# Automatic Digicromatography:
# Colorizing the Images of the Russian Empire

Suporn Pongnumkul
Research Advisor: Prof. Alexei Efros
May 2005

## Abstract:

*A hundred years ago, before color photography existed, a Russian scholar Prokudin-Gorskii attempted to create color photographs by taking three black and white exposures of the same scene through three filters: blue, green and red, to be displayed using a special triple lens projector. Today, 1,902 three-channel negatives of these photographs survive and are available in digital format on the Library of Congress website. A color image can be reconstructed from a three-channel negative by manual alignment and careful color adjustment, termed digicromatography. However, the manual process is too time-consuming to be used on the entire collection, so an automated approach is needed.*

*In this work we have investigated several ways of automating digichromatography. After the initial automatic color alignment using the Gaussian pyramid and affine warping, there are three types of image artifacts that must be fixed: 1) local level artifacts such as extreme red/green/blue spots from dirt particles on the negatives; 2) mid-level artifacts such as gradual region color change; and 3) global artifacts such as unnatural or bleak color in the whole picture. We have applied different image processing techniques for each of these types of artifacts. Our fully automatic approach demonstrates promising results on many of the images.*

(a) (b)

Figure 1: shows a negative (a) from Prokudin Gorskii's collection, and
a manually composite image (b) of the negative by Alex Gridenko
This demonstrates the ideal goal of our work, which is to create
a program that takes a negative and produce a beautiful image like in (b)

## 1. Introduction

Sergei Mikhailovich Prokudin-Gorskii (1863-1944) was a Russian scholar who devoted his career to the advancement of photography. One of his greatest works was the documentation of the Russian Empire before the Russian Revolution. Around 1907, he formulated a plan to document the Russia in color photography to educate the Russian school children about their vast empire. Even though color printing was not available at that time, his plan was to install special projectors that would project a glass plate containing three images of the same picture, taken through red, green and blue filters, on the wall as color slides in multimedia classrooms. His plan won Tsar Nicholas II's support and Prokudin-Gorskii traveled around taking pictures of a wide range of subjects in the entire Russian territories from 1907 to 1915. The subjects of his photographs range from medieval churches and monasteries of old Russia, to the railroads and factories of an emerging industrial power, to the daily life and work of Russia's diverse population. Unfortunately, due to the Russian Revolution, he had to leave Russia in 1918. Luckily, the negatives of his photographs survived and were purchased from his heir by the Library of Congress in 1948. The Library of Congress has recently digitized the entire collection of negatives and hired a photographer to put some of them together as color pictures [3]. These pictures are fascinating and other photographers have become interested and have produced some more images from this collection. The process of manually reconstructing the color images, called "Digichromatography,"[6], produces high quality images, but is time-consuming. There are 1,902 images in this collection; hence, there is a need for good automatic reconstruction. However, the current approaches for automatically colorizing these images do not produce satisfying results. Frank Dellart [2] explored only the alignment of the negatives, leaving many artifacts in the final images to be fixed.

In the next section, we describe the artifacts in the final images after an automated alignment. We follow this with our registration process, which includes three algorithms. Then we talk about the way we approached each artifact. Next, we give a rough idea about the processing time for each picture to show that the whole collection can be processed with our work. Finally we conclude with the future work that could be done to further improve the final process.

## 2. Image Artifacts

As mentioned earlier, Frank Dellaert worked on aligning a three-channel negative to get a color image. He successfully produced color images for the complete collection. However, the resolution of the images was low, and he did not attempt to remove the defects in the final images. The defects are demonstrated in the following list:

1. There is noise and dirt on the images. Many of these images look like someone has spilled red color or blue color on the images. Since the glass plates are old, there can be scratches and dirt on the image plates, which definitely will make spots on the colorized image. For example, when the blue layer of the image has a white dirt spot, it would cause a bright blue spot to appear in the final color image after the alignment. Evidence of this is seen in Figure 2 where there are blue and red spots all over the image.

Figure 2: Picture 760 from Prokudin-Gorskii's collection rendered by Frank Dellaert's program. This demonstrates local image artifacts in the final image.

Figure 3: Picture 657 from Prokudin-Gorskii's collection rendered by Frank Dellaert's program. This demonstrates a gradual change in red color in a bottom part of the image.

2. Some of the images have gradual changes in color, causing a region of the image to have too much of one base color (red, green or blue.) This is probably caused by uneven aging of the negative. Figure 3 demonstrates a gradual change in red color at the bottom right of an image.

