Jon Slade, Christopher B Jones & Paul L Rosin - EuroSDR/ISPRS Workshop on 'Efficient Capturing of 3D Objects at a National Level: with a Focus on Buildings and Infrastructure'

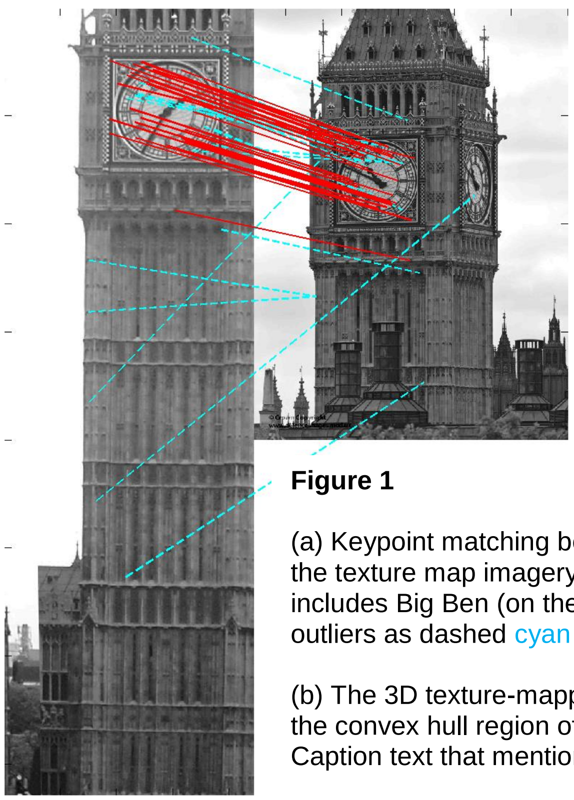
ABSTRACT: Semantic and geometric enrichment of 3D geo-spatial building models with captioned photos and labelled illustrations

Currently 3D building models do not provide much if any descriptive information associated with their geometry (Jones *et al.*, 2014). Consequently users of mobile and desktop applications which make use of such models are unable to obtain information about the architectural features, origins and material properties of the buildings. This research plans to enrich 3D models with annotations from object-specific descriptions from photo captions and corresponding plans and diagrams.

With rich semantic content available imagine a mobile application which allows the user to explore a city, neighbourhood or building via augmented reality learning of important cultural heritage and architectural features through superimposed annotations. Or as a practitioner of the emergent field of Building Energy Modelling (BEM) being able to model building energy loss across entire neighbourhoods or cities using a model attributed with building material. Or as a radio network planner using that material information to produce a more refined planning model which takes account of the variation in radio wave attenuation by material (de Fornel and Sizun, 2006) thus enabling the most effective placement of the smallest number of transceivers for the best radio coverage.

The enrichment will be facilitated through the use of feature detection and matching techniques from the field of computer vision such as SIFT (Lowe, 2004). Feature (or keypoint) matching will be achieved by obtaining correspondence between the texture maps on the surfaces of the existing 3D building models and the supplementary captioned photos, such as those on *Flickr*. See Figure 1. Computer vision template matching allied to machine learning techniques such as mutual information (Walter and Fritsch, 1999) or Bayesian learning (Jones *et al.*, 1999) may be used to identify objects within the texture maps. This matching will present opportunities to enrich the geometry of the building model.

(a)



(b)

Figure 1

(a) Keypoint matching between a *Flickr* image of Big Ben (on the right) and the texture map imagery of a 3D model of the Houses of Parliament that includes Big Ben (on the left). Inlier matches are coloured as red lines and outliers as dashed cyan lines.

(b) The 3D texture-mapped model of the Houses of Parliament in which the convex hull region of the matching inlier key points is highlighted in red. Caption text that mentions clock can be linked to this region.

Should floor plans of the building be available then vectorisation, 2D-3D conflation with the 3D models and OCR text extraction from the plans would provide the potential for enriching building interiors.

Once a match is achieved then the object annotation can be placed in to the building model. Subsequent research may also include the use of natural and spatial language processing to optimally locate the captions.

Acknowledgements

Jon Slade is a PhD student at Cardiff University, GB, funded by an EPSRC Industrial CASE studentship with Ordnance Survey, GB.

References

DE FORNEL, P. AND SIZUN, H., 2006. *Radio Wave Propagation for Telecommunication Applications*. [online]. Berlin Heidelberg. Springer. 436 pages. pp. Available from: http://books.google.co.uk/books?id=x5wWcgYBdI0C&lpg=PA171&ots=EARomx4tq-&dq=radio%20wave%20attenuation%20by%20material&pg=PA170#v=onepage&q=table%206.1&f=false [Accessed 11 September 2014].

JONES, C., ROSIN, P. AND SLADE, J., 2014. Semantic and geometric enrichment of 3D geo-spatial models with captioned photos and labelled illustrations [online]. *In:* I. Press, (Ed). *Workshop on Vision and Language (VL4)*, Dublin. Available from: http://www.aclweb.org/anthology/W/W14/W14-54.pdf#page=74 [Accessed 18 August 2014].

JONES, C. B., WARE, J. M. AND MILLER, D. R., 1999. A Probabilistic Approach to Environmental Change Detection with Area-Class Map Data. [online] *In:* P. Agouris and A. Stefanidis (Eds.) *Integrated Spatial Databases.* Springer Berlin Heidelberg. pp. 8. Available from: http://link.springer.com/content/pdf/10.1007%2F3-540-46621-5_8.pdf [Accessed 11 June 2014].

LOWE, D. G., 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* [online], 60(2): 91-110. Available from: http://link.springer.com/content/pdf/10.1023%2FB%3AVISI.0000029664.99615.94.pdf [Accessed 12 June 2014].

WALTER, V. AND FRITSCH, D., 1999. Matching spatial data sets: a statistical approach. *International Journal of Geographical Information Science* [online], 13(5): 445-473. Available from: http://www.tandfonline.com/doi/pdf/10.1080/136588199241157 [Accessed 11 June 2014].