Color Image Registration

Sai Saradha Kalidaikurichi Lakshmanan[1], Frank Merat[1]

[1]Department of Electrical Engineering and Computer Science,

Case Western Reserve University, Cleveland, OH, Email: sxk1195@case.edu

ABSTRACT

This paper aims to colorize the photographs of the early 20th century Russian empire taken by Sergei Mikhailovich Prokudin-Gorskii. The original photograph of a single scene is available as the stacked negative image of three individual color channels (R, G and B) and we colorize them by registering the red and green channel images (template) with the blue channel image (reference). While there are many standard techniques to register images, we specifically focus on a specific automatic registration algorithm with minimal human interaction, SURF (Speed Up Robust Features), and employ the technique to colorize these images. Further, we also compare and evaluate the performance of two other techniques for image registration, SIFT and Normalized Cross Correlation (NCC). To eliminate the borders in the registered image, we propose a method for automatically cropping the borders off, taking the mean as the cropping metric. Finally, we observe that even after registration, many images are not sufficiently enhanced, so we perform basic image enhancement techniques such as histogram equalization and scaling operation and study their effects on the registered image. All our experiments have been written and tested in MATLAB.

KEYWORDS

SURF, Registration, automatic cropping, SIFT, colorizing.

INTRODUCTION

Image registration is a task that frequently comes up in computer vision and AI related applications. In our paper, we look at techniques to colorize the photographs taken by Prokudin Gorskii and provided by the Library of Congress [1]. The negatives were used by Gorskii primarily for two reasons – to use in his lectures regarding the Russian empire and also to make albums of reference photographs of his journey. These negatives have different geometric alignment and intensity variations due to the fact that they have been taken separately using three different filters and geometric and intensity misalignments are unavoidable. Library of Congress has digitized these negatives by a process called "Digichromatography", where they scan the tripart glass negative, separate crop the edges and align the scan. Registration is then done by placing the blue negative at the bottom, overlaying the green and then the red image subsequently. Contrast adjustment and enhancement techniques are then followed. Registration of these images can be done manually or automatically and the registration process can be broadly categorized into two types, Feature based and Intensity based. Manual registration methods require the user to input the matching points while automatic methods identify the features without human intervention. Area or intensity based methods look at the similarity measure and determine the displacement value by looking for the pixel position and orientation that gives high similarity measure with the reference image. Some of the commonly used metrics for measuring similarity are Sum of Squared Distances (SSD) and Normalized Cross Correlation (NCC). While the primary focus of this paper is to use the SURF algorithm for registration, in the comparison section we have also analyzed and discussed the effectiveness of these other methods on colorizing the photographs. Feature based methods normally consist of the following steps – feature Detection and extraction, feature matching, transform model estimation and image resampling. Two of the most commonly used techniques here are, SIFT and SURF. In the following section we review the formulation of these algorithms.

Normalized Cross Correlation (NCC)

NCC [2] is a simpler method but is considerably effective compared to the other similarity measure methods (SAD, SSD etc.). The NCC algorithm determines the match between two images by performing a discrete cross correlation of the template image and the reference image at every possible location in the reference image. Since this is equivalent to finding the dot product between two unit vectors, the range of the normalized cross correlation value is between -1 and +1. The advantage of this method over the other intensity based methods is that, since it is normalized, NCC is less sensitive to the luminance variation. But these methods produce dense set of correspondences between the two images and can be computationally expensive.

General definition of the normalized cross correlation between two images f and g can be presented as,

$$\hat{f} = \frac{f-\bar{f}}{\sqrt{\sum(f-\bar{f})^2}} \qquad \hat{g} = \frac{g-\bar{g}}{\sqrt{\sum(g-\bar{g})^2}}$$

$$\boxed{\mathrm{NCC(f,g)} = C_{fg}(\hat{f}, \hat{g}) = \sum_{[i,j]\in R} \hat{f}(i,j)\hat{g}(i,j)}$$

Where C_{fg} the correlation is,

$$C_{fg} = \sum_{[i,j]\in R} f(i,j)g(i,j)$$

In this paper, we have implemented the NCC algorithm to register images using the image pyramid approach where large images are down-sampled into smaller size and registration is performed by coarse-to-fine alignment strategy. For analysis purpose, we used a single level pyramid where the image was only downsized to a single level and aligned this image.