3. The colors in the images do not look natural. This might be because the filters that the photographer used were not perfectly red, green and blue. Figure 4 shows a comparison between two images of the same scene. The left image is manually adjusted by a professional photographer Alex Gridenko to look most natural to human's eyes. The right image is from the automated alignment, where the image is bleak and unnatural.


Figure 4: (left) Picture 458 from Prokudin-Gorskii's collection rendered by Alex Gridenko. (Right) Picture 458 from Prokudin-Gorskii's collection rendered by Frank Dellaert's program.

## 3. Image Registration

The first step of reconstructing a color image from the three-channel negative is to align the three channels. Assuming that the camera did not move between three exposures of one image, the registration of two negatives is just finding the best offsets. The offset finding can be done by SSD error minimization, which will be explained in section 3.1. However, the SSD error minimization is an iterative calculation which uses computer's memory proposition to the number of pixels in the image. Therefore, to get a registration of the high-resolution negative, which is about 70 MB per negative, we applied a Gaussian pyramid, described in section 3.2, to reduce the computation. After registration, we discovered some misalignment in some images. By observation, the misalignment is caused by some affine transformation between negatives, so we applied Shi & Tomasi's affine tracking algorithm, explained in section 3.3, to remove the affine transformation. The whole registration process takes a few seconds for low-resolution images, and takes about 8 minutes per image on Pentium M 1.70 GHz with 760MB of RAM.

### 3.1 SSD Error Minimization

A possible way to align two similar images is by exhaustive search for the best ($x$, $y$) displacement vector over a possible window of displacement. This approach assumes that an image is a translation of the other image, i.e. there is no rotation or any change in size. Here, we are looking to minimize the Sum of Squared Difference (SSD) between two channels. The algorithm is described below.

Let $F$ be a 2-D image, and $M$ be the other 2-D image of the same size. The algorithm will give $F'$ and $M'$ that are best aligned with each other:

> Let $x$ direction refers to the direction along the width of the image, and $y$ direction refers the direction along the height of the image.
>
> Let *circshift* (*im*, [*a*, *b*]) be a function that returns an image that is circularly shifted from *im* *a* pixels in $x$ direction, and $b$ pixels in $y$ direction.
>
> Let $SSD\,(im1 - im2) \;=\; \sum_y \sum_x \left(im1(x,y) - im2(x,y)\right)^2$
>
> Let $D$ be a possible window of displacement, for example $D = [-15{,}15] \times [-15{,}15]$.
>
> $$(a,b) \;=\; \arg\min_{(x,y)\in D} \left[SSD\left(F - circshift(M,[x,y])\right)\right]$$
>
> $F' = F$
>
> $M' = circshift\,(M, [a, b]\,)$

If we fix Red channel (let $F$ = Red) and align the Blue channel (let $M$ = Blue) with Red, and fix Red channel and align Green channel with Red, them put them on top of each other, we will get a color image.

### 3.2 Gaussian Pyramid

For high resolution negatives (~70MB per negative), SSD Error Minimization algorithm has too much computation to be efficient. Since SSD error minimization algorithm uses memory and time for computation propositional to the number of pixels, a big image takes too long, and sometimes cannot be completed due to the insufficiency of the computer's memory. Therefore, we use an image pyramid technique to help speeding up and reducing the memory needed. The concept of the pyramid is to do a computation on a low-resolution image first, then the higher resolution can be done with a smaller offset, and hence reduce the number of computation needed.

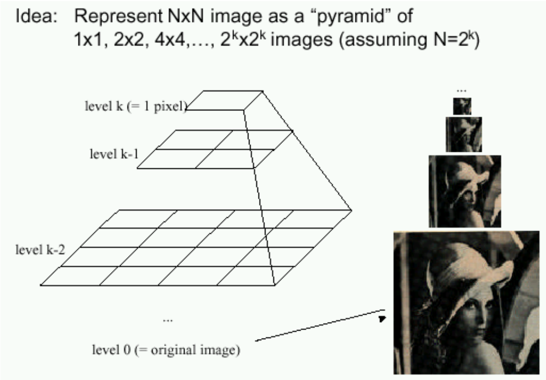
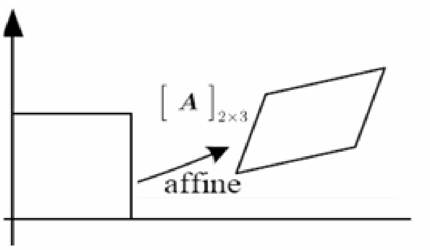


