UNIVERSITY OF CAMBRIDGE
Centre for Digital Built Britain

2017/18 Mini-Project
Autonomous Image Recapture
Final Report

Michael Ramage, Department of Architecture, mhr29
Joan Lasenby, Department of Engineering, jl221
Hugo Hadfield, Departments of Engineering & Architecture, hh409
Chris Doran, Sidney Sussex College, chrisjldoran@gmail.com

Abstract

This research project has used advanced image processing and computer vision techniques to recreate geospatial viewpoints of some historic, 17th century David Loggan line drawings of Cambridge Colleges. Subsequent drone flights then captured current images of the colleges from these same viewpoints. Using some Cambridge colleges as examples, we show that we are then able to overlay the image and the drawing, highlighting differences and similarities. This leads to interesting questions in the field of history-of-architecture and paves the way for more complete structural and architectural analysis of such historical buildings.

1 Research Question

David Loggan (d.1692) was engraver to Oxford and Cambridge Universities, and is known for his birds-eye etchings of Cambridge and Oxford Colleges. Using computer vision and image processing techniques, this project aims to take Loggan's 17th century etchings and determine the best fit to the geospatial coordinates that he imagined them from. With these coordinates, we then program a drone to fly to the correct location and record images to recreate Loggan's line drawings. The immediate outcomes are threefold:

- Computer vision techniques to determine geospatial viewpoints: such techniques exist for photographs, but are not known to exist for aerial etchings.
- Drone flight automation for reconstructive image capture integrated with existing 3D buildings/terrain.
- Direct comparison between Loggan's views and contemporary views to highlight urban growth, change and stasis. Additionally, we can determine the extent to which Loggan's perspective drawing was correct.

Initially the product of this research will be useful for time and space comparisons of cities and infrastructure, including urban growth, building condition surveys and similar. We imagine other uses will grow as our research develops.

As above, our aim is to identify the viewpoints in a selection of Loggan drawings and take similar images from drone flights; as time allows we will extend our analysis to as many of the Loggan etchings as feasible. We plan to use Loggan's iconic image of Trinity College to test the software and efficacy of our approach. Google Earth points toward a proof of concept; our research would implement a higher accuracy, customisable and more specific algorithm. Our research will be extendible to places Google has not yet reached, and will be specifically targeted, higher resolution/accuracy and more visually useful.

Figure 1: Desired pipeline: a sketch is input into GArch and a GPS waypoint for a drone shot is output.

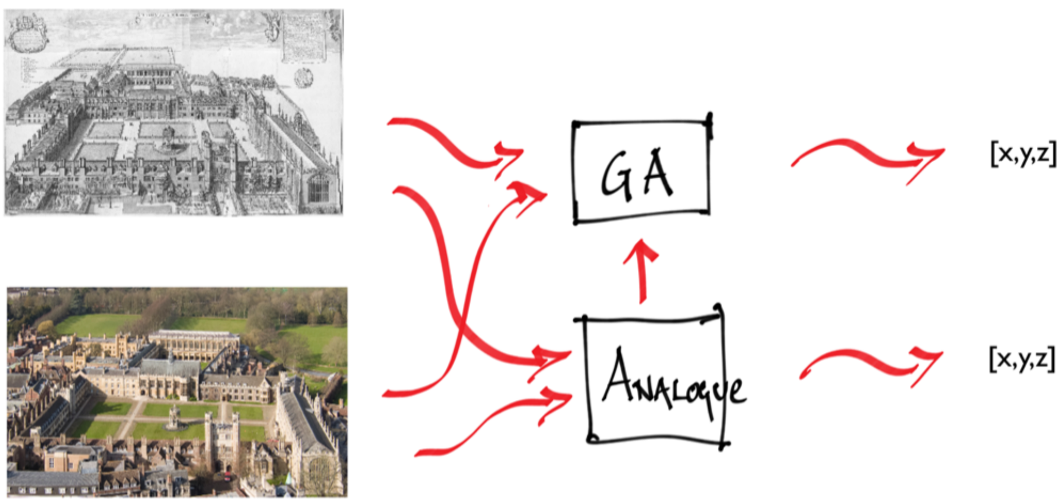

Figure 2: Manually assisted pipeline: input from a human guides the initial camera rotation and translation estimation.