Scale Invariant Feature Transform (SIFT)

SIFT is a method of extracting scale and rotation invariant features that can then be used in a wide range of applications such as registration, object detection and recognition. As described in [3], the first step in SIFT is to identify the points of interest by computing the difference of gaussian (DOG) that searches over the image to find these points. These identified points are then localized by fitting a model using the Hough transform. Orientations are assigned to the keypoints based on their gradient directions. Since the operations using the SIFT depend on these gradient directions, luminance variance is handled to a good extent. In the next step, a descriptor for representing these keypoints is designed. Once the features and the descriptor design is done, matching is done using nearest neighbor algorithm followed by a hough transform that identifies objects of the same image. Outliers are removed by the hough transform that matches the image and the model, discarding points that are not in agreement with the model.

In this paper, we have implemented the SIFT algorithm using the VLFeat library [4] and their SIFT functions. Their code on Image mosaicking [5] using SIFT has been used as a reference. The vl_sift function is used to obtain the feature points and descriptors of the images. We then use the vl_ubcmatch to create the point correspondences. Another important function that is used is the vl_imbackward function to map the points in the meshgrid.

Speed Up Robust Features (SURF)

SURF is a scale and rotation invariant point detector and descriptor that is claimed to be faster than the area based methods and SIFT as well. At a high level, from [6], Bay et all detail the SURF and how it works as – Identify the points of interest in the image using the determinant of Hessian matrix, use integral image technique to get these hessian matrices as this saves time. This matrix with thresholded determinants is called blob response map. Outliers are removed by using non-maxima suppression with the map itself and with the maps above and below it. To determine the direction of the features identified, haar transforms are used and finally features vectors are generated.

For this paper, we have focused mainly on the SURF algorithm and demonstrated that we can effectively colorize the negatives by implementing SURF. We use the functions in MATLAB for implementing the algorithm.

Automatic Cropping

Once the images are registered, noisy border is also accompanied and it would be desirable to eliminate them. We unfold our approach to eliminate the unnecessary border here. It was difficult to identify a pattern in the registered color image that could be generalized as a border across all images. We apply Otsu’s thresholding on the image to obtain a black and white thresholded image. The thresholded image carries good information about the border. The graph we then determine the distribution of the average intensities across the rows and columns in the image. As we can notice, the mean of the intensities is very low, mostly less than 0.4 along the borders. This proved to be a useful technique to eliminate the border. The algorithm can be summarized in the pseudocode below:

1. *Threshold the registered image and create a binary image (Otsu’s method, threshold value = 0.51)*
2. *Row and column size to look for border fixed as 5% of the original size of the image*
3. *Four regions of interest of the size in the previous step are selected - top, bottom, left and right*
4. *Find the mean of the rows and columns in these regions*
5. *Cropped values of the parameters are selected as follows*

   toprow and left column – Max value of row and column respectively

   bottomrow and right column= number of rows/column – min value of row/column respectively

6. *Final image = registered_image (toprow:bottomrow, leftcolumn:rightcolumn,:)*

Cropping has good effects on the image appearance. We also noticed that SURF was able to handle images that are not cropped before registration well and remove the outliers correctly, matching the correct set of points. However, with SIFT and NCC, removing the border before registration helped improve the matching. Another point to note here is that, while this method of removing the border by hardcoding the size (as 5% in our paper) to look for the border might turn out to be disadvantageous – there is a possibility of cropping the significant image region too, hence a safe value has to be chosen.

