

Diachronic Manuscript Registration and Visualization

Stephen Parsons, Brent Seales, Seth Parker

Abstract

This paper presents a software framework for the registration and visualization of layered image sets. To demonstrate the utility of these tools, we apply them to the St. Chad Gospels, relying on images of each page of the manuscript as it appeared over time. An automated registration pipeline is used to perform nonrigid registration on each series of images. To visualize the differences between copies of the same page, a registered image viewer is constructed that allows for direct comparisons of registered images. The registration pipeline and viewer for the resulting aligned images are generalized for use with other data sets.

1. Introduction

The St. Chad Gospels is an ancient manuscript held in the Lichfield Cathedral's library in Lichfield, England. The document combines Irish and Anglo-Saxon heritage and contains the oldest surviving Old Welsh writing in its margins.

In 2010, a group of researchers from the University of Kentucky traveled to Lichfield to scan the manuscript. The team chose to study the St. Chad Gospels not only due to its historical significance, but also because its pages have been photographed multiple times since 1912. This collection of historical images offers researchers the opportunity to explore how best to organize documents of different ages diachronically, or how they have changed over time.

The Chad Gospels dataset consists of a variable number of images of each page of the document. The manuscript was photographed in 1912, 1929, 1962, 2003, and 2010, and the number of pages imaged each time totaled 9, 240, 238, 181, and 264 respectively. The 2010 set contains pictures of every page in the manuscript, including the covers and front and end matter; and it includes multispectral images of the pages that reveal details invisible to the naked eye.

The objective with this dataset is to combine the data sources, creating a complete version of the document that encompasses all known images. Each year's set of images provides unique information; but the value of that information increases when it is viewed in the context of other years' photographs, because differences can then be observed. Specifically, this project focuses on the diachronic axis of the image sets (that is, their change over time) and attempts to organize them so that it becomes possible to meticulously observe the evolution of the manuscript over time. Doing so requires pinpointing physical parts of the manuscript and visualizing their changes over the last century. The known problem of image registration is well suited to aligning the images, and a viewer was built to help make sense of the registered results. The images viewed in visual context of each other can add information beyond what a simple facsimile would provide [Seales and Crossan].

2. Related Work

2.1 Image Registration

Image registration is a well-studied problem with a wide array of applications, largely in medical imaging. Image registration is the process of aligning two images by mapping a *sensed* or moving image S into the coordinate system of a *reference* or fixed image R [Lisa Brown, Saxena, Zitová], yielding a registered image S_R . The images are of the same scene or object, but they may differ by time, modality, viewpoint, illumination, or a number of other parameters. In the case of The Chad Gospels, the 2010 photograph of each page is treated as the reference, and each additional image of that page is individually registered to the reference.

The generalized registration process can be broken down into the following steps, each of which can be implemented in a various ways: feature detection, feature matching, transform model estimation, and image resampling and transformation [Zitová]. In transform modeling, the transformation is typically either *rigid-body* (uniform affine transformations such as rotation, translation, and scaling) or *non-rigid* (also known as non-linear, local, or deformable). The appropriate method depends on the data and how it has changed between the sensed and reference images.

2.2 Registration of Manuscript Images

In 2008, the Archimedes Palimpsest project took a number of image sets, including Heiberg's 1906 photography of the palimpsest, and registered them to high-resolution multispectral photos taken in 2007. It was also a registration problem over a diachronic axis, meaning the images displayed a change over time. This manual registration process involved scaling and warping [Toth and Emery], and presumably relied upon an affine transform, meaning it would not have accounted for any non-linear warping that might have occurred between images.

In 2009, Baumann and Seales applied registration tools to manuscript pages in an effort to overcome a problematic camera sensor. While creating a set that included a photo and a 3D scan of each page (taken simultaneously), they discovered that a dirty sensor on the camera had produced some images of unacceptable quality. They rephotographed those pages, but, due to time constraints, they were unable to perform new 3D scans of the pages. Instead, they registered the high-quality sensed images to the "dirty" reference photos taken earlier. Doing so allowed them to texture the 3D scans with the high-quality images and then perform digital flattening techniques to remove imperfections in the images due to warped pages [Baumann and Seales]. They used a triangle-based warping, which was an improvement over the affine transform used with the Archimedes Palimpsest.