2 Methodology

To tackle this research question we have created Geometric-Architect (GArch). GArch is a suite of software tools for multi-image point and line extraction, 3D reconstruction, model matching and pin-hole camera viewpoint finding. This package builds on previous work in the area of multi-image reconstruction and extends it with novel line-based matching algorithms designed in the framework of conformal geometric algebra (CGA).

The GArch pipeline is illustrated in figure 3 and is as follows: first, stills of the building of interest are taken from multiple viewpoints by flying a drone around the building. We then match points between the stills using a multi-image SIFT technique and use these points to reconstruct a wireframe of the building by utilising the metadata encoding the GPS position and attitude of the camera and the algorithms described in [1] and [2]. At a later stage in the research, this technique will be extended to use both points and lines.

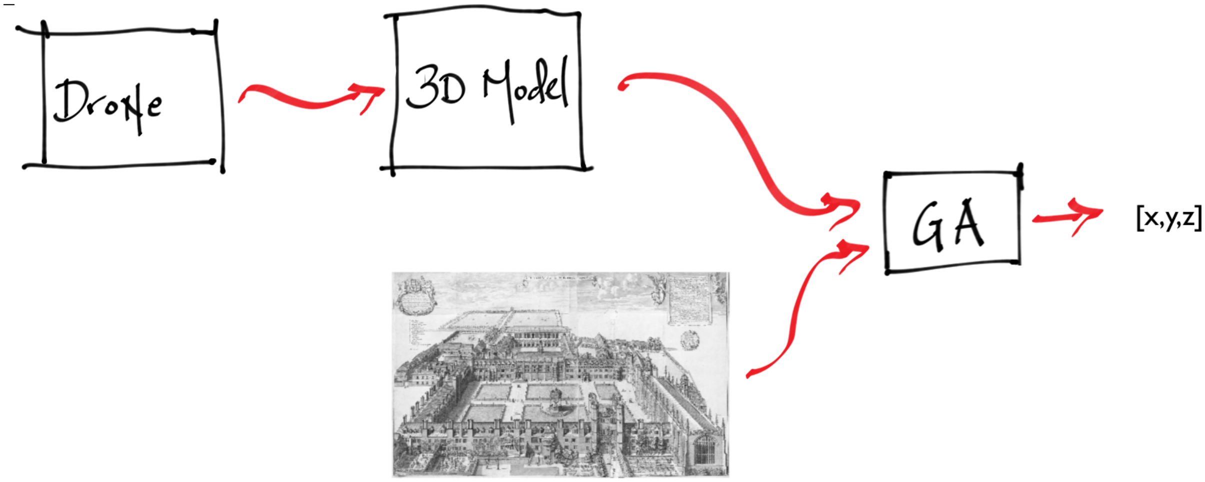

Figure 3: 3D reconstruction pipeline: a drone is used to construct a 3D model of the building that is the subject of the sketch, this is then used to find the viewpoint of the sketch.

Given we have a reasonably accurate wireframe of the building, our task now becomes to find the query image relative to this wireframe. First, lines are extracted from the query image with the probabilistic Hough transform: figures 4 and 5 show the results of this process for the Loggan drawing of Clare college. Each of these extracted lines are then matched with a line from the wireframe model. Next we define a simulated camera and project lines from the wireframe into the image plane of a simulated camera. We define a cost function in the image plane of the camera as the sum of the squared distances, both translational and rotational, between each object and its match. The rotation and translation of the camera from the origin form the search space for our optimisation and the minimum of the cost function represents our best guess at the location of the camera. The best guess camera location is output in the form of a GPS waypoint, altitude measure and camera attitude information to allow a drone to fly to this exact point and recapture a photograph with the same view as the query image. This algorithmic process requires several new techniques, for example, fast & robust line matching and cost functions for how close one line is to another – the mathematical development of these techniques has been submitted as papers the abstracts of which are reproduced in the appendices and described in a little more detail below.