Figure 5: demonstrates the idea of image pyramid

The pyramid technique we used is the Gaussian pyramid, which is described in section 2 of [1]. "A sequence of low-pass filtered images $G_0$, $G_1$,…,$G_N$ can be obtained by repeatedly convolving a small weighting function with an image. With this technique image sample density is also decreased with each iteration so that the bandwidth is reduced in uniform one-octave steps. Sample reduction also means that the cost of computation is held to a minimum." [1]

### 3.3 Affine Tracking

Nonetheless, some problem remains for very fine resolution images. We found that, for some images, some parts of the image align well, while other parts of the same image do not well-registered. Therefore, we apply Shi & Tomasi's algorithm [7] for affine transformation to undo the affine tilting. Shi and Tomasi Tracking algorithm is a feature based tracking, that is extends to work with affine transformation. The affine transformation can be visualized as in Figure 6a, and can be mathematically described in Figure 6b.




$$\begin{bmatrix} x' \\ y' \end{bmatrix} = A \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where A is a 2x3 matrix

Figure 6: shows the idea of affine transformations

## 4. Image Enhancement

### 4.1 Hole Fixing

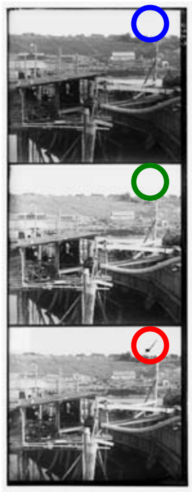


Figure 7: shows a negative with dirt particles on red channel

In this section, a hole refers to a really red/green/blue spot that appears on an image, where it does not belong. Many of this holes were shown earlier in Figure 1. Hole Fixing is intended to get rid of these holes. The idea is that the hole is caused by really high pixel values in one channel from scratches and dirt, and the corresponding spots on other two channels do not have a high value. Figure ?? shows a negative that has a blue spot causing by a black hole in the red channel. Normally, in natural scenes, colors are not pure red, green or blue, so if there is a change of pixel values in one color, which we will call an edge, there is also an edge in another color, which can differ in magnitude. Therefore, in gradient domain, we can detect a false edge in one channel by looking at other two channels.

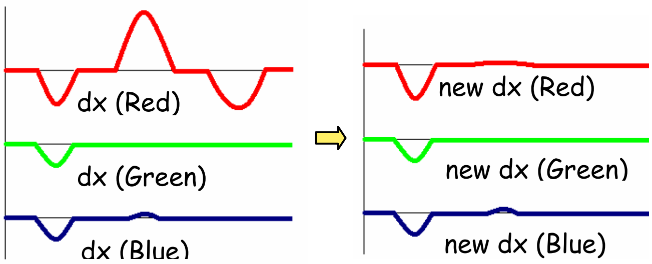



Figure 8: demonstrates the idea of Hole Fixing

The dx and dy are detected for each color channel, then if there is a huge edge (high bump) in one channel, but not the other two, we replace the dx there by the average of the dx values from other two channel. After the whole process is done, we reconstruct the image using Levin *et al.*'s reconstruction algorithm [4].

Figure 9: demonstrates the result (right image) of Hole Fixing algorithm on the left image

### 4.2 Region Color Fixing

Region Color artifacts are gradual changes in color, likely caused by uneven photograph aging. We fixed them by computing an image pyramid and applying the Hole Fixing algorithm (section 4.1) on the very coarse level of the pyramid. Figure 10a shows an image with a red region. Figure 10b shows the same image after applying the fixing. This works because if we look at the coarse level of the pyramid, the problem is just the same as Hole Fixing.

Figure 10: a (left image) shows a naïve combined image with a red region artifact
b (right image) shows the result of the left image after applying the region color correction.