Image Enhancement

For balancing the exposure in many images, histogram equalization is performed. However, this equalization might not be suitable for all images. Only those images that are overexposed, with poor white balance have an impressive effect after equalization.

RESULTS AND DISCUSSION

The first step in any method is to separate out the three channel negatives separately. Once that is done, we colorize the images in one of the methods discussed above and qualitatively analyze the registration. In general, we perform registration by fixing the blue image as the reference image and align the red and green images separately with the blue and then overlay the three images. In the following sections, we discuss the implementation of the different algorithms discussed above and their results. Fig. 1 shows the original stacked image and fig. 2 shows the combined image that is not aligned yet.

Figure 1: Original stacked negative image - 00200u (Order of the channels - B, G, R from top to bottom)

Figure 2: Combined image (Not Aligned)

NCC

In the Normalized cross correlation method, the red and the green channels were aligned with the blue channel image by computing the L^2 norm. Instead of computing it on the original image, we smooth the image and then apply a canny edge detector to the image (shown in fig. 3). Correlation in the edge image is estimated. It can be observed that NCC is very slow when the image is very large because the search window becomes very large and it takes too long. For large images, implementing an image pyramid would be a better option. For this analysis, we reduced the size of the image to 1024 x 1024 (approx.) and aligned this downsized image. The time taken for this alignment is about 16 seconds. Doing the same operation on the whole image took about 234 seconds for the full image to be registered. Hence, it is desirable to use an image pyramid and perform coarse-to-fine alignment. Figure 4 shows the result of the image aligned in this method. Our implementation only identifies translation offset while rotation and subpixel displacement can also be identified using the same concept.

Figure 3: Canny edge on the red channel image

Figure 4: Aligned Image – Using NCC

SIFT

We used the Vlfeat library for registering the images using the SIFT algorithm. Running the SIFT on the entire image was quite slow and it took about 234 seconds for the image to be registered. Using an image pyramid here would help reduce the computation time. We used the RANSAC model to remove the outliers. Fig. 5 shows the aligned image, aligned using the SIFT functions in vlfeat library.

Figure 5: Image aligned using SIFT

SURF

SURF algorithm proved to be very fast as registering the full image took about 16 seconds. The quality of the colorized image also was better in SURF. The algorithm was so efficient in handling the discrepancies in the image and matching was done perfectly. Outliers are also efficiently handled as we can see from fig. 6, initially a number of points are detected but the outliers are removed effectively to include only the important features in the image (shown in fig. 7). The final image aligned using SURF is shown in fig. 8.

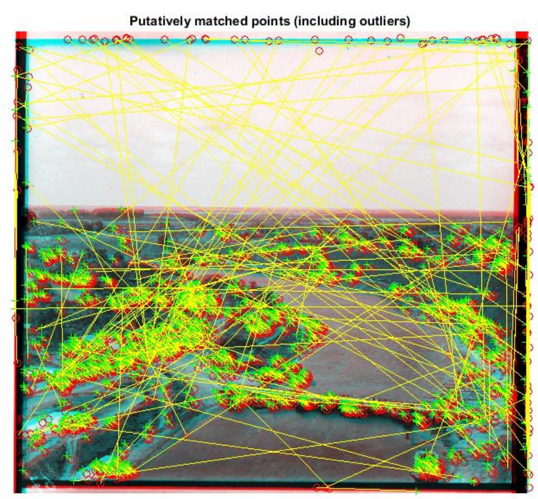

Figure 6: Keypoints detected (including outliers)