There are a number of challenging technical aspects to this project, the most significant being the automated matching of 3D wireframe lines to 2D image plane lines and construction of accurate 3D wireframe models from drone footage. To attack the first of these issues, the matching problem, we have designed a probabilistic iterative matching algorithm, **REFORM** (**R**otor **E**stimation **F**rom **O**bject **R**esampling and **M**atching) [3]. This algorithm goes a long way towards solving the matching problem but at present when attempting to match the 3D to 2D lines it requires a good initial guess of the

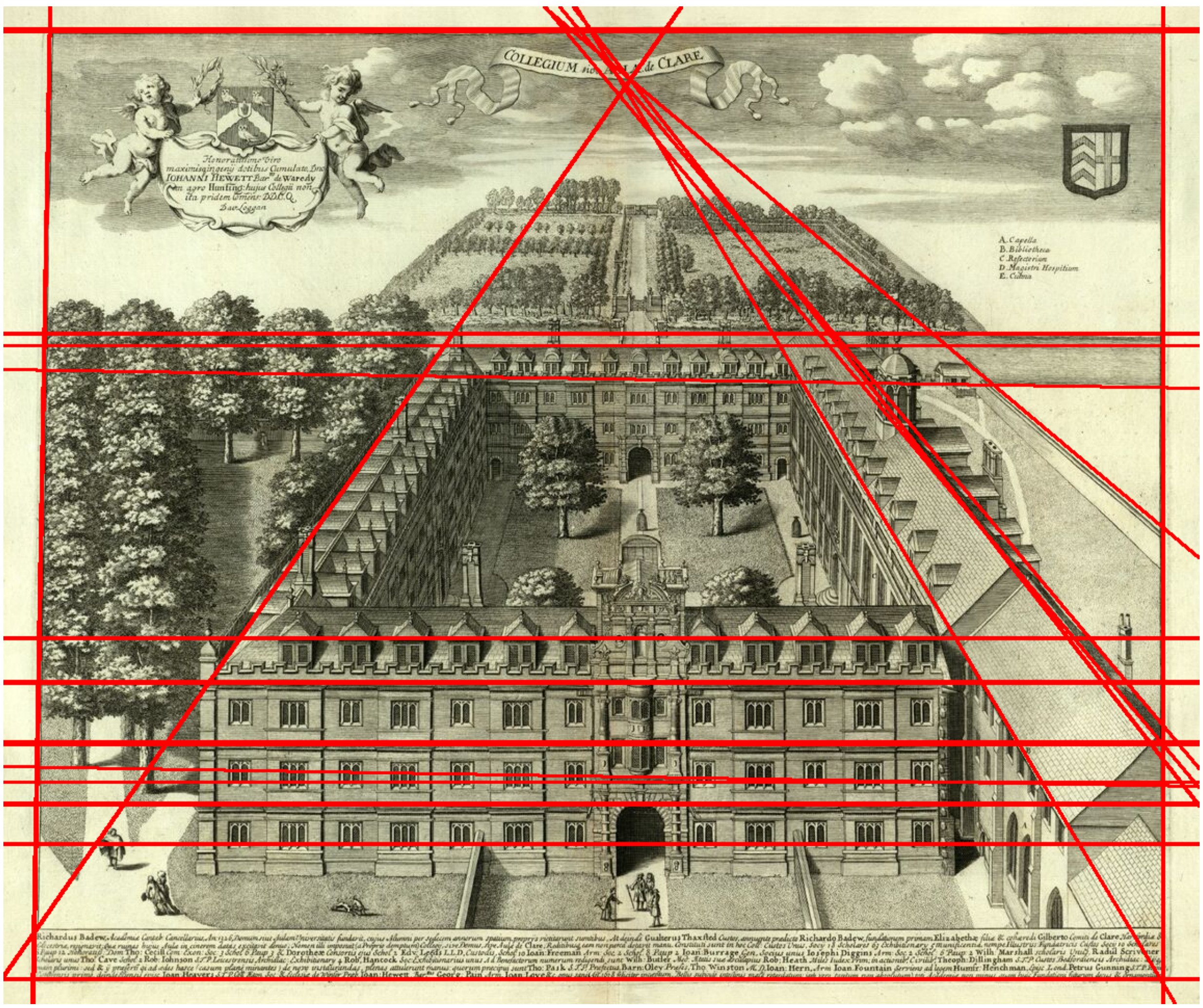

Figure 4: Clare College, extracted lines overlaid.

