



Improving Image Registration Accuracy through User Interaction

Vincent Groen

**Supervisors: Ruben Wiersma, Ricardo Guerra Marroquim, Elmar Eisemann
EEMCS, Delft University of Technology, The Netherlands**

23-6-2022

**A Dissertation Submitted to EEMCS faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering**

Improving Image Registration Accuracy through User Interaction

Vincent Groen¹

Supervisor(s): Ruben Wiersma¹, Ricardo Guerra Marroquim¹, Elmar Eisemann¹

¹EEMCS, Delft University of Technology, The Netherlands

Abstract

The current state-of-the-art image alignment techniques and their input parameters are often unintuitive to those without the required background knowledge. This work aims to provide users with a graphical user interface through which they can intuitively influence the parameters and results of the algorithm by excluding certain areas from the images. The application is also capable of automatically optimizing the parameters of the underlying registration algorithm. Experimental results indicate that the automatic parameter optimization results in alignments at least as good as the default parameters, but that the introduced user interaction does not provide any major benefits.

1 Introduction

Modern research into classical paintings uses various non-invasive imaging techniques. These scans capture details about the composition of the painting in different modalities without damaging it. An example of this is the recent study into *Girl with a Pearl Earring* by Vandivere et al. [2019].

The next step is registering all the scans to a base scan of the painting, which is usually a colour image. Although recent studies such as the ones by Klein et al. [2010], Farmer [2013] and Conover et al. [2015] have mostly focussed on the automated image registration process, and more specifically the *control-point selection*, manual registration by experts has the potential to yield more accurate results than fully automated registration in certain settings [Hedlund et al., 2010; Nakajima et al., 2020].

In this paper, an application is presented which involves a human operator in the state-of-the-art algorithm for registering and mosaicking images presented by Conover et al. [2015]. Though Conover et al. provide an excellent image registration algorithm, there are some shortcomings with it in practice. First is that the parameters of the algorithm can be hard to understand for the user, leading to suboptimal registration results. Second is that the algorithm works on the entire image with the same parameters, this can result in features not being found in the desired locations.

The goal of the application presented by this paper is to provide an abstraction layer for the state-of-the-art, imple-

mented as a *Graphical User Interface* (GUI). Through this GUI, users can influence the parameters of the algorithm to address the previously mentioned flaws. The parameters can be influenced with several user interactions, such as asking the user for clarification in the *feature matching* step and having the user specify which regions should be aligned to the other image. This paper implements both methods to some extent in an application where the user manually selects *regions of interest* and can subsequently adjust or remove the found *control-point* pairs.

The structure of this paper is as follows. First, the relevant problem will be explored in more detail in Section 2, this section will also provide the necessary background information. Section 4 presents the main application and after that, Section 5 describes the evaluation of the presented application and discusses the results. Section 6 will reflect on the ethical aspects of the research and finally, Section 7 summarizes the work and provides some possible directions for future research.

2 Background

The most prominent methods for image registration are covered in literature by Brown [1992] and more recently by Zitová and Flusser [2003]. Both papers categorize the sub-steps of the registration process in the following groups, *feature detection*, *feature matching*, *transform model estimation* and *image resampling and transformation*. The combination of the first two groups will be referred to as the *control-point selection* step in this paper.

The algorithm presented by Conover et al. [2015] consists of several steps, the most important of which are highlighted here. First, it finds features in the entire template image, often based on the craquelure of the painting since these features are generally visible across modalities. After this step, regions that should be close to these features in the reference image are searched. The search strategy that Conover et al. employ is based on the phase correlation algorithm [Gonzalez and Woods, 2008] and uses the phase images of both the reference and the template [Conover et al., 2011, 2015]. If a valid candidate is found within such a region, it is paired with the feature in the same approximate location in the template image. These pairs are also referred to as *control-point* pairs. Both of these steps act on some fixed parameters, which are static during the runtime of the algorithm. This can cause

problems if the images are particularly difficult to align. It could be the case that no matching features are found in certain regions of a painting or that both images contain some completely different features, which could happen with X-Ray scans where the frame of the canvas is visible.