These projects show a marked improvement in the preservation and visualization of ancient manuscripts, but leave room for further work. To be most practical, a diachronic edition of a manuscript should be complete, well-aligned with deformable registration, and possess an interactive and useful interface.

2.3 Visualization of Registered Images

Making use of diachronic registration requires a method for viewing the reference image R and the registered image S_R so their differences become readily apparent. In many cases, an observer wants to see how a specific part of the manuscript has changed over time. Such discoveries can be made by identifying a point on R and asking what the corresponding point looks like on S .

The most obvious way to accomplish such comparison is to examine R and S_R side by side, e.g. each occupying half of a screen. While this method places both images directly in front of the observer at the same time, it remains difficult to directly compare the specific points of interest. In switching attention from one image to the other, the observer loses their visual context and position. This method does not make use of the images being registered to the same coordinate system, and an observer could just as well use R and the original sensed image S .

This limitation can be overcome by “flipping” between R and S_R . The observer can view one image and then switch the display to show the other in its place. Since S_R is registered to R , a point of interest will occupy the same spot on the screen regardless of which image is being viewed. By focusing on one point and switching back and forth repeatedly, an observer is able to see how that area changes across the images. However, entirely switching out the images is still overwhelming to the observer, who is tasked with mentally tracking the points and trying to remember what they looked like on the other image. The result is a confusing visualization that does not clearly illustrate differences.

Composite images are one of the most common methods for visualizing registered image data, and they have been proven to overcome some of the above mentioned challenges. This method combines the two images into a third image that contains information from both. A simple composite image is the difference $|R - S_R|$, or “diff,” which is simply the absolute difference between the images at each pixel. Where the images are identical, meaning no change is revealed across registered images, the diff image is dark. Bright spots indicate a difference between the images that could then be examined on either image. More complex composite images exist that offer an enhanced visualization when compared to that of the absolute diff [Wallace, Rehm]. However, no composite image can ever fully display the original data, since the combining process reduces the data of multiple images into one impression.

3. Approach

3.1 Image Registration

The registration task for the Chad Gospels dataset was to produce a set $P = \{R, S_{R1}, S_{R2}, \dots, S_{Rn}\}$ of aligned images for each page of the manuscript, where R is the 2010 image and $0 \leq n \leq 4$ is the number of sensed images for that page, with each S_{Ri} from 1912, 1929, 1962, or 2003. This set was produced by beginning with the 2010 image R for each page, and generating a pair (S, R) for each non-2010 image S of the

same page. Each pair (S, R) was then passed through the registration pipeline to produce an S_R , which were used to produce P for each page.

3.1.1 Feature Detection and Matching

The nature of The Chad Gospels dataset made it possible to manually select matching feature points for each pair of sensed and reference images. The dataset is small enough to make this tractable, and the consistent image structure allows for a straightforward process.

For each image pair (S, R) to be registered, a team member used a visual interface to select five points on each image: each of the four corners, plus a more central point. The corners were identified as the actual corners of the physical manuscript page, while the central point was chosen to be some identifiable feature, such as the tip of the serif on a central letter. This process resulted in a set of landmarks $L_{S,R}$, mapping pixel indices on S to their corresponding indices on R :

$$L_{S,R} = \{((x_{i_S}, y_{i_S}), (x_{i_R}, y_{i_R})) \mid 0 \leq i < l\}, \quad l = \# \text{ of landmarks for pair } (S, R)$$

Occasionally after registration, particular images exhibited regions that did not converge to alignment. In these cases, it was possible to manually add additional feature points in the problematic region, which proved adequate for improving the resulting registration.

The use of automated feature detection and matching, using an algorithm such as Scale Invariant Feature Transform (SIFT) [Lowe], was explored but ultimately proved unnecessary for this particular set of images. However, should the same methods be applied to an image set of a larger scale or with more variation in content, an automated approach would be appropriate.