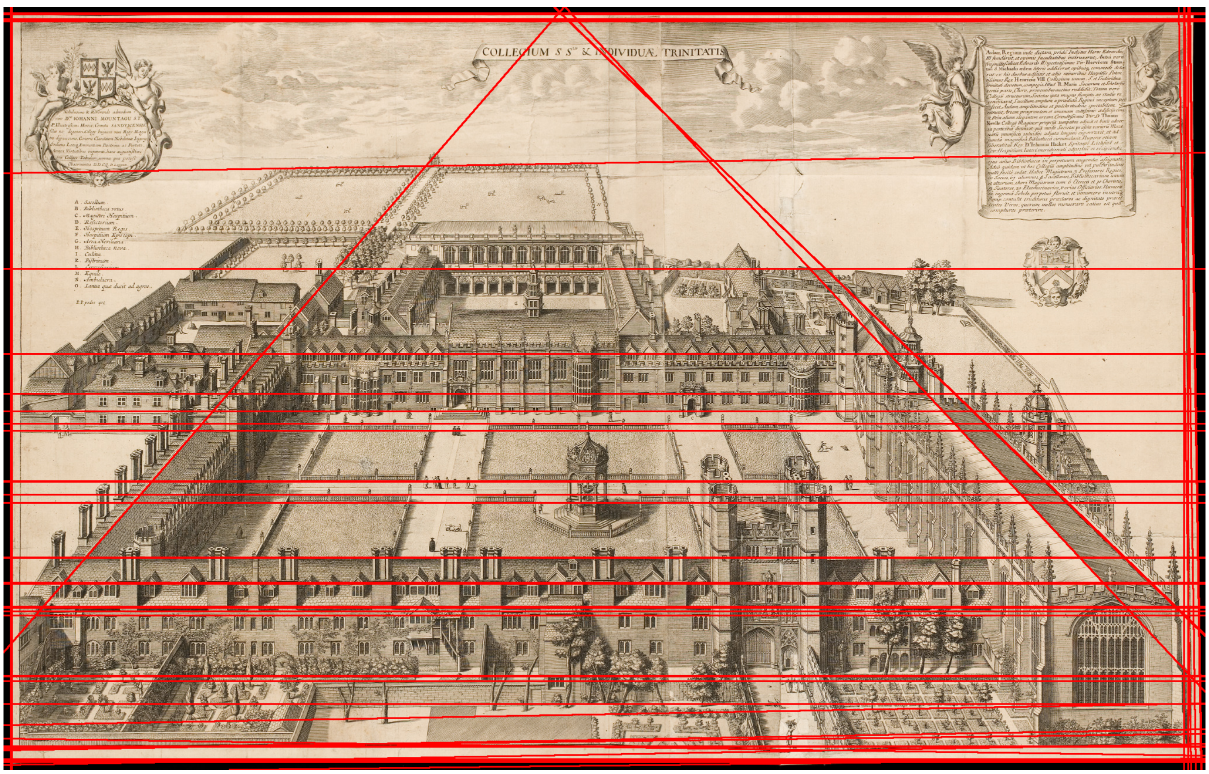

Figure 5: Trinity College, extracted lines overlaid.

camera translation and rotation relative to the model and as such our current workflow 2 requires a manual estimate of the camera position to start the process. To tackle the wire-frame creation problem we have developed new techniques for averaging and clustering geometric objects [4].

3 Discussion

Using historic images to recreate viewpoints of buildings, infrastructure and cities has advantages for monitoring and research. Change over time can give insights into both maintenance and urban development, and the ability to recreate specific viewpoints will enable easier comparison. In this specific project, close inspection of David Loggan's drawings also offers some research questions about David Loggan's methods. There are major and minor differences, many due to time, of course. Some can be put down to perspective, or artistic license. But some are less easily explained. In the Clare College bird's eye view, for example (`https://tinyurl.com/ClareLoggan`), the lantern on the right-hand side is drawn as an octagonal element, whereas in fact it is hexagonal. He also gets the far right hand chimney over the wrong window. These are interesting errors, because neither is present in his ground-drawn perspectival elevation of Clare's hall viewed from within the court (figure 6). Was he paying more attention in one, or were less important elements of the bird's eye drawn by an apprentice? These questions will require deeper comparisons between images, and across more colleges.

In the Trinity image, Loggan straightens out most of the buildings (`https://tinyurl.com/TrinityLoggan`) to give a more organised view. This appears to be clearly artistic license, rather than an error. Is it based on preference for regular courts, or a desire to portray Trinity College in a particular way? Clare College is elongated it is a square court viewed from almost any angle, but Loggan's reproduction is elongated.

