

Divya Udayan J, HyungSeok Kim, Jee-In Kim, 2015. An image-based approach to the reconstruction of ancient architectures by extracting and arranging 3D spatial components. *Frontiers of Information Technology & Electronic Engineering*, **16**(1):12-27. [doi:10.1631/FITEE.1400141]

An image-based approach to the reconstruction of ancient architectures by extracting and arranging 3D spatial components

Key words: Digital reconstruction, 3D virtual world, 3D spatial components, Vision and scene understanding

Contact: HyungSeok Kim,

E-mail: hyuskim@konkuk.ac.kr

 ORCID: <http://orcid.org/0000-0003-4816-2992>

Introduction

- Recovering 3D geometry from 2D photograph has infinite possibilities.
- To resolve the problem, we propose a semi-automatic and interactive reconstruction methodology that leverages techniques from modern digital technologies and image processing algorithms.
- Validation of our approach and test results of three well-known monuments from the Mughal architecture style have been presented in this work.

Pipeline of the reconstruction methodology proposed in this work

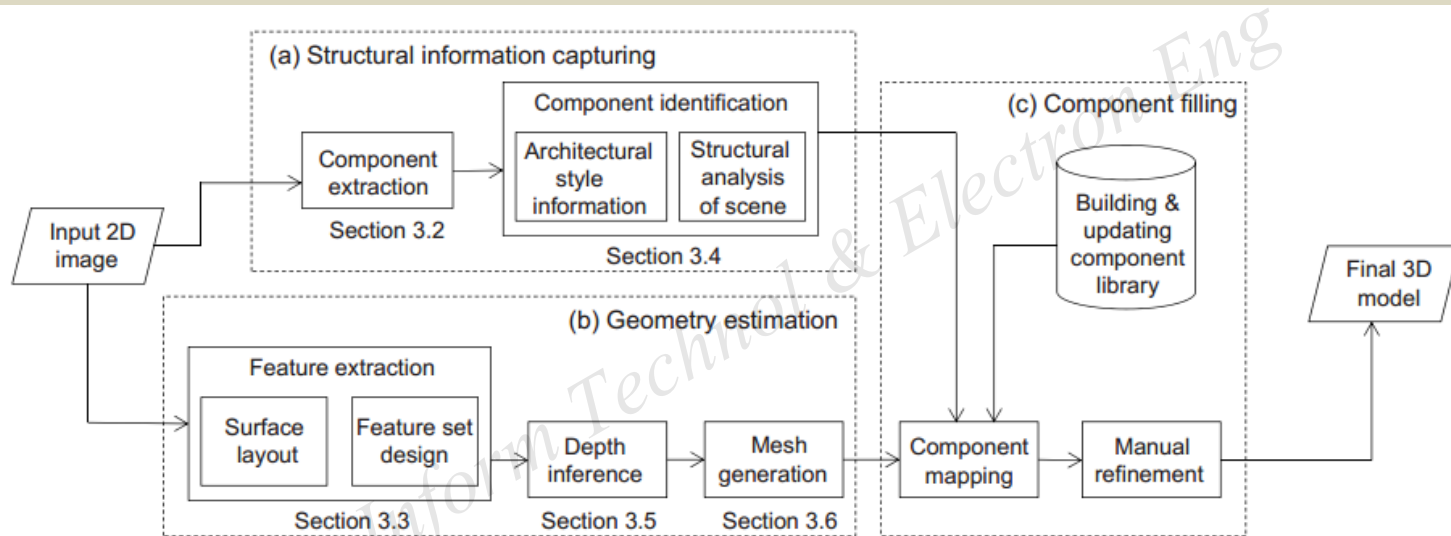


Fig. 2 Pipeline of the overall process: (a) structural information capturing, (b) geometry estimation, and (c) component filling