3.1.2 Registration

Registration methods vary considerably, and their utilization depends on the underlying data. For example, in some cases it may not be appropriate to introduce nonlinear transformations to a sensed image. In our study, however, such deformation is considered acceptable due to the value of registering letterforms on the page to their counterparts in images of the same page taken at different times. The Chad Gospels photographs were taken with different camera sensors and lenses, from different distances and angles, and the pages are warped differently each time they are opened. These subtle deformations combine to create a registration problem that not only allows, but actually requires, nonlinear transformations in order to achieve the desired effect.

Our registration framework uses the Insight Segmentation and Registration Toolkit (ITK) [ITK] and consists of landmark warping and deformable registration stages. The process begins with the input images S and R and the landmarks $L_{S,R}$. The pixel indices in $L_{S,R}$ are converted to points in the physical space used for registration in ITK.

The images are converted to grayscale for the registration process, which makes it easier to compute a similarity metric at each iteration:

$$S' = \text{grayscale}(S), \quad R' = \text{grayscale}(R)$$

For the landmark warping stage, an initial transformation is applied to S' that aligns each of its landmark points with its corresponding point on R . The remaining points on the images are interpolated using a displacement field built by a kernel-based spline function. This transformation T_1 creates S_{LW}' , which is roughly aligned with the reference image R in size and space. While the landmark feature points are all registered perfectly at this point, the rest of the page remains poorly aligned. Due to the nature of the deformation, different regions of the page are misaligned independently of each other.

$$\text{Landmark warping stage: } S' \xrightarrow{T_1} S_{LW}'$$

The deformable registration stage is then performed using a regular step gradient descent optimization with a relaxation factor of 0.85. The deformation is computed using a basis spline or B-spline transform with spline order 3. The alignment is evaluated using the Mattes mutual information metric [Mattes] with 50 histogram bins and the number of samples at each iteration being 1/80 the number of pixels in the reference image. This sampling volume can be increased for more accurate registration at the cost of registration program runtime.

There are multiple possible stop conditions in this configuration:

- The maximum number of iterations (default 100) is reached.
- The optimizer step size falls below the threshold value.
- The change in metric between successive iterations falls below the threshold value.

The deformable registration stage yields the transformation T_2 , which gives a grayscale registered image S_R' from S_{LW}' :

$$\text{Deformable registration stage: } S_{LW}' \xrightarrow{T_2} S_R'$$

The full color registered image is generated by applying the transformations to the original sensed image:

$$\text{Registration process: } S \xrightarrow{T_1+T_2} S_R$$

The transformation $T = T_1 + T_2$ is also recorded in a file.

3.2 Visualization and Preservation of Registered Images

3.2.1 Visualization

A layered image viewer was built for the visualization of registered image sets. This tool represents an improvement over previous methods by allowing for the direct

comparison of original images while still maintaining their visual context. The viewer consists of a background image and a foreground image. The background image is featured prominently in the interface. A circular “flashlight” then allows a user to “shine” the foreground image onto the background, which is aligned with the foreground image. In Figure 1, the “background” image is from 2010 and the “foreground” image seen with the flashlight is from 1962:

