Music Score Image Registration for Annotation Extraction

Joseph Grealy (Durham University, United Kingdom)

Abstract - Musicians of all levels enrich musical scores by adding handwritten annotations reflecting their interpretation of the score and providing additional playing instructions. The digital extraction, identification, and classification of these is of interest to musicians and historians alike.
This paper will explore the use of feature-based image registration techniques to align an annotated and clean version of a score in order to extract the annotations.
We will first discuss different techniques used for image registration in fields such as satellite imaging and discuss their applicability to musical scores. We will then implement our chosen method in Python, and test it on a selection of test scores. Finally, we will analyse the results before discussing areas for future work.

1 Introduction

Professional and amateur musicians in many musical traditions enrich written musical scores by adding handwritten annotations. These annotations change the interpretation of the score by indicating where performers make real-time adjustments to the printed score allowing for variations in style and highlighting historical practices.

Given the tendency of libraries and archives to digitize their musical scores, the application of annotation extraction techniques to musical score images could provide score-annotation data sets of interest to digital libraries (such as the New York Philharmonic Orchestra) and musicologists [1]. Additionally these extracted annotations would be of use within the field of Optical Music Recognition[1] to recover musical semantics, creating a more rich and accurate output [2].
This paper will explore the use image registration as a method to extract sheet music annotations hopefully to form part of a pipeline to classify and interpret annotations in the future.

Image registration aims to align two or more images with overlapping scenes captured at different times, by different sensors, or from different viewpoints [3].
In order to extract annotations, we can compare un-annotated ground-truth scores to annotated scores. However, visual distortions - normally some combination of translation, scaling, rotation, and skew - pose a problem when attempting to match up the two clean and annotated copies. Here we can use image registration techniques to estimate a geometrical transformation that aligns points from one viewpoint of a scene with the corresponding points in the other viewpoint [4].
In simple terms this works by considering the visual distortions as a mathematical transformation between two coordinate systems which we seek to identify in order to align the clean score with the annotated image, we can then make comparisons and "subtract" the two images to obtain the annotations on their own.

This is currently utilised in many areas such as medical image analysis, facial recognition, and surveillance and satellite imaging [4], and techniques from these fields can be analysed to see their relevance to our music score problem.

For the duration of this paper we will refer to the annotated score image as the target or reference image, and we will refer to the clean score image as the source or moving image. Once we have identified and applied the transformation to the moving image we will obtain the registered image (which will be the clean score aligned to the annotated one).

2 Types of Image Registration

There are two main types of image registration: intensity-based and feature based. The former compares the pixel intensities in the two images using a specified similarity metric, and a geometrical transformation between images which minimises/maximises this metric is found [5].
The latter extracts features such as points or lines present in both the images and uses their correspondences to estimate a geometrical transformation to get from one image to the other (More mathematically: estimates a coordinate map between two coordinate systems).

Which of the two above types is most appropriate depends on the problem at hand. For example, we must take into account: the modality of the images (whether the images were obtained by the same sensor type) [6], whether the images have different spectral contents (i.e. if the images have actual differences due to being captured at different points in time), the type of visual deformation present/type of transformation required (e.g. rigid, affine, or projective transformations), illumination changes, or noise in the images. [3]

In our case, we would hope that the clean and annotated copies of the score would be the same edition and high quality scans. However, this is not always the case as scores of interest are normally older editions where a clean copy may not exist (or be of low quality). Additionally, clean copies of scores can most easily be obtained from sheet music libraries such as IMSLP which features many user scans, often with noise or visual distortions from the scanning process.
It is also worth noting that our images will have spectral differences as one score will have annotations and the other will not meaning we do not expect the registered images to have the same pixel intensities. As a result of this feature-based methods are much more applicable to our problem.

2.1 Feature-Based Image Registration

Feature-based methods all tend to follow the same basic steps:

1. Identify feature points in both the images.
2. Match the extracted features and find the correct matching pairs.
3. Estimate the transformation parameters.
4. Apply the transformation to obtain the registered image.

There are many ways of achieving each of the above steps and additional steps can be taken to improve the final registration.

A popular feature extraction method is David Lowe's SIFT (Scale Invariant Feature Transform) algorithm [7] which is invariant to scaling, rotation and noise. It uses a difference-of-Gaussian function to search over all scales and locate stable keypoints. Local image gradients of the keypoints are computed and a keypoint descriptor is returned. These descriptors are highly distinctive and allow features to be recognised correctly with high probability [7].

There are also various adaptions to improve the performance of SIFT such as Principal component analysis (PCA)-SIFT [8], Bilateral filter SIFT (BF-SIFT) [9], and Adaptive binning (AB-SIFT) [10] as surveyed by [3]; however, SIFT-like methods struggle to provide meaningful matching results when applied to images with significantly different spectral contents [3]. SIFT keypoints are often concentrated on salient texture details such as text of street names on maps [11] which are inconsistent between images. This is could be a problem for us as our images have inconsistent contents - the annotations - which are likely to result in concentration of keypoints like the map labels. We can check if this is the case by examining the selected matches between the points; if none of the matches involve the annotations and are only between features present in both the images then we have no problem.

