

CS180: Intro to Computer Vision and Computational Photography — Project #1

Introduction

This project reconstructs color photographs from Sergei Prokudin-Gorskii's digitized glass-plate negatives from the early 1900s. Each glass plate contains three monochrome exposures taken through blue, green, and red filters, stacked vertically. By carefully extracting these three channels, aligning them with high precision, and combining them into a single RGB image, we can automatically produce stunning full-color photographs that bring history to life.

The pipeline handles both low-resolution JPEG images (using single-scale alignment) and high-resolution TIF scans (using multi-scale pyramid alignment for efficiency). The key challenge is achieving precise alignment despite slight camera movements, glass plate imperfections, and varying lighting conditions between exposures.

Approach Overview

The reconstruction process follows a systematic pipeline:

1. **Channel Extraction:** Split the vertically-stacked BGR image into three separate grayscale channels
2. **Preprocessing:** Apply border cropping and optional edge detection for robust feature matching
3. **Alignment:** Register the green and red channels to the blue reference using translation search
4. **Reconstruction:** Combine aligned channels into final RGB image with contrast enhancement

The alignment phase uses either single-scale exhaustive search for small images or multi-scale coarse-to-fine pyramid search for large images.

Alignment Metrics

Two primary similarity metrics were implemented and compared for channel registration:

Sum of Squared Differences (SSD)

The SSD measures pixel-wise squared differences between aligned image regions:

$$\text{SSD}(\mathbf{I}_1, \mathbf{I}_2) = \sum_x \sum_y (\mathbf{I}_1(x, y) - \mathbf{I}_2(x, y))^2$$

Lower SSD values indicate better alignment. This metric is sensitive to intensity variations but provides smooth optimization landscapes for gradient-based refinement.

Normalized Cross-Correlation (NCC)

NCC normalizes images by their magnitudes, providing invariance to linear brightness and contrast changes:

$$\text{NCC}(\mathbf{I}_1, \mathbf{I}_2) = \frac{\sum_x \sum_y \mathbf{I}_1(x, y) \mathbf{I}_2(x, y)}{\sqrt{\sum_x \sum_y \mathbf{I}_1(x, y)^2} \sqrt{\sum_x \sum_y \mathbf{I}_2(x, y)^2}}$$

Higher NCC values indicate better alignment. This metric proved more robust for historical photographs with varying exposure conditions and lighting artifacts.

Preprocessing Strategy

To improve alignment robustness, images undergo preprocessing before similarity computation. Edge detection using Sobel filters emphasizes structural features over smooth regions, while intensity normalization (zero mean, unit variance) reduces sensitivity to illumination variations.

Part 1 — Single-Scale Alignment

For low-resolution images, exhaustive search over small displacement ranges (± 15 pixels) is computationally feasible. The algorithm evaluates all possible integer translations and selects the offset that optimizes the chosen similarity metric. Border cropping ensures alignment focuses on reliable central image content rather than potentially corrupted edges.



Cathedral

G:(2, 5), R:(3, 12)



Monastery

G:(2, -3), R:(2, 3)



Tobolsk

G:(3, 3), R:(3, 6)

Part 2 — Multi-Scale Pyramid Alignment

For high-resolution images, single-scale search becomes computationally prohibitive. The pyramid approach implements a coarse-to-fine strategy that dramatically reduces complexity while maintaining precision:

1. **Pyramid Construction:** Generate multi-resolution image hierarchies through iterative $2\times$ downsampling
2. **Coarse Alignment:** Perform initial alignment at the smallest scale with large search windows
3. **Progressive Refinement:** Propagate and refine alignments at successively higher resolutions
4. **Final Registration:** Apply accumulated displacements to achieve sub-pixel precision



Church

G:(4, 25), R:(-4, 58)



Emir

G:(24, 49), R:(40, 107)



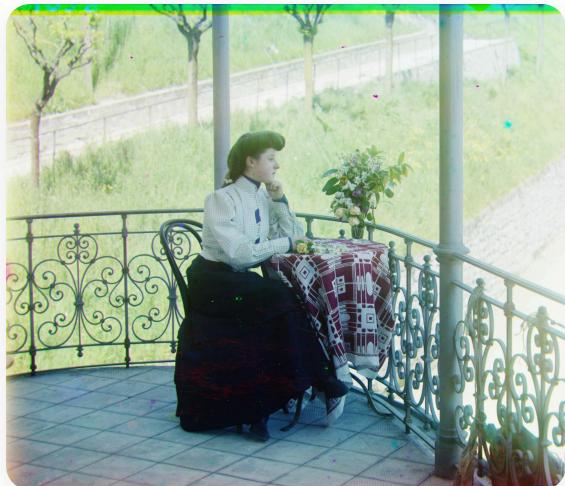
Harvesters

G:(17, 60), R:(14, 124)



Icon

G:(17, 42), R:(23, 90)



Italil

G:(22, 38), R:(36, 77)