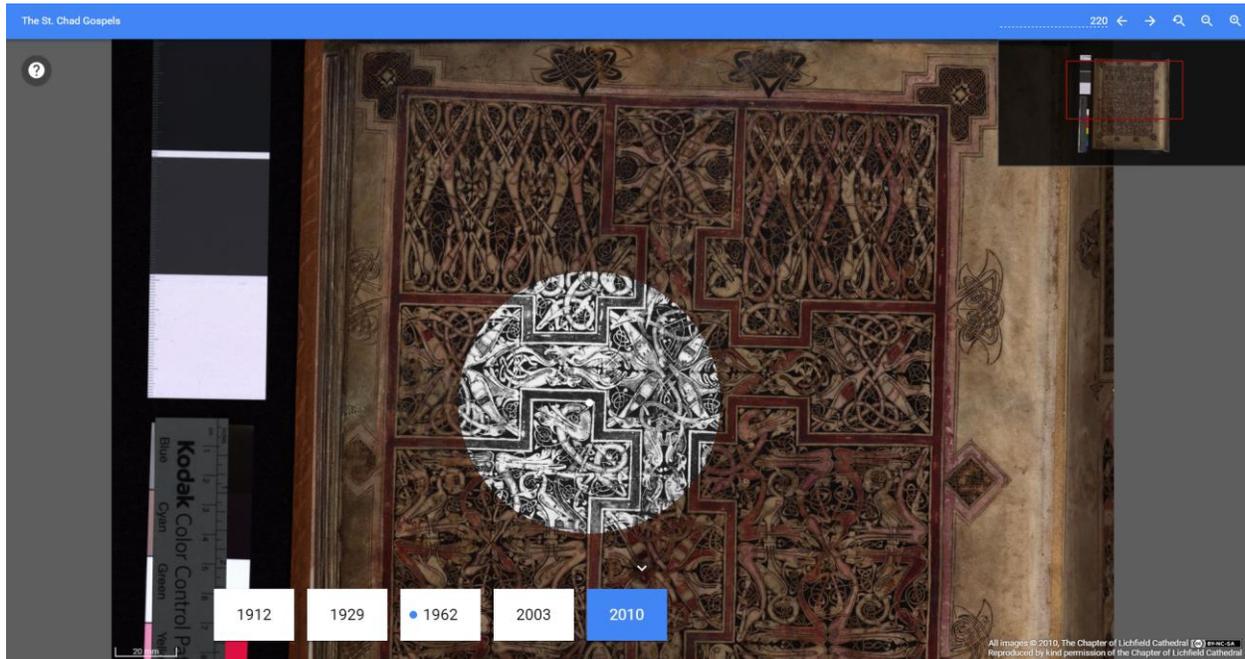


Figure 1: Registered image viewer.

The viewer understands a dataset as a set of pages, each page being comprised of the aforementioned mutually registered set P . The following is an example of the JavaScript Object Notation (JSON) file format accepted by the viewer for the visualization of a registered dataset:

```
{
  "pages": [
    {
      "name": "Front Cover",
      "layers": [
        {
          "dzi": "data/dzi/2003_Reg2010/ChadGospels-001-FrontCover-2003_Reg2010.dzi",
          "version": "2003",
          "pixelsPerMeter": 16033.33
        },
        {
          "dzi": "data/dzi/2010/ChadGospels-001-FrontCover-2010.dzi",
          "version": "2010",
          "pixelsPerMeter": 16033.33
        }
      ]
    },
    ...
  ]
}
```

While only two images are shown at a time, keyboard shortcuts allow quick switching of the foreground and background images. Other commands enable rotating through the various pages of the manuscript. The flashlight can also be resized and is of variable shape. Together these functions allow a user to quickly navigate the pages of the manuscript and pinpoint regions of interest. This is an improvement over the aforementioned methods of viewing registered images, which are not well-suited for more than two layers.

The viewer uses the Deep Zoom Image (DZI) [Microsoft] tiled image format so that the user can zoom in and achieve high resolution without needing to load the entire full resolution images in the beginning. OpenSeadragon [OpenSeadragon] is used to display the DZIs. The JSON file describing the dataset simply points the viewer to the appropriate DZI files and directories, which can be hosted locally or elsewhere. The viewer hosted at infoforest.vis.uky.edu shows The Chad Gospels, but the viewer can be used with other datasets simply by cloning the repository and using another JSON file to describe the dataset.

3.2.2 Preservation

The Internet Archive [archive.org] was chosen as a long-term repository for our diachronically registered images of The Chad Gospels for a number of reasons. As the name implies, the Internet Archive is designed for long term preservation. It also has a data ingestion pipeline that processes the images and makes them publicly available for download in various formats. Uploads were performed with a command line tool from archive.org that allows for batch transfer specified by a comma-separated values (CSV) file.