If the problem persists, one solution is to implement a segmentation stage. Image segmentation partitions an image into regions according to a given criteria, then binarises the image[2] to distinguish between objects and background [3]. We could use this to exclude the annotations by segmenting our score image into lines using a line detector; Grampone von Gioi et al present the fast Line Segment Detector (LSD) which has been shown to outperform other line detectors [12]. We have a large number of lines such as the note stems, staff lines, and bar lines which should be present and match up in both of the images. These should hopefully provide reliable keypoints to calculate the transformation. It is also worth noting that the line segments (and corresponding support regions) are only used for segmentation, line segments are not sufficient to support SIFT extraction and matching as they have limited texture details [3].

One issue with the LSD is that line features with details or noise are easily detected as multiple line segments. It has been proven [13] that analysing with the LSD at a coarser resolution helps to detect global structures. This is because line-support regions are extracted using similarity of gradient which is highly dependent on noise which decreases at coarser resolutions.

Zhao et al propose the use of a iterative strategy to ensure the accuracy of registration by down-sampling the image to a coarser resolution, and then re-segmenting and re-matching until an expected accuracy is achieved [3].

To support SIFT there are a variety of algorithms which refine the selection of keypoints or matches. These vary widely tackling the problem in very different ways. The classical RANSAC algorithm treats the problem as outlier detection using parametric models, whereas the Locally Linear Transform (LLT) uses a non-parametric Bayesian model. Other methods include the Weighted Graph Transformation Matching (WGTM) and Geometric Outlier Removal (GOR) algorithm which take a geometrical, graph-based approach. Though there may be a slight variation in performance between these methods, it is easiest to start by using the freely available RANSAC algorithm and switching if the matches selected are decreasing the quality of the final registration.

2.2 Types of Transformation

When performing image registration, we must also consider what sort of deformations are present and hence what sort of transformation is required to obtain a sufficiently registered image. This will depend on whether the images were obtained from different viewpoints or by different sensors [3].

More complex transformations involving more parameters require longer to calculate and often need more data (matching keypoint pairs) to return a result. There are two broad categories of transformation model.

1. Linear transformations:
   These are a combination of translation, rotation, global scaling, shear and perspective components [5]. Without the perspective component this is referred to as an affine transformation which can be represented as a 3x3 matrix using homogeneous coordinates. With the perspective component it is a perspective transformation [5].
   Linear transformations are global in nature, thus not being able to model local deformations.

2. Non-Linear transformations:
   Also referred to as elastic or non-rigid transformations, these allow for local deformations. Non-rigid transformation approaches include polynomial wrapping, interpolation of smooth basis functions (thin-plate splines and wavelets), and physical continuum models [5].

For our problem, we will assume we are dealing with an affine transformation, this is because we would expect global translations, rotations, scaling and perhaps shear, but we would hope to not have any perspective deformation as our images should all be computer scans, and we expect the staff lines to remain parallel [14].

We also would not expect any non-linear or local deformations as they are most often introduced from lens distortions which should be at a minimum as these are just scans of the scores [14]. The actual matrices we will be using, the parameters present, and the matrix equations we will form and solve using Python are well described by Kheng in [15].

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

Eq. 1: The 2-D affine transformation matrix equation using homogeneous coordinates.

3 Initial Method and Results

We will now recap the method we will initially be implementing.

1. Fist the SIFT keypoints are extracted from both of the images and brute force matched using their corresponding keypoint descriptors.
2. The RANSAC algorithm will then be used to refine the selected matches.
3. The affine transformation matrix will be then calculated from the correct matches using a least squares method.
4. We can then warp the clean score using the matrix to obtain the registered image.
5. We can then explore the results using a number of plots.

All of the above steps will be implemented in Python using popular image libraries OpenCV and scikit-Image, all code is also available on GitHub here.

We will be working with a test set of 17 annotated score images taken from the New York Philharmonic Digital Archives which can be accessed here. The clean versions of the scores were obtained by scraping IMSLP using an API. This process was carried out by Eamonn Bell from Durham University.

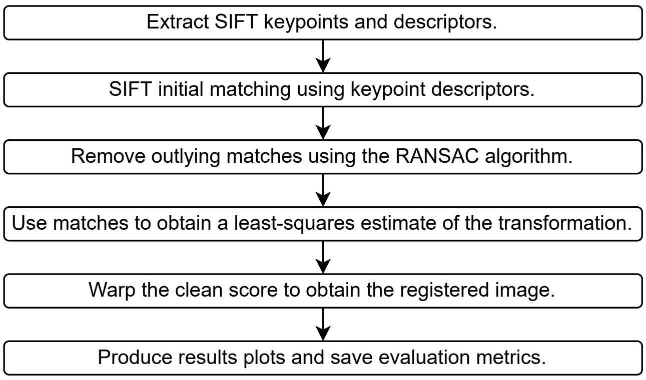
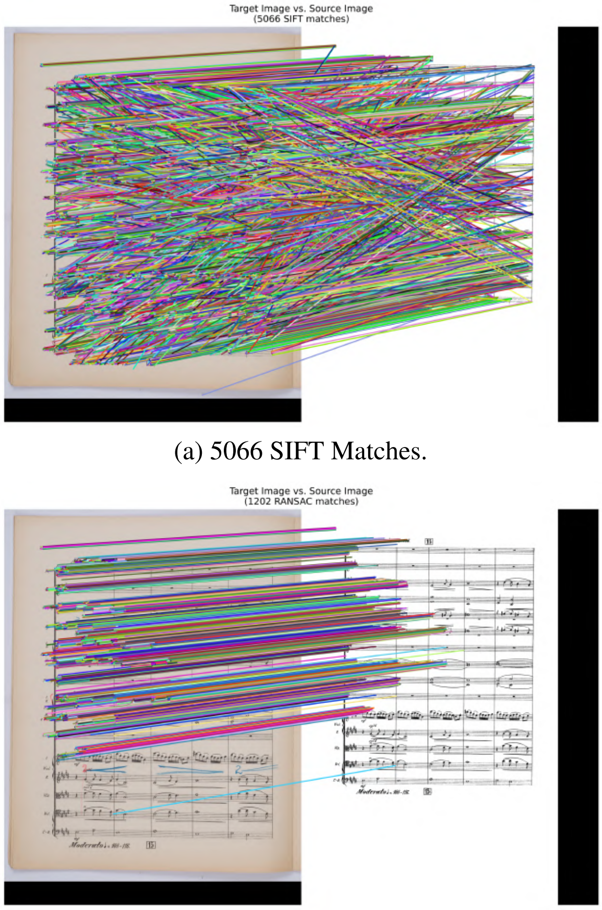