3 Related Work

Graphical user interfaces for image registration

Elastix is an image registration toolkit with the primary purpose being the medical field which was introduced by Klein et al. [2010]. Though this tool usually requires some programming knowledge to be used, there have been some GUI extensions developed for it. Notable are *SlicerElastix* [Lasso, 2022] and *elastix_napari* [van der Valk et al., 2021]. However, these tools are not always applicable to every use-case. SlicerElastix is an extension to the Slicer3D software¹, which is used to register three-dimensional volumes from medical scans. Elastix_napari does work on two-dimensional data and provides several presets for different transformation techniques, but these are not always applicable. For the parameters of certain datasets, the user will have to either manually set them or download them from the *Model Zoo*², which provides good parameters for several medical datasets. Neither option will work great for art registration by users who are not concerned with the technical aspects of the underlying algorithm.

Photoshop [2022] and *PTGui* [2022] both provide some features that allow users to align and register images, but it is the main purpose of neither program. Photoshop is a general purpose image editor and can certainly be useful as image registration software, but it is unlikely that its auto align features are specialized for multi-modal scans of classical works of art. PTGui suffers from a similar problem, as it is a software used for stitching normal colour pictures into panoramas.

4 Method

This work provides an abstraction layer for the algorithm presented by Conover et al. [2015]. This abstraction layer, which is presented as a *Graphical User Interface* (GUI) and can be seen in Figure 1, allows the user to influence the algorithm and view the results without requiring any prior knowledge on the inner workings of the original work. The user input is used to create a mask to exclude certain areas from the image registration algorithm. A genetic algorithm optimizes the parameters to find those that result in the best alignment by executing the algorithm a number of times while mutating the current best set of parameters. It should be noted that this process does not guarantee optimality, due to the relatively long execution time of the image registration. Below, the contributions of this work are split in two parts and explained in more detail. The first part introduces the algorithm responsible for optimizing the parameters of the image registration algorithm and the second part presents the GUI and how it is used to affect the optimization algorithm.

¹<https://www.slicer.org/>

²<https://elastix.lumc.nl/modelzoo/>

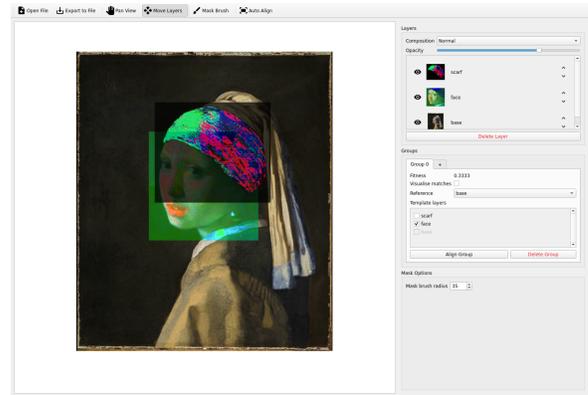


Figure 1: The GUI presented in this paper. In the screenshot, three images are loaded and can be influenced as independent layers, the user also aligned the active group but disabled the visualization of matches.

4.1 Parameter Optimization

The image registration parameters are tuned in an iterative process that aims to yield the best alignment. In this process, the program executes the registration algorithm by Conover et al. [2015] in two phases and applies only to the parts of the *template* image that are not covered by the mask.

First, the algorithm optimizes the parameters related to the *feature size* and *feature density*. It does this by performing the steps of the algorithm by Conover et al. that correspond to identifying the initial set of *control-points*. If a certain set of input parameters results in a greater number of control-points, these parameters will be ranked higher. The input parameters and the resulting control-points are ordered, decreasingly, by the number of control-points. After this step, only the top ranking input parameters are passed on to the next phase, since having more control-points likely results in more control-point pairs making it through the next phase.