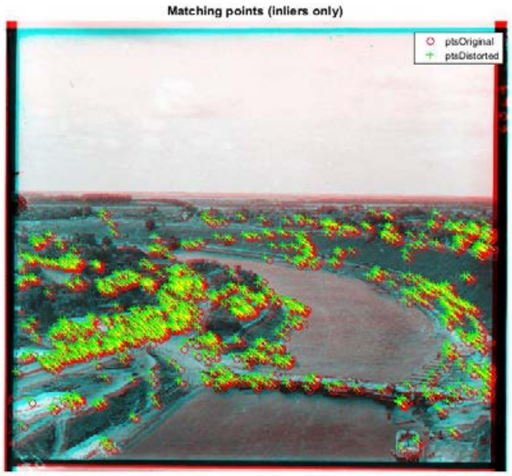

Figure 7: After removing the outliers

Figure 8: Image aligned using SURF

Automatic Cropping

Figure 9 shows the image thresholded by Otsu's method. As we can see the noisiness in the border is not uniform, but is scattered. Hence we adopted the method discussed in the previous section and as we can see, the borders are effectively removed. Comparison of fig. 8 and fig. 10 would clearly show that the borders have been removed. Fig. 11 shows the distribution of the mean intensity along the right side region of the image. We select 5% of the number of columns from the right end as our region of interest and compute the mean. As we can see from the graph, when the border begins, the mean intensity drops very close to 0. In this way, we can crop off the rows and columns that would not be required. This method though effective in most cases, is not universal as some of the images might have very small or very large border and we might not be able to adequately crop them.

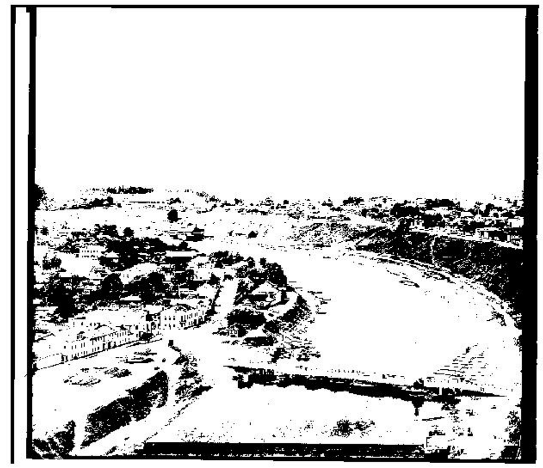

Figure 9: Threshold applied on the aligned image

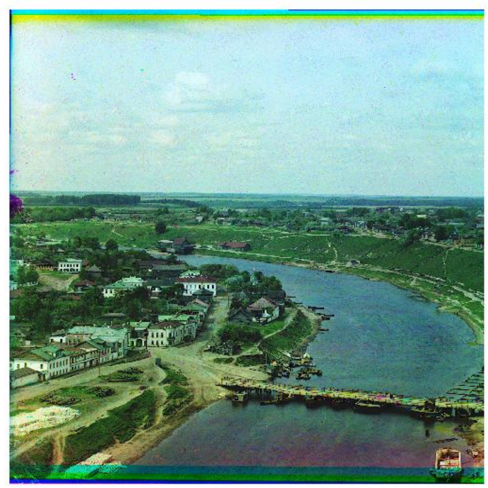

Figure 10: Cropped Image

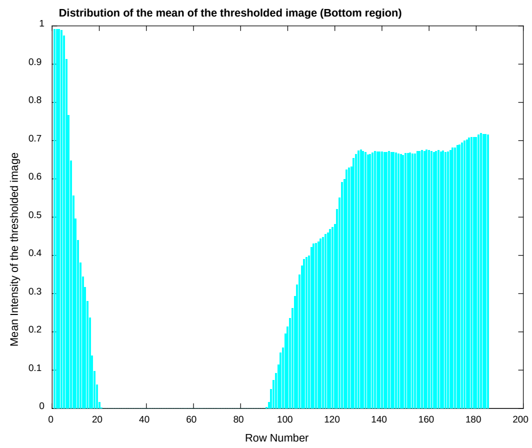

Figure 11: Mean intensity of the right columns in the image (0.5% of the total columns from the right)

Image Enhancement

As we can see from fig. 12 and 13, histogram equalization balances the exposure of the color channels. However, this equalization might not be suitable for images that are already balanced, for instance histogram equalization on fig. 18 would have no effect. Fig 14 shows the image after equalization.