4. Results

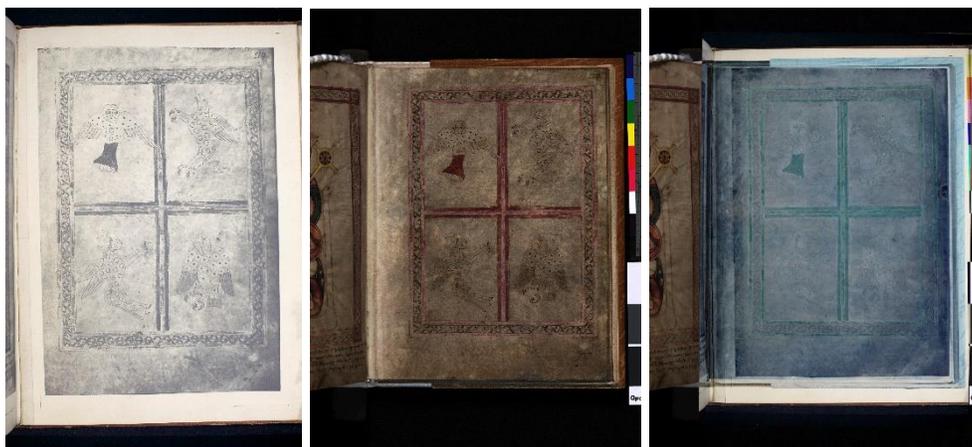


Figure 2: Page 219. (a) 1929 image. (b) 2010 image. (c) Diff after registration.

Results of the registration process for the alignment of Chad Gospels pages are shown in Figure 2, with page 219 used as an example. Diachronic registration was

successful for pages with text as well as for those containing illuminations. These registered images along with the layered image viewer allow researchers to quickly navigate the document and identify and examine points of interest that reveal diachronic changes in the artifact. For example, Figure 3 reveals pigment that has chipped off between the years 1929 and 1962, and again from 1962 to 2003.

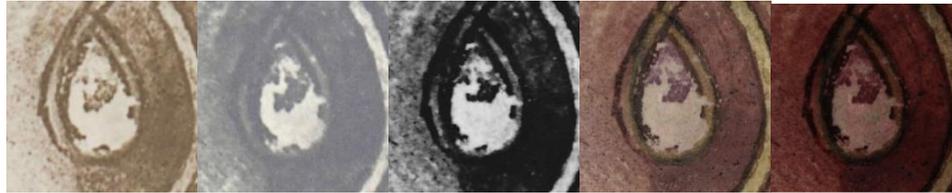


Figure 3: Pigment chipping on page 218. (a) 1912. (b) 1929. (c) 1962. (d) 2003. (e) 2010.

In some cases, the registration process failed to yield perfect alignment because the original pages were so severely warped. Where a misalignment was deemed unacceptable, it was corrected by manually adding landmark points in the offending region and then re-registering that image pair, such as with page 234 in the below Figure 4. The text is initially brightly outlined where it is not aligned, but as features are added to help correct this, this bright outline melts away.

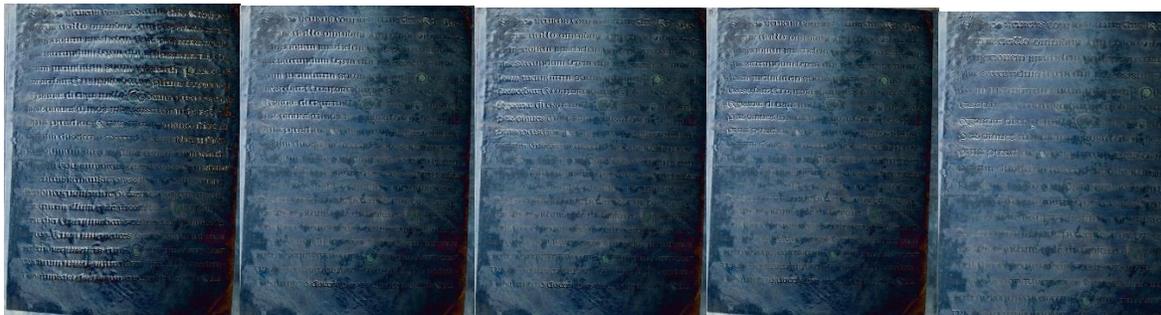


Figure 4: Successive diffs of results for page 234, as landmarks are manually added for correction.