The next phase of the algorithm uses the parameters that yielded the most feature points and attempts to find their corresponding locations in the *reference* image, any points that it can not match to the reference image are discarded. The resulting control-point pairs are used in the final step of the algorithm, which finds a bilinear mapping from the feature points in the template image to their counterparts in the reference image by finding a bilinear function which fits to the points with an iteratively reducing the threshold. Results of this phase are evaluated by how many control-point pairs were found and the final threshold of the mapping. The order is firstly defined by the lowest threshold, followed by the greatest number of control-point pairs.

4.2 Graphical User Interface

The GUI provides the user with several ways to interact with the algorithm and visualizes the results of their actions. Due to the limited toolset of the application, it is possible to annotate all options with icons and text and not hide any in sub-menu's. It is, in some features, inspired by existing image ma-

nipulation programs such as Adobe Photoshop³ and GIMP⁴. In the following paragraphs, the most important features of the GUI are presented.

Importing and exporting images supports the *PNG* and *TIFF* formats. Both formats are lossless and therefore logical choices for highly detailed images. The *TIFF* format is mainly used to allow importing and exporting multi-layer images.

The brush tool allows the user to directly alter the mask applied to the object before aligning the images. This way, the user can focus the registration on specific parts of the image and prevent other perhaps more detailed parts of the image from dominating the control-points. An example of this can be seen in Figure 2, the right image shows that the algorithm found most control-point pairs in an area seemingly devoid of features. To counteract this behaviour, the user can paint a mask on the undesired areas which prevents the program from finding any control-points there.

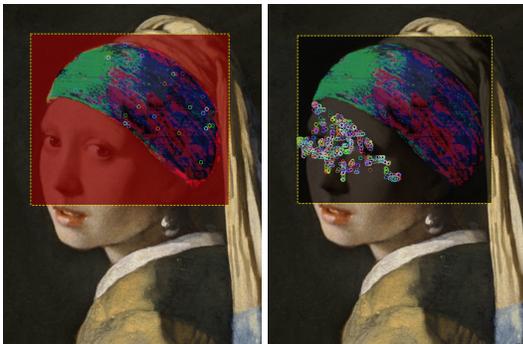


Figure 2: Image alignment with and without user interaction, the left image shows the area that the user excluded from the registration in red. The coloured dots represent the control-points that the algorithm found.

Image alignment is applied to *alignment groups*, which are collections of a single reference layer and one or more template layers. Each group is independent of any other groups and can be used to quickly align multiple templates to the same reference.

5 Experiments

Several experiments are executed to evaluate and compare the performance of the application with and without user interaction. The following sections go over the setup, motivations, and conclusions for each of these experiments.

5.1 Implementation

All experiments use the application that this paper presents, which internally uses an implementation of the algorithm by Conover et al. [2015]. The different setups compare the accuracy of the algorithm when used through the application in various scenarios. First a baseline measurement will be taken by executing the algorithm by Conover et al. with some

³<https://www.adobe.com/products/photoshop.html>

⁴<https://www.gimp.org>

arbitrary default parameters on the entire template image, to simulate the experience of using the algorithm without setting any parameters. This will be compared to a second measurement in which the genetic optimization algorithm is used and a third measurement where the GUI is used to focus the registration on a specific region.

Each experiment will also define a region of interest, where the alignment should be at its best. Results are evaluated by manually comparing the alignment in these regions of interest. Though this is a qualitative metric, it provides insight into the relative performance of the user guided algorithm over the classical algorithm. If the alignment using the optimized parameters is notably better than the that of the default parameters, then this indicates that the application is helpful but not necessary. However, when the user guided results are better than those achieved with the optimized parameters then the application has successfully transformed the user interaction into an improved accuracy of the image registration.

The images in all experiments are down-scaled from the original source, since the application in its current form can use a lot of memory due to several copies of the images being made. The exact setup and motivation of each experiment is detailed in its respective section.

5.2 Colour to Colour Alignment

This experiment examines the performance of the algorithm when registering regions of colour images to their full size source image. Registrations of this type should give the best results given that the subjects have roughly the same scale and rotations. With this experiment, the main goal is to discover whether the optimized algorithm performs at least as well as the base algorithm in terms of final alignment.