Structural information capturing

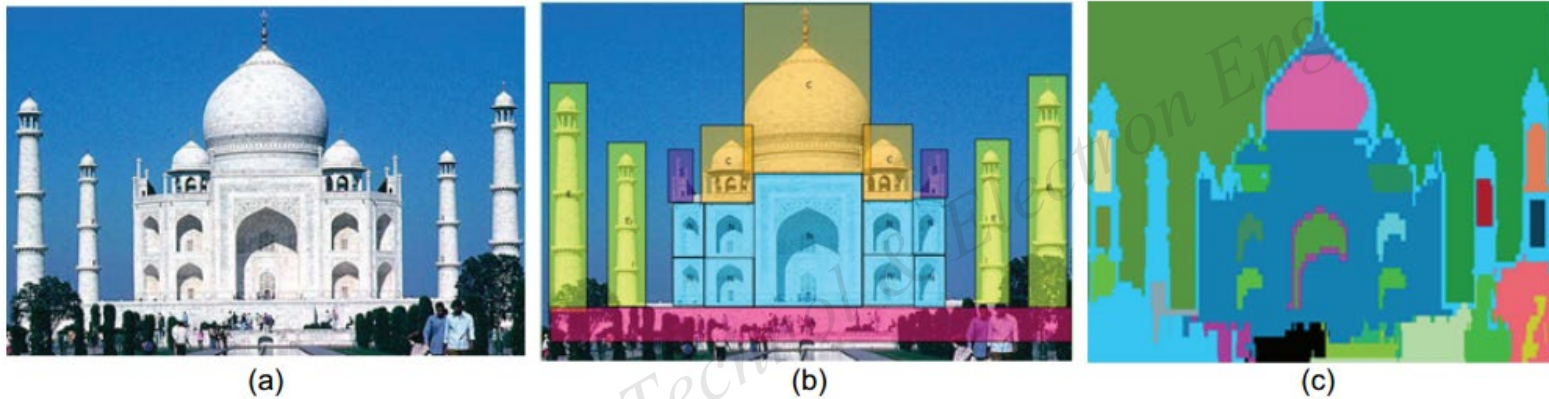


Fig. 3 Component extraction results of the input image: (a) digital input image, (b) identified components and their repetition shown in a unique color, and (c) feature-based surface layout results

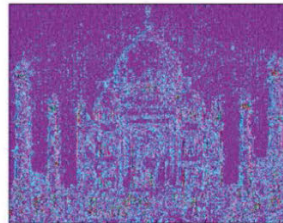
Feature extraction



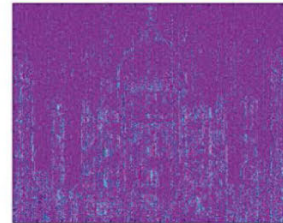
(a)



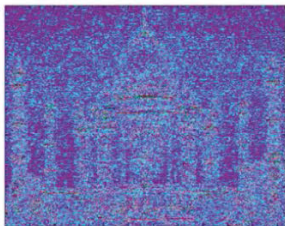
(b)



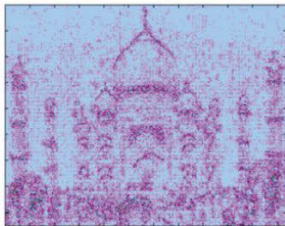
(c)



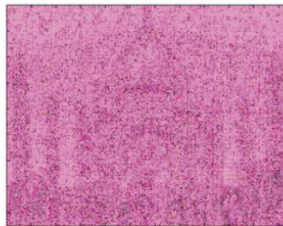
(d)



(e)



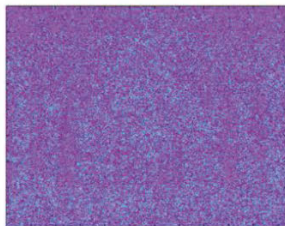
(f)



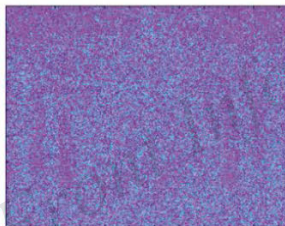
(g)



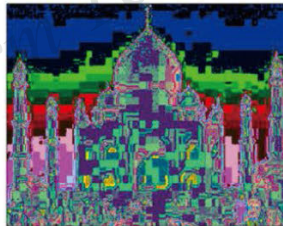
(h)



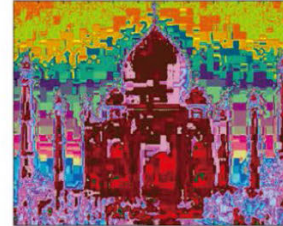
(i)



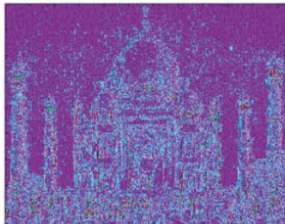
(j)



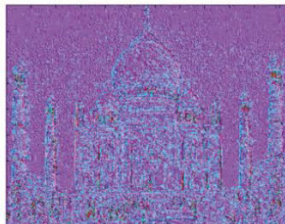
(k)



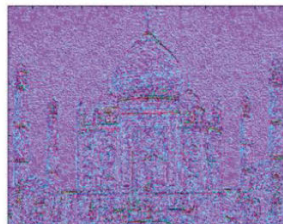
(l)