Figure 1: Flow diagram demonstrating method.

Upon inspecting the results we can decide whether it is necessary to add a segmentation stage, an iterative strategy, or replace the RANSAC algorithm with a better one.

Figure 2 and figure 3 show the returned SIFT and RANSAC matches from two score images chosen randomly from the test dataset (images 10 and 17). It is clear from the large number of criss-crossed lines in the SIFT match images that many are incorrect. Criss-crossed lines are an indicator of this since it implies the match is not between corresponding points in the two scores. Upon examining the second RANSAC match images, the result is much better in both cases. There are a large number of matches appearing to be correct, between corresponding points in the images. There also does not appear to be any matches involving the annotations which means the segmentation stage discussed before is not necessary.

Figure 4 displays all the SIFT keypoints selected, and while there are some keypoints concentrated at the annotations, they do not seem to be leading to many of the incorrect SIFT matches seen in figures 2 and 3. As we are getting meaningful matching results, it is not necessary implement any of the changes previously discussed.



(a) 5066 SIFT Matches.

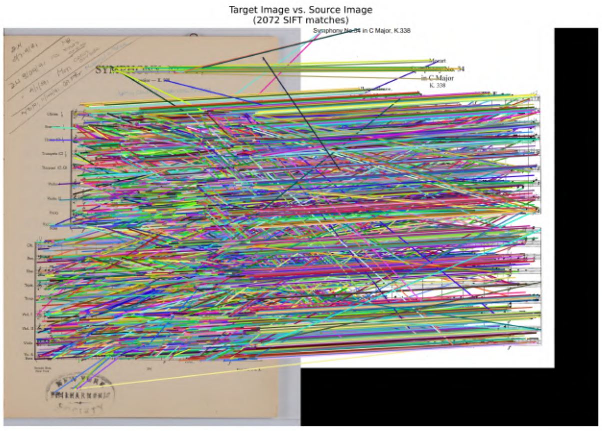
(b) 1202 RANSAC Matches.

Figure 2: Comparison of the SIFT and RANSAC matches for image 10 of the dataset.

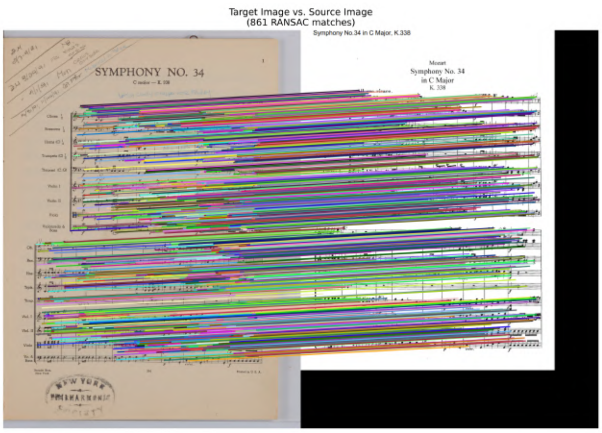
4 Results Analysis

To evaluate the success of image registration it is common to use metrics such as the cross correlation coefficient to judge success. In our case even with a perfect registration we will not have a perfect cross correlation coefficient of 1 because of spectral differences between the images (i.e. the annotations), but it is still worth considering it to judge the success. It is also likely that the cross correlation will be dominated by the large amount of "white" space in the images and still obtain a relatively high result.

We will also calculate the root-mean-square error (RMSE) of the registration, this. This considers the final selection of keypoints and calculates the average euclidean distance between the points in the annotated image, and the corresponding points in the clean image after they have been transformed. In a perfect transformation these points should all match up perfectly so we would expect the RMSE to be 0. Using RMSE compared to the cross-correlation is beneficial since it will not be affected by spectral differences between the two images since it only uses the keypoints; however if

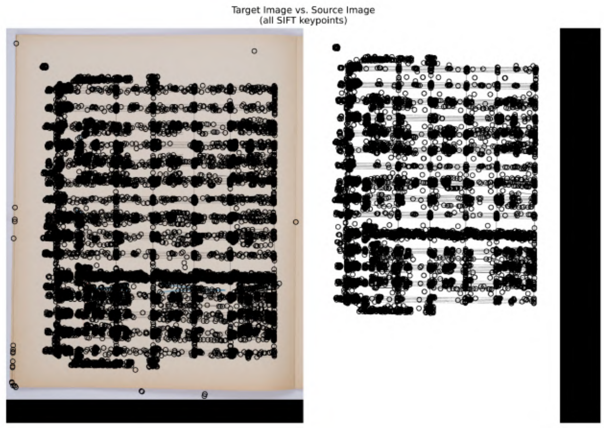


(a) 2072 SIFT Matches.