The images that are used in this experiment are a down-scaled version of the complete painting as the reference and a subregion of that same image around the lips as the template. These specific images were chosen because the combination proved to be difficult for the base algorithm, due to the template image being relatively small compared to the values of certain parameters. This would result in the algorithm finding matching control-point pairs outside the bounds of the template image, as can be seen in Figure 3.

In this case, user intervention was not needed, and the new application already improved the alignment over the base parameters on its own. However, the improved accuracy is not a result from the genetic algorithm itself but rather from a calculation that bases certain parameters on the size of the template.

5.3 MA-XRF to Colour Alignment

This experiment examines the performance of the algorithm when registering a MA-XRF scan to a colour image. The aim of this experiment is to find compare the alignment accuracy of the new application to the base algorithm.

None of the methods gave sufficient alignments, especially around the top of the headscarf as can be seen in Figure 4. The exact cause of this is not known for sure, but does not appear to be a lack of control-points. Figure 5 shows that the optimized parameters yield more control points, even when a



Figure 3: Comparison between the alignment of the lips to the original image. The left image uses the base parameters, the right image has been optimized using the genetic algorithm.

mask limits the algorithm to only the foreground. It is probable that the alignment fails due to the bilinear mapping of the template to the reference image, which was described in Section 4.1. A function with more parameters and more degrees of freedom might be able to stretch the template into a better alignment.



Figure 4: Comparison between alignment of an infrared scan and a colour image. The left image uses the base parameters, the right images has a masked applied to the background and has been optimized using the genetic algorithm.



Figure 5: Comparison between the control-points found during the alignment. The left image uses the base parameters, the middle and right images have been optimized using the genetic algorithm, the right image has a mask applied to it.

5.4 X-Ray to Colour alignment

This experiment examines the performance of the algorithm when registering x-ray images to a colour image. The aim of this experiment is to find compare the alignment accuracy of the new application to the base algorithm. The template image in this case is an x-ray scan of the subject in the reference image, but with a scaling of 95% compared to the reference.

This results of this experiment expose another problem with the current application in that it sometimes fails to find

a correct estimation of an initial transformation (Figure 6), in which case it defaults to no transformation at all. It is important that this initial alignment is close to the final alignment, since it is used to find the set of control-point pairs. The parameter optimization is currently not used for this initial alignment, but a similar solution could be applied to this problem. Another possibility is to add a new interaction which would allow the user to specify the rough alignment that the algorithm should use.



Figure 6: The algorithm could not find a valid initial transformation, causing the template to be aligned incorrectly. A correct alignment should have scaled the template to be the same size as the reference.

6 Responsible Research

As yet, the source code of the works discussed in this paper are not openly available. The code will first have to be finalized and cleaned, but it is expected that the complete source code will be publicly available a short time after this paper is published. The data used in the experiments of this research was acquired from the Mauritshuis, so the availability of it is uncertain but currently improbable.

When the source code is available, it is predicted that anyone could be able to achieve the same results as presented in this paper. Since all experiments are performed within the GUI, anyone with the application should be able to reproduce and validate them.

7 Conclusions and Future Work

An abstraction layer for the algorithm by Conover et al. [2015] was presented, which allows the user to influence the algorithm through a *Graphical User Interface* (GUI). The GUI provides several tools for the user to emphasize or exclude regions from the image registration algorithm, which affects how the genetic algorithm tunes the parameter. When the algorithm has finished, the user can view the matching *control-point* pairs and has the option to adjust them for more accurate results. Experiments show that the main advantage of the new application is the automatic optimization of the parameters, which prevents the user from having to know what each parameter does to get good alignments. Alignments in which certain areas of the image have been masked by the user, show no accuracy gain over the alignment of the full template using automatically optimized parameters.

The application in its current state suffers from several problems. The initial rough alignment is sometimes incor-

rect, which leads to an incorrect alignment. This could be solved by letting the user specify the rough alignment and skipping, since they might have a better idea of where the template should be registered. Another problem with the registration is that a bilinear mapping might not provide enough freedom to stretch the template into a good alignment, this might be solved using a function with more parameters like a bicubic mapping.