Lastochikino

G:(-2, -3), R:(-8, 76)



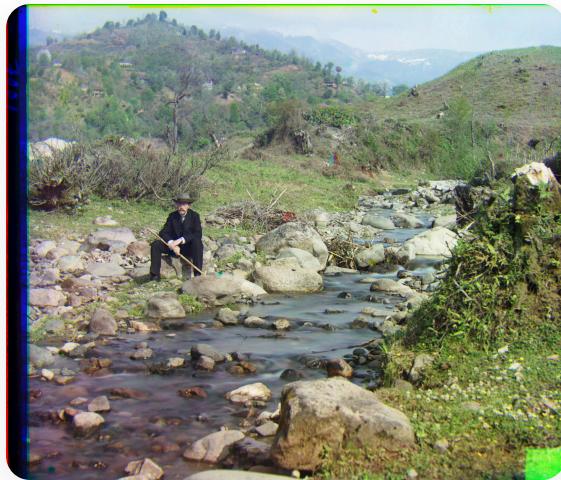
Lugano

G:(-17, 41), R:(-29, 92)



Melons

G:(10, 80), R:(13, 177)



Self Portrait

G:(29, 78), R:(37, 176)



Siren

G:(-6, 49), R:(-24, 96)



Three Generations

G:(12, 54), R:(9, 111)

Personal Image Selection

To further validate the algorithm's robustness, four additional images were selected from the Prokudin-Gorskii collection based on their diverse photographic characteristics. These examples demonstrate the pipeline's performance across different scene types, lighting conditions, and structural complexities.



Candle

G:(1, 48), R:(-7, 106)



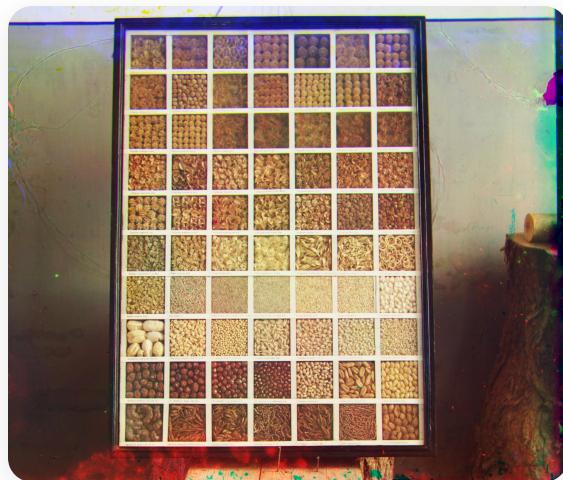
Cathedral

G:(7, 44), R:(-2, 95)



Rocks

G:(-16, 62), R:(-45, 117)



Beans

G:(-38, 42), R:(-82, 108)

{ Results and Discussion

Alignment Quality

The implemented pipeline successfully reconstructs high-quality color images from century-old glass plates. Displacement vectors show the extent of misalignment in the original captures, ranging from small corrections (2-3 pixels) for well-preserved plates to significant adjustments (>100 pixels) for damaged or shifted negatives.

Computational Efficiency

The multi-scale approach enables processing of high-resolution scans (3000+ pixels) in reasonable time. Automatic threshold-based selection between single-scale and pyramid methods ensures optimal performance across the full range of input image sizes.

Bells & Whistles — Advanced Enhancements

Beyond basic channel alignment, several automatic enhancement techniques were implemented to improve image quality and robustness without manual intervention.

Automatic Cropping

The pipeline automatically removes border artifacts through a two-stage cropping strategy. During alignment, 10% borders are cropped to focus similarity computation on reliable central content. After reconstruction, an additional 5% margin is removed from each edge to eliminate alignment artifacts and inconsistent border regions. This approach effectively detects and removes problematic border areas that often contain color fringing or misaligned content.

Automatic Contrast Enhancement

Dynamic range optimization is performed by stretching image intensities to utilize the full [0,1] range. The algorithm identifies the darkest and brightest pixel values across all channels, then linearly rescales intensities such that the minimum becomes 0 and maximum becomes 1. This automatic contrast adjustment significantly improves visual quality, especially for historical photographs with limited dynamic range.

Edge-Based Feature Alignment

Instead of relying solely on raw pixel intensities, the alignment process can utilize Sobel edge detection to emphasize structural features over uniform regions. Edge preprocessing helps the algorithm focus on geometrically relevant features like building corners, facial contours, and object boundaries, rather than being misled by smooth areas with inconsistent lighting across channels.

Combined Impact: These enhancements work synergistically to produce higher quality reconstructions with improved color fidelity and better contrast compared to basic channel alignment alone.

Search Window Parameter Analysis

The choice of search window size (`max_shift_base`) impacts alignment quality, as demonstrated by the melons image. Contrary to intuition, larger search windows don't always produce better results:



Melons (max_shift_base=15)

Poor alignment - visible color fringing



Melons (max_shift_base=50)

Correct alignment - clean registration