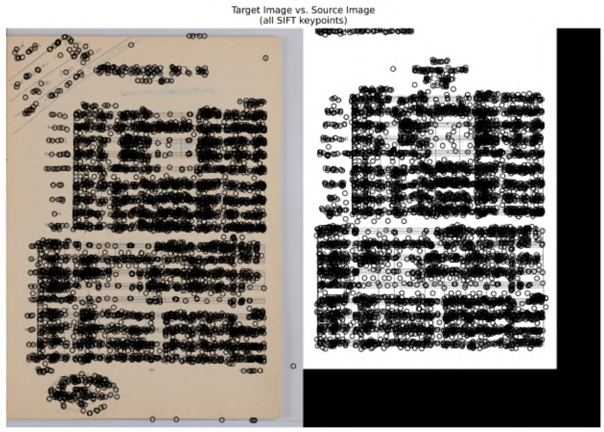


(b) 861 RANSAC Matches.

Figure 3: Comparison of the SIFT and RANSAC matches for image 17 of the dataset.



(a) SIFT keypoints in image 10.



(b) SIFT keypoints in image 17.

Figure 4: All SIFT keypoints found in images 10 and 17.

matching keypoints are only in certain parts of the image, the success of the registration in other parts of the image would not be taken into account.

In addition to these metrics, we can consider plots to judge success, this will allow us to visually inspect that the registration worked confirming the results shown by the metrics. We will use an overlap plot (where the red channel is the annotated target image and the blue/green channels are from the registered image) and a difference plot (formed by pixel-by-pixel subtracting the registered clean image from the annotated image) in an attempt to extract the annotations.

It is also worth noting that we will discuss the success of the registration in the context of annotation extraction (i.e. how good is our registration for the extraction of annotations).

The evaluation metrics for all images in the test dataset can be seen in table 1. The cross-correlation is above 0.9 in all cases with an average of 0.97158 which suggests the registered image and the annotated score match closely.

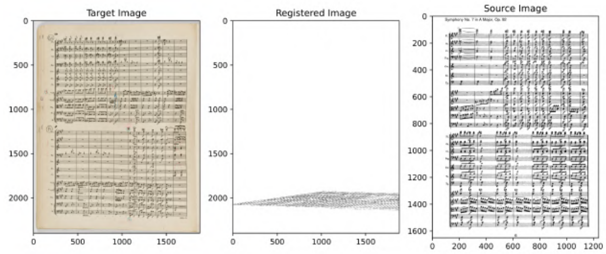
In many cases, the RMSE is also quite low which implies the registration did a good job of finding the mapping between the sets of selected keypoints. However, in some cases the number of RANSAC matches (which are the keypoints used to estimate the transformation) are very low, specifically images 2, 3, and 4.

Image Pair	RMSE	xcorr	SIFT matches	RANSAC matches
1	10.46098	0.97798	4843	548
2	38.18941	0.96711	3218	8
3	45.63586	0.94815	2910	18
4	8.85700	0.94569	3297	15
5	8.51994	0.98475	3517	1123
6	10.00546	0.95240	4780	449
7	16.43558	0.97835	2076	227
8	6.94984	0.98209	5822	2208
9	3.77236	0.98462	4954	2065
10	11.02791	0.97989	5066	1207
11	6.10498	0.94671	6713	1249
12	16.38385	0.95879	6246	1464
13	4.55601	0.98467	4801	884
14	3.88342	0.99006	3011	1200
15	29.05146	0.97693	3468	1060
16	28.59131	0.96947	4553	1466
17	3.68187	0.98923	2072	866
Average	14.82984	0.97158	4197	945

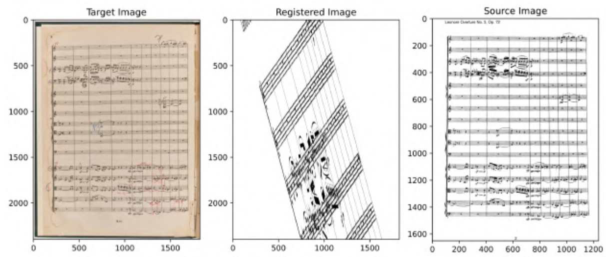
Table 1: Evaluation metrics for each of the 17 image pairs.

Upon examining the images it turns out the registration was not successful at all with very distorted results (see figure 5) despite having a reasonable cross-correlation.