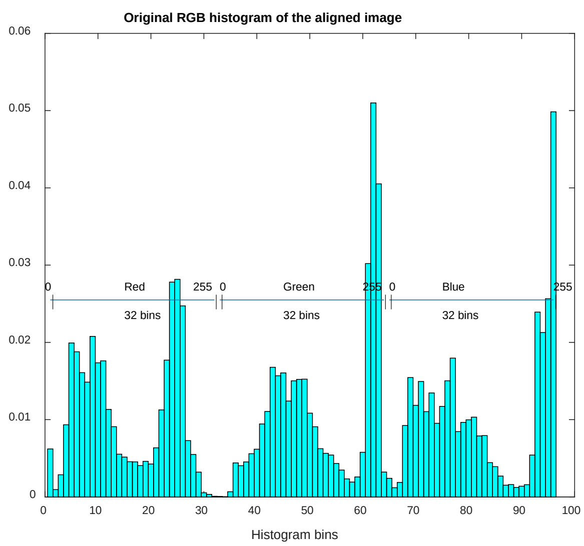

Figure 12: Concatenated histogram of the aligned image

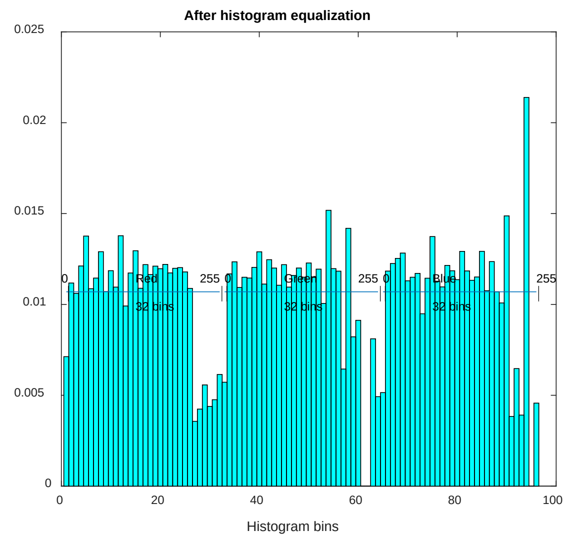

Figure 13: After histogram equalization

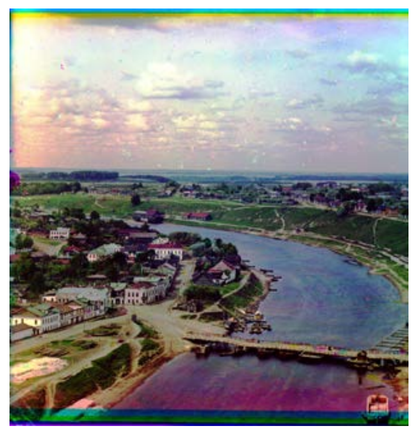

Figure 14: Image after histogram equalization

We compare the running time of the different algorithms as shown in fig. 15. SIFT (using vlfeat) takes 234 seconds on the full image while SURF takes about 16 seconds on the full image. This shows that SURF is much faster than SIFT. NCC on full image was very slow too but with image pyramid coarse alignment was achieved in about 16 seconds. This indicates that pyramidal approach would be very effective.