5. Conclusions and Future Work

High resolution imaging of manuscript pages is an emerging method for improving document longevity and accessibility. In addition to providing facsimile resources, scans of these pages offer new opportunities for discovery about the original ancient manuscripts. Multimodal and diachronic image sets, for example, can reveal patterns or changes that were not visible to the naked eye working only with the physical manuscript or a digital facsimile.

We have presented a method for the organization, registration, and visualization of diachronic manuscript image sets. Registered images combined with the viewer allow researchers to quickly yet precisely navigate a manuscript, search for changes that have occurred over time, and then pinpoint those for a closer inspection. We believe these registration and visualization methods will prove to be powerful tools not only for

the St. Chad Gospels, but for other manuscripts and, more generally, other datasets with images showing some change, whether in time, modality, or other elements.

We imagine several opportunities for future work. The registration process could be fine-tuned to completely eliminate problematic areas that are not perfectly aligned. In this work, the alignment was considered successful in the context of comparing images of text over time. For other applications, the registration may need improvement. The feature detection and matching process could be automated, a step that was not necessary for this data but would be worthwhile with a larger or more varied set of images. As 3D data exists for these pages, digital flattening techniques could be applied to the 2010 images before using them as reference images. This step would remove any imperfections or warps due to the physical warping of the manuscript pages under the camera. By registering sensed images to a flattened reference image, the sensed images would have this corrected as well. Finally, the registration process and viewer can be extended to a variety of other datasets.

This project builds on and anticipates other work that goes beyond the simple advancement of digital techniques used. Ultimately, the methods we describe lead to the improved preservation and understanding of ancient documents, adding to an old body of knowledge in ways that were not before possible.

6. Acknowledgements

Chapter of Lichfield Cathedral, University of Kentucky Center for Visualization and Virtual Environments, Furman University Department of Classics

National Science Foundation Grants No. IIS-0535003, 0916148, 0916421, and EAGER-1041949

7. References

Ryan Baumann and W. Brent Seales. 2009. Robust registration of manuscript images. In Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries (JCDL '09). ACM, New York, NY, USA.

Lisa Gottesfeld Brown. 1992. A survey of image registration techniques. ACM Comput. Surv. 24, 4 (December 1992).

Michael S. Brown, Mingxuan Sun, Ruigang Yang, Lin Yun, and W. Brent Seales. 2007. Restoring 2D Content from Distorted Documents. IEEE Trans. Pattern Anal. Mach. Intell. 29, 11 (November 2007).

Michael S. Brown and W. Brent Seales. 2004. Image Restoration of Arbitrarily Warped Documents. IEEE Trans. Pattern Anal. Mach. Intell. 26, 10 (October 2004).

"The Insight Segmentation and Registration Toolkit." www.itk.org.

"The Internet Archive." www.archive.org.

Rehm, Kelly et al. "Display of merged multimodality brain images using interleaved pixels with independent color scales." Journal of Nuclear Medicine 35.11 (1994).

David G. Lowe. 1999. "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision. pp. 1150–1157. doi:10.1109/ICCV.1999.790410.

D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. 2003. PET-CT image registration in the chest using free-form deformations. IEEE Transactions on Medical Imaging. 22, 1 (January 2003).

Microsoft. Deep Zoom Image format.

R. Netz and W. Noel. *The Archimedes Codex*. Da Capo Press, 2007.

"OpenSeadragon." OpenSeadragon.

W. Brent Seales and Steve Crossan. 2012. Asset digitization: moving beyond facsimile. In SIGGRAPH Asia 2012 Technical Briefs (SA '12).

Saxena, Siddharth, and Rajeev Kumar Singh. A survey of recent and classical image registration methods. Target 7.4 (2014).

M. Toth and D. Emery. Infinite possibilities: Archimedes on the web.

D.P. Wallace, Y.-P. Hou, Z.L. Huang, E. Nivens, L. Savinkova, T. Yamaguchi, M. Bilgen, Tracking kidney volume in mice with polycystic kidney disease by magnetic resonance imaging, *Kidney International*, Volume 73, Issue 6, 2 March 2008.

Barbara Zitová, Jan Flusser, Image registration methods: a survey, *Image and Vision Computing*, Volume 21, Issue 11, October 2003.