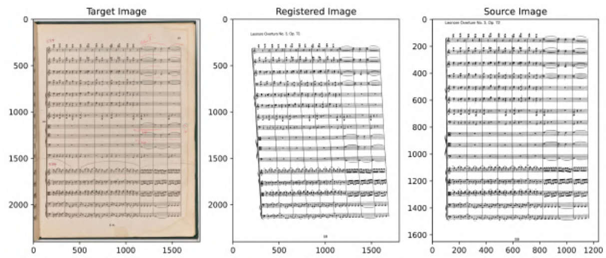
It turns out that the scores in image pair 2 are not actually the same which explains why the registration failed, this is probably due to some fault from the API scraping process. However, this is not the case for image pairs 3 and 4. In these cases a likely explanation for the bad result is a concentration of SIFT keypoints and matches on extraneous image features from the scanning process (the other pages visible at the top of the annotated score) as can be seen in figure 6. This could have led to the RANSAC algorithm classifying more points than it should have as outliers.



(a) Registered Sequence for image 2.



(b) Registered Sequence for image 3.



(c) Registered Sequence for image 4.

Figure 5: Unsuccessful registered Sequences for images 2, 3, and 4.

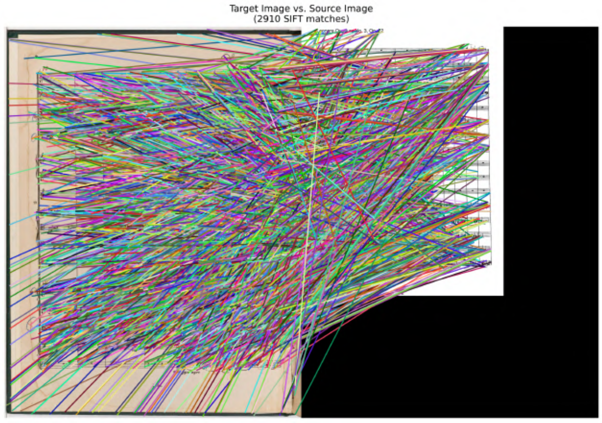


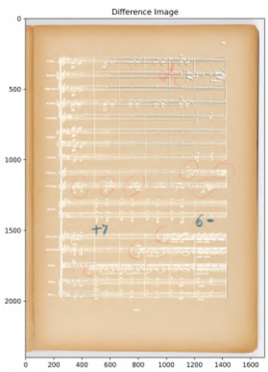
Figure 6: SIFT matches for unsuccessful registration of image 3 demonstrating large number of incorrect matches from border due to scanning.

Having explained the failed cases, we should examine some more successful cases. Figures 7 and 8 contain the difference and overlap plots for image 9 (second lowest RMSE), and image 14 (highest cross-correlation). In both cases the registration appears almost perfect, the only slight error is in the top right section of image 9.

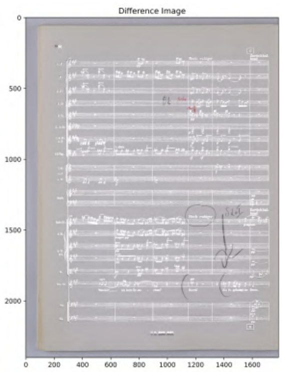
We can go further and attempt a basic extraction of the annotations by binarising the images before taking the difference (setting all pixels above a certain threshold to black/1, or setting them to white/0 otherwise), we used the Otsu Thresholding method. This will hopefully remove set all pixels to white apart from the annotations as we have removed the staves and notes by taking the difference; however it is likely that light coloured annotations will not be extracted as the background itself is dark which will raise the threshold.

The results can be seen in figure 9. As expected lots of the light coloured annotations were not extracted, but a number of the darker annotations stand out; however there is also a lot of noise left over from the stave lines, bar lines, and notes. This noise is especially prevalent in areas of the image where the registration didn't look perfect in the previous plots (notably in the top right of image 9). The noise is likely also to be caused by slight pixel differences between the two images, likely introduced by the scanning process.

This result could likely be improved by de-noising and other edits to the image, but we will discuss ideas improvement after the results. This also reveals that even in the most successful registrations, actually extracting the annotations (the next stage in the pipeline) is a challenging problem of its own.



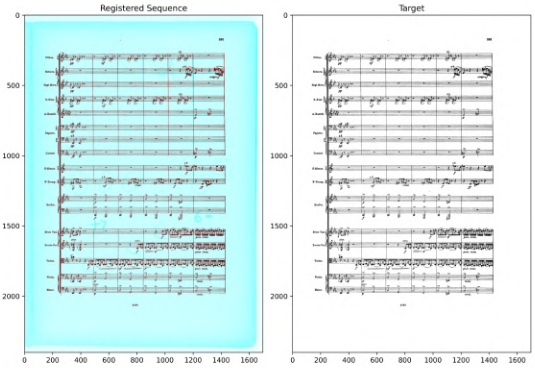
(c) Difference plot image 9.



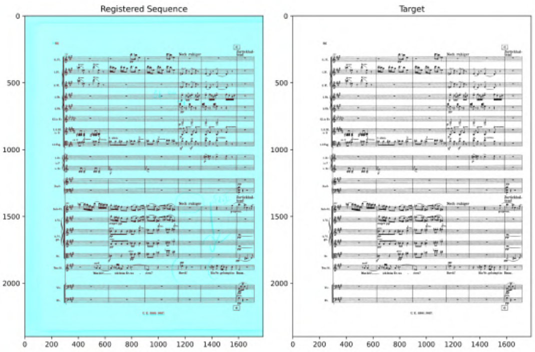
(d) Difference plot image 14.