The application is not memory efficient and can therefore only be used with scaled down versions of the scans. Execution time is also a problem, as it can take several seconds to optimize the parameters of even a small image. Reducing the time it takes to optimize the parameters, would have multiple benefits. First, it would improve the user experience, but more importantly it would allow more optimization iterations which should lead to better results.

Future work can also investigate the user experience more deeply and focus on adding more advanced tools. Another option is to consider other image registration algorithms, which should be relatively trivial to implement since the GUI is independent of the algorithm that is used in the back-end.

References

- Lisa Gottesfeld Brown. 1992. A Survey of Image Registration Techniques. *ACM Comput. Surv.* 24, 4 (dec 1992), 325376. <https://doi.org/10.1145/146370.146374>
- New House Internet Services B.V. 2022. PTGui. <https://ptgui.com>
- Damon M. Conover, John K. Delaney, and Murray H. Loew. 2015. Automatic registration and mosaicking of technical images of Old Master paintings. *Applied Physics A* 119, 4 (01 Jun 2015), 1567–1575. <https://doi.org/10.1007/s00339-015-9140-1>
- Damon M. Conover, John K. Delaney, Paola Ricciardi, and Murray H. Loew. 2011. Towards automatic registration of technical images of works of art. In *Computer Vision and Image Analysis of Art II*, David G. Stork, Jim Coddington, and Anna Bentkowska-Kafel (Eds.), Vol. 7869. International Society for Optics and Photonics, SPIE, 92 – 100. <https://doi.org/10.1117/12.872634>
- Michael E. Farmer. 2013. Illumination invariant intensity-based image registration using chaos theory. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2094–2098. <https://doi.org/10.1109/ICASSP.2013.6638023>
- Rafael C. Gonzalez and Richard E. Woods. 2008. *Digital Image Processing, 3rd Edition*. Pearson.
- Elisabeth Hedlund, Jan-Erik Karlsson, and Sven-Ake Starck. 2010. Automatic and manual image fusion of In-111-pentetreotide SPECT and diagnostic CT in neuroendocrine tumor imaging - An evaluation. *JOURNAL OF MEDICAL PHYSICS* 35, 4 (OCT-DEC 2010), 223–228. <https://doi.org/10.4103/0971-6203.71766>
- Adobe Inc. 2022. Adobe Photoshop. <https://www.adobe.com/products/photoshop.html>
- Stefan Klein, Marius Staring, Keelin Murphy, Max A. Viergever, and Josien P. W. Pluim. 2010. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging* 29, 1 (Jan 2010), 196–205. <https://doi.org/10.1109/TMI.2009.2035616>
- A. Lasso. 2022. SlicerElastix. <https://github.com/lassoan/SlicerElastix>
- Yujiro Nakajima, Noriyuki Kadoya, Takayuki Kanai, Masahide Saito, Satoshi Kito, Shimpei Hashimoto, Katsuyuki Karasawa, and Keiichi Jingu. 2020. Evaluation of the effect of user-guided deformable image registration of thoracic images on registration accuracy among users. *Medical Dosimetry* 45, 3 (2020), 206–212. <https://doi.org/10.1016/j.meddos.2019.12.004>
- V. van der Valk, M. McCormick, and J. Kiggins. 2021. elastix_napari. https://github.com/SuperElastix/elastix_napari
- Abbie Vandivere, Jørgen Wadum, Klaas Jan van den Berg, Annelies van Loon, and The Girl in the Spotlight research team. 2019. From ‘Vermeer Illuminated’ to ‘The Girl in the Spotlight’: approaches and methodologies for the scientific (re-)examination of Vermeer’s Girl with a Pearl Earring. *Heritage Science* 7, 1 (29 Aug 2019), 66. <https://doi.org/10.1186/s40494-019-0307-5>
- Barbara Zitová and Jan Flusser. 2003. Image registration methods: a survey. *Image and Vision Computing* 21, 11 (2003), 977–1000. [https://doi.org/10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9)