4 Conclusions

This project has been an ambitious interdisciplinary blend of the old and new. We have used cutting edge 21st century technology to bring insight into the work of a 17th century engraver sketching 16th century buildings. In order to solve the technical problems that we have faced we have blended hardware, software and historical intuition. Along the way we have identified promising avenues for future research in the areas of automated aerial photogrammetry, geometric scene analysis and infrastructure monitoring.

Figure 6: Clare Hall drawn from within the court

Related and Further Work

On the software and computing front, future work is likely to revolve around improved 3D reconstruction methods incorporating points, lines and planes into a single framework, as well as extending our Geometric Architect system into a fully fledged Simultaneous Localisation And Mapping (SLAM) package for UAV navigation and mapping.

From the historical and architectural perspective there are many interesting questions thrown up as to the working practices of architects and engravers in the 16th century, and David Loggan in particular. Despite being a prolific engraver and artist (the national portrait gallery credits him with 191 portraits) very little is known about him, his techniques or his workshop. The inconsistencies between the sketches and the oldest college buildings are intriguing. Did he go to the buildings themselves or work from incomplete plans? Are sections of the colleges younger than previously thought? Was he simply wrong? Further work is likely to concentrate on attempting to recreate more Loggan engravings from his 1690 book *Cantabrigia Illustrata* as well as delving into the historical archives of the colleges to compare his sketches with architectural drawings of the buildings at a range of time points in the last 300 years.

Acknowledgements

The authors would like to thank Sidney Sussex College, Caius College and the Old Schools for giving permission to fly our drone from their sites. Thanks are also extended to the staff of the Wren Library, Trinity College for providing high resolution images of the Loggan engravings. This project was supported by a mini-projects award from the Centre for Digital Built Britain, under InnovateUK grant number 90066.

References

[1] J Lasenby, WJ Fitzgerald, AN Lasenby, CJL Doran - International Journal of Computer Vision, 1998. *New Geometric Methods for Computer Vision: An Application to Structure and Motion Estimation*

[2] Lasenby, J. and Stevenson, A.X.S. (2001) *Using geometric algebra for optical motion capture*. In: Geometric Algebra with Applications in Science and Engineering. Birkhauser, Boston, MA, USA, pp. 147-167. ISBN 0817641998

[3] Hadfield et al. *REFORM: Rotor Estimation From Object Resampling and Matching* Attached as Appendix A. Submitted to AGACSE2018.

[4] H. Hadfield and J. Lasenby *Direct linear interpolation of conformal geometric objects* Attached as Appendix B. Submitted to AGACSE2018

Appendix A

Direct linear interpolation of geometric objects in conformal geometric algebra

Typically we do not add objects in conformal geometric algebra (CGA), rather we apply operations that preserve grade, usually via rotors, such as rotation, translation, dilation, or via reflection and inversion. Direct linear interpolation of conformal geometric objects can be both intuitive and useful. Here we discuss the empirical results of interpolating objects of a range of grades and present two algorithms of interest for computer vision applications that benefit from the direct averaging of geometric objects.

Appendix B

REFORM: Rotor Estimation From Object Resampling and Matching

Our goal is to find the correct rotation and translation between a reference model and a query model. Both models are made up of geometric primitives. Our reference might be, for example, a CAD model, and our query model might represent the output of fitting primitives to LIDAR data or structure from motion point clouds. It was shown in [**?**] that by respresenting our objects in conformal geometric algebra (CGA) and using an appropriate rotor magnitude-based cost function, we can calculate the conformal rotor that takes one model into another, assuming the objects within them are correctly matched. Here we consider the case in which there is an incorrect initial matching of objects within the models. The cases that we consider are those in which our query model: contains additional primitives not present in the reference; is missing primitives that are present in the reference. We will also look at cases in which there are a large number of primitives per model. These are all common issues facing any SLAM-type (Simultaneous Localisation And Mapping) systems. To overcome these problems we introduce an *inter-object rotor magnitude-based matching function* and a *subsampled iterative rotor estimation and matching algorithm* - **REFORM**. REFORM is robust to noise in the models, artefacts and, due to the subsampling step, is computationally cheap for large models and easily parallelisable.