Figure 7: Difference plots for image 9 (second lowest RMSE), and image 14 (highest cross-correlation).

Considering some of the other image pairs we can examine why they got a lower cross-correlation/RMSE. Examining the difference plots of images 1 and 8 in figure 11 they both look to be largely accurate registration where only small parts of the score still appear, this is backed up by low RMSEs (< 11) and high cross correlation (> 0.97). However, upon binarising the difference images (figure 10) there is a very large amount

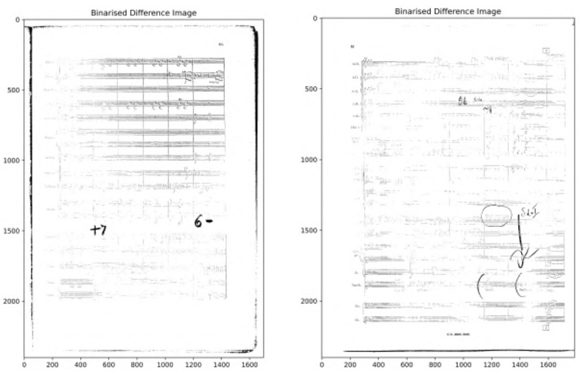


(a) Overlap plot for image 9.



(b) Overlap plot for image 14.

Figure 8: Overlap plots for image 9 (second lowest RMSE), and image 14 (highest cross-correlation).

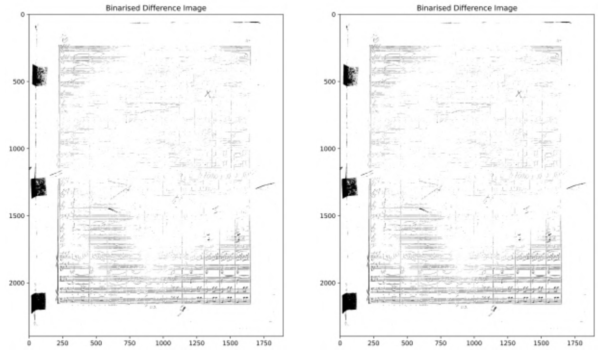


(a) Binarised difference image 9. (b) Binarised difference image 14..

Figure 9: The binarised difference images for successful registration using Otsu thresholding.

of noise left by the score revealing there are only small areas where taking the difference successfully removed the printed score. Despite this, some dark annotations still appear clearly in image.

This shows just how precise the registration must be for the subtraction to truly remove the printed score even when the result appears successful. It also highlights how important the scanning step is, as small pixel differences due to scan quality mean that even with a perfect registration, the subtraction will leave large amounts of noise.



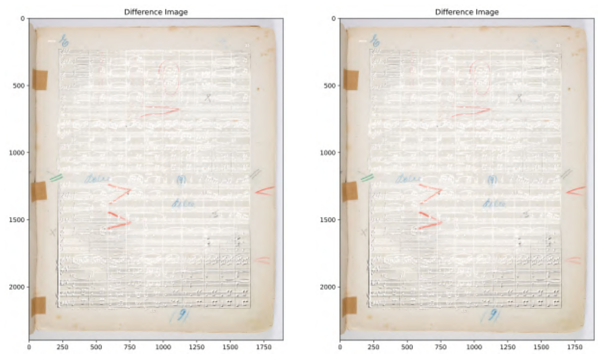
(a) Difference image 1. (b) Difference image 8.

Figure 10: Noisy binarised difference images for largely successful registrations.

The results for images 12 and 13 look less successful than those of the images discussed above. Considering the difference and binarised difference plots in figure 12, it is clear that there is some kind of perspective distortion in the images as there are areas where the registration looks successful and gradually gets more dissonant.

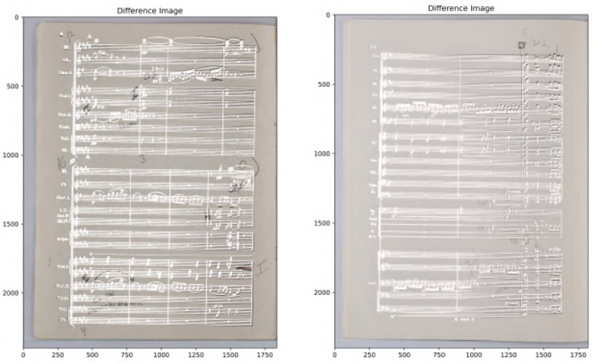
This means that our assumption we would have to only deal with an affine transformation was incorrect. Perspective warping is likely due to the scanning process for one or both of the images, and it is also likely it affected other images in the dataset but to a less noticeable extent.

The binarised difference image for image pair 13 highlights this showing a very successful registration in the left half of the image, with the largest area without noise from the printed score out of all the test pairs. However, a very low RMSE despite the noise in the right of the image begs for more investigation. Upon examining a plot of the RANSAC matches (Figure 13) used for estimating the transformation, many of them are from the left half of the image explaining why the registration was so successful on this side. The difference in success between the parts of the image also shows that an affine transformation isn't enough and that we have local-deformations. We will discuss solutions to this in the next section.

