Application of Image Registration Techniques to Musical Scores

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.

This project will form part of wider research being undertaken at Durham University to automatically identify, segment, extract, and classify handwritten annotations as they appear in digital images of music scores. 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 (Bell and Pugin, 2019, p49).

In order to extract annotations, unannotated ground-truth scores can be compared to annotated scores in order to identify the annotations. 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. This issue is an image registration problem and the focus of this project.

Document or image registration is a challenging problem in image analysis, defined as the estimation of a geometrical transformation that aligns points from one viewpoint of a scene with the corresponding points in the other viewpoint (Argyriou et al, 2015, pp.1-2). In simple terms this considers the visual distortions as a mathematical transformation which we seek to identify and invert in order to retrieve the undistorted image and make comparisons. This is currently utilised in many areas such as medical image analysis, facial recognition, and surveillance and satellite imaging (Argyriou et al, 2015, pp.1-2). Optical Music Recognition (OMR) is the field of research that investigates how to computationally read music notation in documents (Calvo-Zaragoza et al, 2020, p.3). This project falls within the bounds of OMR, computer vision and image processing.

The objective of my research will be to try to find image registration techniques that are effective when applied to music scores. By carrying out a review of current registration literature I hope to identify techniques from other areas such as medical imaging which have potential to be modified to work on musical scores. I will then curate a selection thirty or so annotated and clean (reference) scores which can be used for testing. I will implement the two or three established techniques with the most potential to work on music scores in Python and test them on the sample of scores. I will then quantify the error in the different techniques and evaluate their effectiveness as well as identifying areas for further research. All of this would be detailed in a research report which I would produce at the end of the research.

Having conducted some preliminary research, I have ideas on a number of techniques which could be applied to musical scores. One simple technique originating from office automation would be to identify the transformation necessary by using results from projective geometry (Safari et al, 1997, p.1337), by considering a general transformation, we can choose points to find the parameter values. Another method is the scale-invariant feature transformation (SIFT) which is a well-known computer vision algorithm (Lindeberg, 2012, p.1) that is used in many areas such as satellite imagery. By extracting key points from the reference image, we can compare structural features to identify the same points in the new image and proceed to make further statistical comparisons to estimate the parameters of the mathematical transformation of the image. There are also numerous deep learning techniques used in medical imagery (Devine et al, 2020, p.246) which combine methods to reduce error compared to conventional image registration workflows.

In my research I would hope to explore these further and adapt the most relevant techniques for application to musical scores to aid research into the automated classification of handwritten music annotations.

References:

Bell, E. and Pugin, L., 2019. Heuristic and supervised approaches to handwritten annotation extraction for musical score images. International Journal on Digital Libraries, 20(1), pp.49-59.

Argyriou, V., Martinez, J., Villarini, B. and Roche, A., 2015. Image, Video and 3D Data Registration. John Wiley & Sons, pp.1-14

Calvo-Zaragoza, J., Jr, J.H. and Pacha, A., 2020. Understanding optical music recognition. ACM Computing Surveys (CSUR), 53(4), pp.1-35.

Safari, R., Narasimhamurthi, N., Shridhar, M. and Ahmadi, M., 1997. Document registration using projective geometry. IEEE transactions on image processing, 6(9), pp.1337-1341.

Lindeberg, T., 2012. Scale invariant feature transform. Scholarpedia 7(5), p.10491.

Devine, J., Aponte, J.D., Katz, D.C., Liu, W., Vercio, L.D.L., Forkert, N.D., Marcucio, R., Percival, C.J. and Hallgrímsson, B., 2020. A registration and deep learning approach to automated landmark detection for geometric morphometrics. Evolutionary Biology, 47(3), pp.246-259.