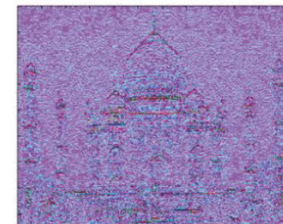
(m)



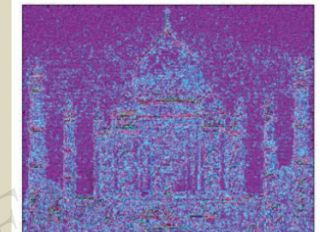
(n)



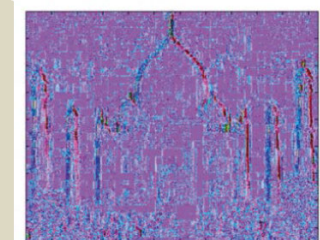
(o)



(p)



(q)



(r)

Fig. 7 Outputs of the filters of a digital input image of the Taj Mahal. (a) is the input image; (b)–(d), (e)–(g), and (h)–(j) are the outputs of the Laws filter mask for local averaging, edge detection, and spot detection, respectively; (k)–(l) are the outputs of the local averaging filters applied to the Cb and Cr channels, respectively; and (m)–(r) are the filter outputs for the six different directions of the texture gradients

Depth and mesh generation

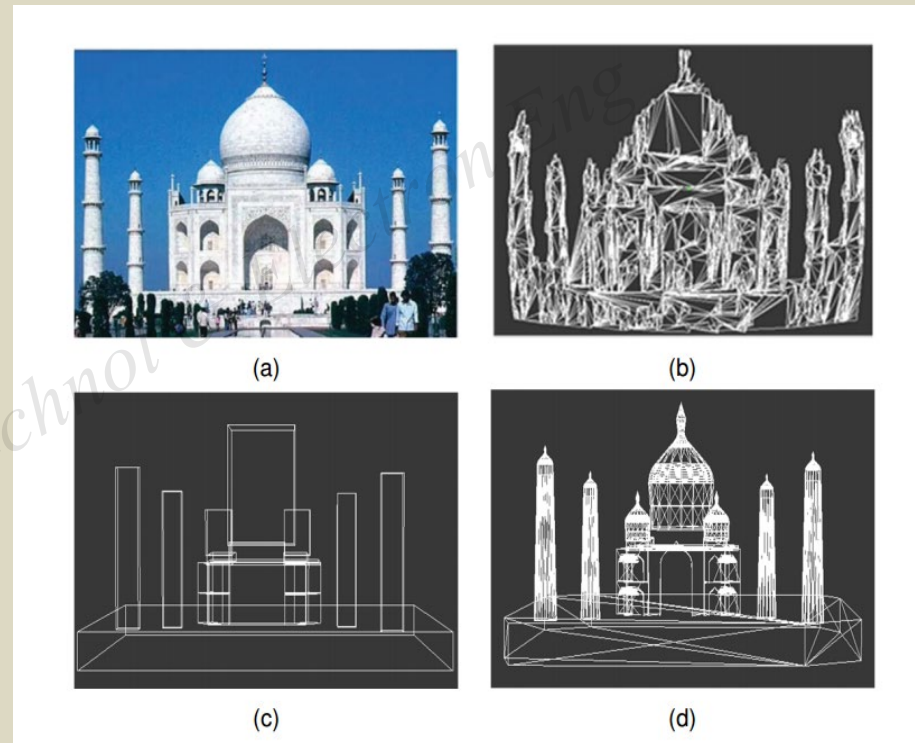
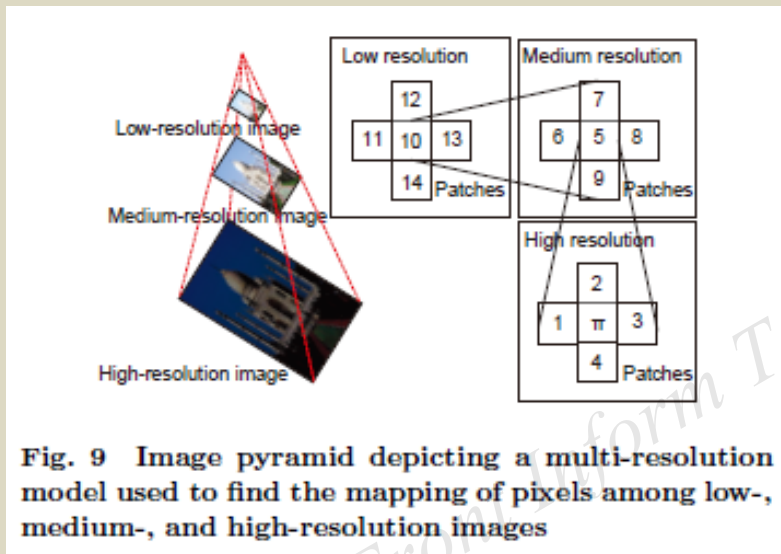