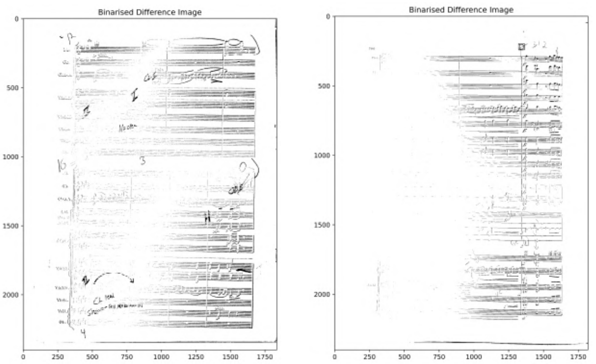


(a) Binarised difference image 1. (b) Binarised difference image 8.

Figure 11: The binarised difference images for images 1 and 8 showing what appears to be a largely successful registration.



(a) Difference image 12. (b) Difference image 13.



(c) Binarised diff. image 12. (d) Binarised diff. image 13.

Figure 12: The difference and binarised difference images for pairs 12 and 13 showing perspective distortions.

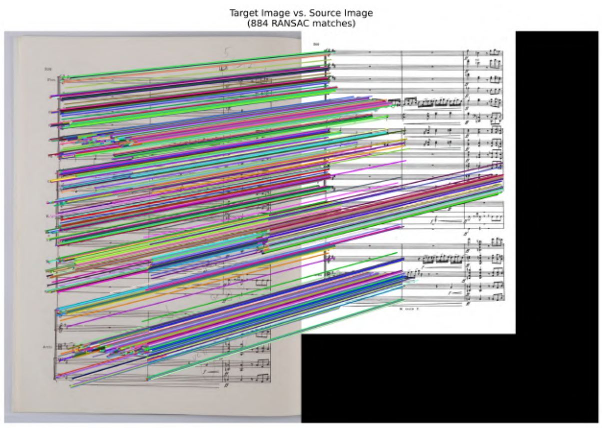
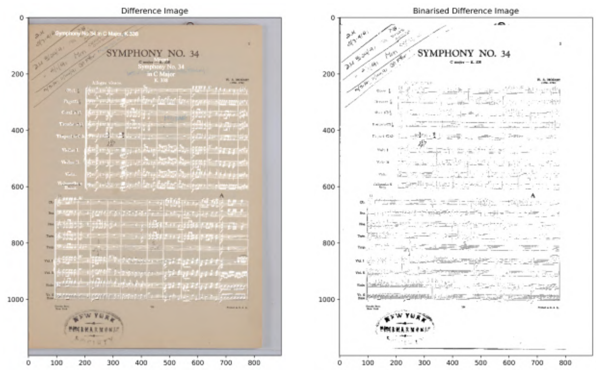


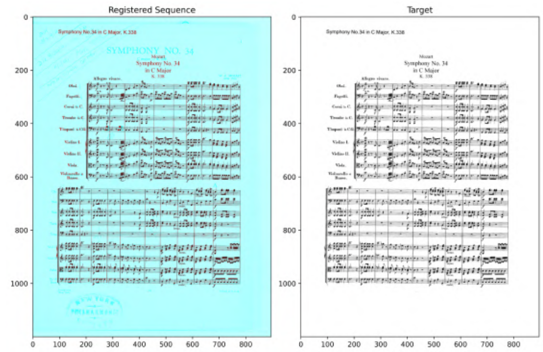
Figure 13: RANSAC matches for image pair 13 showing concentration of matches on the left of the image.

Finally, we will examine image pair 17. The difference, binarised difference, and overlap plots are shown in figure 14. Image 17 has the most spectral differences between the clean and annotated score out of the pairs in the test dataset, this is due to the editions being slightly different (differing title/stave instrument label/note placement) as well as a large number of annotations including a stamp. This is clear when examining the comparison plots. The clean score is also of notably lower quality than most of the other clean scores in the dataset.

Despite all of these issues the registration is largely successful with some of the best RMSE and cross-correlation of the whole dataset. Upon examining the RANSAC matches for pair 17 (shown earlier in figure 3), all the matches are between printed score details from all over the two scores and results in less noise than expected in the binarised image.



(a) Difference image 17. (b) Binarised diff. image 17.



(c) Overlap image 17.

Figure 14: The results for pair 17 showing largely successful registration results despite large spectral differences between images.

This example highlights issues we are likely to have to deal with where the exact edition cannot be found for the clean score or where there is a large number of spectral differences. The method we have used has proved handle cases like this well at the registration stage, it is the extraction of the annotations which still poses an issue.

5 Conclusions and Future Work

Having analysed our results, our method was mostly successful in registering the scores. Failed cases all had bad scans with a large border which disrupted the RANSAC algorithm. Most cases had a mostly accurate registration but unexpected perspective warping from the scanning stage meant that the affine transformation model was not sufficient. Despite this success, the primitive annotation extraction method (taking the difference of the binarised images) revealed a lot of noise left behind from the printed scores and often did not show any light coloured annotations.

Even with improvements to the registration pipeline (such as using a perspective transformation model), it is almost impossible to remove the resulting noise completely.

This is because the noise it likely to result from spectral pixel differences between the two images due to them being editions and different quality scans.

In an ideal world, the clean and annotated score would not only be exactly the same edition, but also would be scanned at the same time by the same scanner to reduce this quality difference. However, in order to perform annotation extraction on the large number of annotated scores of interest in archives - such as the 40000 scores and parts in the New York Philharmonic Digital Archives - another solution is needed.

