенных копиях изображения. Таким образом, оригинальное изображение совмещается не в диапазоне  $-15 \dots 15$  пикселей, а в меньшем, уточненном с помощью уменьшенных копий изображения.

## Интерфейс программы, данные и скрипт для тестирования

Необходимо реализовать функцию `align`, принимающую на вход изображение, полученное сканированием фотопластинки, и возвращающую совмещенное изображение. Скрипт `align_test` получает на вход две директории, прогоняет функцию `align` на файлах из первой директории и сохраняет результаты во второй.

Данные для тестирования — 6 картинок в двух вариантах: `small` и `big` для тестирования обычной реализации и реализации с пирамидой соответственно. Время работы пирамиды на одной картинке не должно превышать 20 секунд. Кроме того, обе реализации должны совмещать картинки достаточно качественно — не должно быть явно видимых сдвигов (т.е. сдвигов величиной больше 10 пикселей).

## Полезные ресурсы

[Выставка](#) о Прокудине-Горском на сайте библиотеки конгресса.