**4.3 Border Region Fixing**

Considering the old glass plate negatives, the parts of images that are most likely affected are the edge of the negative. Therefore, we can try to fix the upper part of the blue channel and the lower part of the red channel. The idea is that for every patch in the lower edge, we will find the best matching patch to the green & blue channels in the center region, then we replace the center of the patch with a red pixel value calculated by the following formula:

Let B = a patch that we are looking to correct the red color.
$B_b$ = blue channel of B.
$B_g$ = green channel of B.
Let G = a patch of the same size in the good region that is most similar to patch B in green and blue channel.
$G_b$ = blue channel of G.
$G_g$ = green channel of G.
$G_r$ = red channel of G.
Let r' = new red pixel in the center

$$r' = \frac{1}{2} * \left[ \frac{\left(\sum B_b\right) * \left(\sum G_r\right)}{\left(\sum G_b\right)} + \frac{\left(\sum B_g\right) * \left(\sum G_r\right)}{\left(\sum G_g\right)} \right]$$

The motivation behind this is that there is some similar patches in the middle region, and we can learn the red channel pixel from the patch.

Figure 11: a (left image) shows a naïve combined image with red boarders
b (right image) shows the result of the left image after applying the boarder red color correction.

**4.4 Color Balancing**

Figure 4 shows that we can get a better picture from the naively aligned image. One possibility of improving an image is by improving the color balance of the images. Some assumptions are usually made for color balancing algorithm. One of the popular assumptions is the Gray World Assumption. The Gray World Assumption states that, given an image with sufficient amount of color variations, the average value of the red, green, and blue components of the image should average out to a common gray value. This assumption is in general valid since in any given real world scene, it is often the case where we have lots of different color variations.

Figure 12 shows the result of applying the gray world assumption on an image. The original image (left) is bluish. The result of the gray world assumption has less blue and looks more natural.

Figure 12: a (left image) shows a naïve combined image which looks bluish
b (right image) shows the result of the left image after applying the gray world assumption.

### 4.5 Learning Global Color Transformation

This section has the same motivation as section 4.4, which is to improve the color of the whole image. We are working on this image enhancement a learning algorithm. We have a collection of negatives which is aligned using the image registration in section 4, and the corresponding image which was manually reproduced by a specialist hired by the Library of Congress (about 120 images from 2000 images in the collection.) Since the filters that Prokudin-Gorskii used are not available to be studied anymore, it is not certain that the filters are the same as our red/green/blue filters. Therefore, we tried to model the difference in the filters from the images that the photographers believe that looks good. There are some correlations between images so we model the transformation using a 3x3 matrix *A* and a 3x1 vector *b* as follows:

$$newim = A * im + b$$

Figure 13: a (left image) shows a naïve combined image which looks redish
b (right image) shows the result of the left image after applying the learned transformation.

## 5. Results

A picture is worth a thousand words. I have 2 pictures that show the results of our work here.

Figure 14: Shows two results of our work on some two negatives. The left image is the negative. The middle image is Frank Delaerd's composition, and the right one is our result.

## 6. Conclusion & Future Work

In this paper, we have described some image registration techniques and many approach of fixing image artifacts. The result of our work shows a promising way to achieve good color images from the whole Prokudin-Gorskii's collection. However, future work can be done to improve the running time of our algorithms.

Some of the observation on techniques to improve a photograph is to brighten people in the images. This gives more emphasis on people, which generally makes a photograph more interesting. A future work can be done by detecting figures of people and lighten the blob of the figure. This should improve some images in this collection.

## Reference


[1] P. Burt and E. Adelson. *A multiresolution spline with application to image mosaics*. *ACM Trans. on Graphics*,2(4):217-236, October 1983.

[2] F. Dellaert. *Prokudin-Gorskii Images*. <http://www-2.cs.cmu.edu/~dellaert/aligned/> (retrieved 13 September 2004)

[3] A. Gridenko. *Digichromatography: Restoring Prokudin-Gorskii's Photographs*. <*http://www.gridenko.com/pg/*> (retrieved 13 September 2004).

[4] A. Levin, A. Zomet, S. Peleg, and Y. Weiss. *Seamless Image Stitching in the Gradient Domain.* Proc. of the European Conference on Computer Vision (ECCV), Prague, May 2004.

[5] Library of Congress. (9 December 2003) *The Empire that was Russia.* <http://www.loc.gov/exhibits/empire/> (retrieved 13 September 2004).

[6] Library of Congress. (9 December 2003) Making Color Images from Prokudin-Gorskii's Negatives. <http://www.loc.gov/exhibits/empire/making.html> (retrieved 13 September 2004).

[7] C. Tomasi and J. Shi, Good features to track, in *IEEE Int. Conf. Computer Vision and Pattern Recognition, 1994*, pp. 593–600.