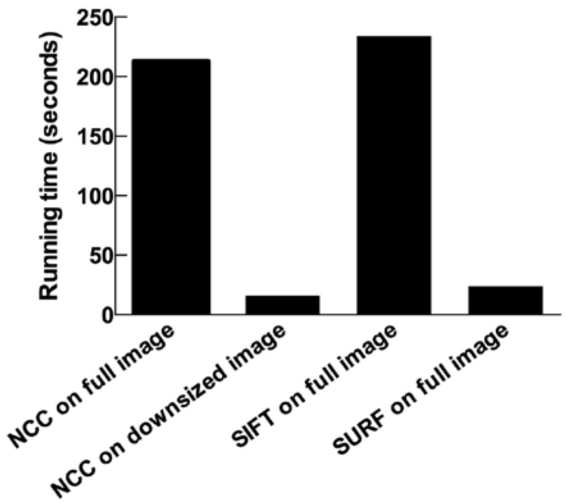

Figure 15: Comparison of running time of differenct algorithms

Figures 16 to 18 show some of the other small and large images aligned using the SURF method.

Figure 16: Aligned image - 01620

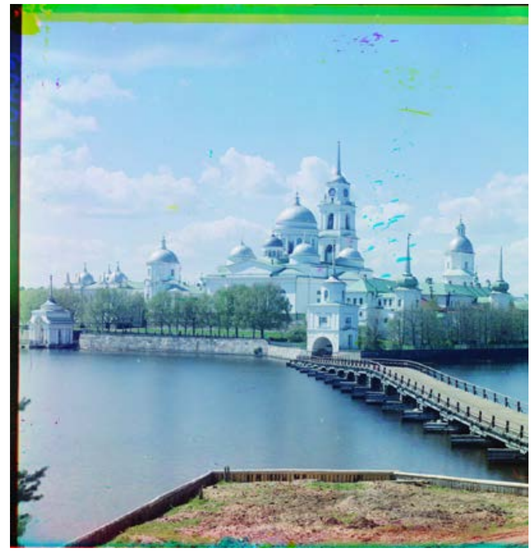

Figure 17: Aligned image - 01115

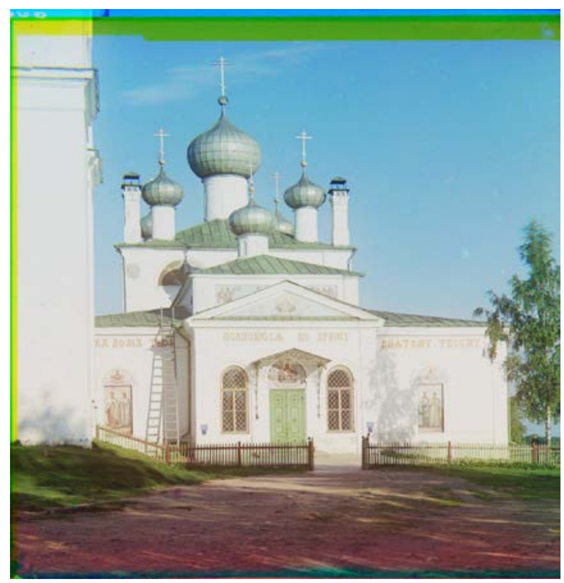

Figure 18: Aligned image - 01133

SUMMARY

Thus the Russian photographs were colorized by using the SURF algorithm. We also performed a qualitative analysis of the SIFT and NCC algorithms. Borders were automatically cropped by looking at the mean of a set of rows and columns. We also enhanced the image by performing basic enhancement operations. While our methods proved to be effective in colorizing the photographs, we notice that cropping and image enhancements did not generalize well over all images. Coming up with an automatic process of doing this would make the task much easier and efficient.

ACKNOWLEDGMENTS
This work was submitted as an assignment in EECS 490, Prof. Francis Merat's class in Digital Image Processing. Prof. Merat provided extensive support in doing this assignment. The images for our experiments were provided by the Library of Congress.

REFERENCES
[1] http://www.loc.gov/exhibits/empire/making.html.
[2]
[3] D. G. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91–110, 2004.
[4] http://www.vlfeat.org/
[5] http://www.vlfeat.org/applications/sift-mosaic-code.html
[6] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. In European Conference on Computer Vision, May 2006. 1, 2