More advanced de-noising methods than just a binarisation such as spatial filtering, transform domain filtering and wavelet thresholding [16] could be used on the scores at different stages of the pipeline. Another option could be to try other types of thresholding for the binarisation, or locally thresholding different sections of the image. Another option could be to use machine learning techniques; by obtaining a small dataset of ground truth annotations (by manually drawing over the annotations in the images using something like Adobe Photoshop) supervised learning could be implemented to learn what in the difference images is noise and what is annotation. Although there are many possible solutions to the de-noising problem, it could prove difficult to truly automate this process.

Returning to the registration process itself, improvements could be made by iterating the process until the RMSE falls below a certain threshold which could improve the final precision. A segmentation and iterative approach such as that discussed in the literature review could also improve the final selection of matches.

Alternatives to the RANSAC algorithm could also help reduce the number of failed registrations by doing a better job of removing outliers in cases where there are extraneous details in the scan. However, it is unlikely that improvements to the point-based image registration pipeline (including those discussed in the literature review) would have a large effect on the results if there is an underlying issue with the transformation model or large pixel differences in the printed scores from low quality scans.

However, using a perspective transformation model to help with perspective warps from the scanning process, or considering non-linear deformations using machine learning are both likely to have a improvement on the results. Even with these changes, exploring different types of denoising techniques to help with the annotation extraction should be the focus to aid the development of novel annotation extraction software.

Notes

1. OMR - the field of research that investigates how to computationally read music notation in documents
2. Converting to a binary black and white image.

References

[1] E. Bell and L. Pugin, *'Heuristic and supervised approaches to handwritten annotation extraction for musical score images'*, Int. J. Digit. Libr., vol. 20, no. 1, pp. 49–59, Mar. 2019, doi: 10.1007/s00799-018-0249-7.

[2] J. Calvo-Zaragoza, J. Hajič Jr., and A. Pacha, *'Understanding Optical Music Recognition'*, Jul. 2020. doi: 10.1145/3397499.

[3] M. Zhao, Y. Wu, S. Pan, F. Zhou, B. An, and A. Kaup, *'Automatic Registration of Images with Inconsistent Content Through Line-Support Region Segmentation and Geometrical Outlier Removal'*. arXiv, Apr. 02, 2022. Accessed: Jul. 05, 2022. [Online]. Available: http://arxiv.org/abs/2204.00832

[4] V. Argyriou, *'Image, video & 3D data registration medical, satellite and video processing applications with quality metrics'*. Chichester, West Sussex, United Kingdom: Wiley, 2015.

[5] S. Boda,'Feature-Based Image Registration', Department of Electronics and Communication Engineering National Institute of Technology, Rourkela, 2009.

[6] R. Safari, N. Narasimhamurthi, M. Shridhar, and M. Ahmadi, 'Document registration using projective geometry', IEEE Transactions on Image Processing, vol. 6, no. 9, pp. 1337–1341, Sep. 1997, doi: 10.1109/83.623198.

[7] D. G. Lowe, ' *'Distinctive Image Features from Scale-Invariant Keypoints'*, International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: 10.1023/B:VISI.0000029664.99615.94.

[8] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-Like algorithm for SAR images," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 1, pp. 453–465, Jan. 2015.

[9] S. H. Wang, H. J. You, and K. Fu, "BFSIFT: A novel method to find feature matches for sar image registration," IEEE Geosci. Remote Sens Lett., vol. 9, no. 4, pp. 649–653, Oct. 2012.

[10] A. Sedaghat and H. Ebadi, "Remote sensing image matching based on adaptive binning SIFT descriptor," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 10, pp. 5283–5293, Oct. 2015.

[11] G. Yammine, E. Wige, F. Simmet, D. Niederkorn, and A. Kaup, *'Novel Similarity-Invariant Line Descriptor and Matching Algorithm for Global Motion Estimation'*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 8, pp. 1323–1335, Aug. 2014, doi: 10.1109/TCSVT.2014.2302874.

[12] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, *'LSD: a Line Segment Detector'*, Image Processing On Line, vol. 2, pp. 35–55, Mar. 2012, doi: 10.5201/ipol.2012.gjmr-lsd.

[13] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, 'LSD: A Fast Line Segment Detector with a False Detection Control', IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 4, pp. 722–732, Apr. 2010, doi: 10.1109/TPAMI.2008.300.

[14] A. A. Goshtasby, 'Transformation Functions', Springer London, pp. 343-400, 2012, doi: 10.1007/978-1-4471-2458-0.

[15] L. W. Kheng, 'Image Registration: S4243 Computer Vision and Pattern Recognition', Department of Computer Science School of Computing National University of Singapore, pp. 1-39.

[16] P. B. Alisha, and K. Gnana Sheela, 'Image Denoising Techniques-An Overview', IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), vol. 11, no. 1, Ver. I, Jan. - Feb. 2016, pp. 78-84.

Joseph Grealy [joseph.p.grealy@durham.ac.uk] is a student studying Mathematics and Statistics at Durham University. This report was carried out as part of the Laidlaw Leadership and Research programme under the supervision of Dr Eamonn Bell [eamonn.bell@durham.ac.uk], an assistant professor in the Department of Computer Science at Durham University.