Fig. 10 Implementation results of the 3D reconstruction pipeline: (a) digital input image, (b) initial automatic reconstruction results, (c) bounding box showing the estimated symmetry, (d) wire-frame model of the final reconstruction

Component filling and reuse for same style buildings

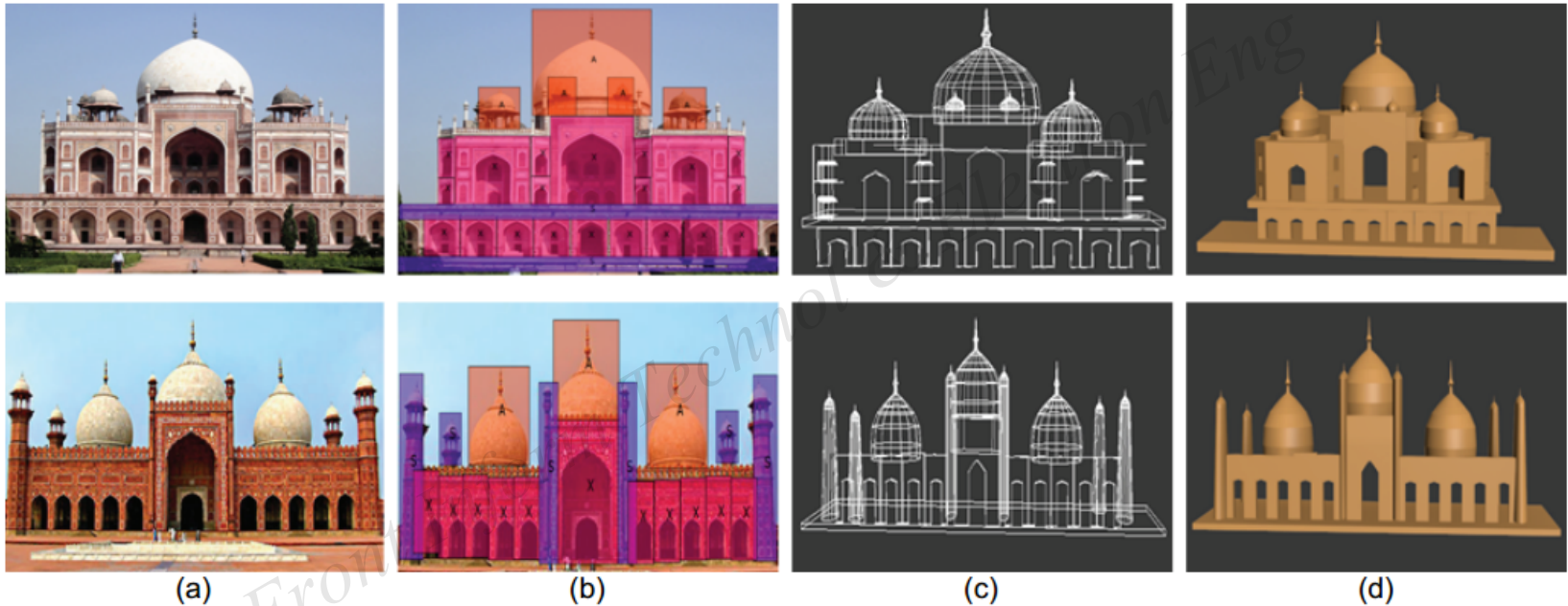


Fig. 12 Two typical examples of Mughal architecture. The first and second rows show the reconstruction results of Humayun's Tomb and the Badshahi Mosque, respectively. (a) A single digital image, (b) semi-automatic component identification (where each randomly generated color represents one unique component and its repetition), (c) the wire-frame of the reconstructed model, and (d) the final reconstructed model.

Performance evaluation

Table 3 Comparison of the displacement errors measured using Hausdorff distances relative to the reference model

| User | Mean (%) | Maximum (%) | RMS (%) |
|------|----------|-------------|----------|
| A | 4.583654 | 5.566855 | 4.623318 |
| B | 4.578231 | 5.566855 | 4.618944 |
| C | 4.584550 | 5.566855 | 4.624117 |
| D | 4.586463 | 5.540374 | 4.623213 |
| E | 4.580921 | 5.540374 | 4.617344 |

RMS: root mean square

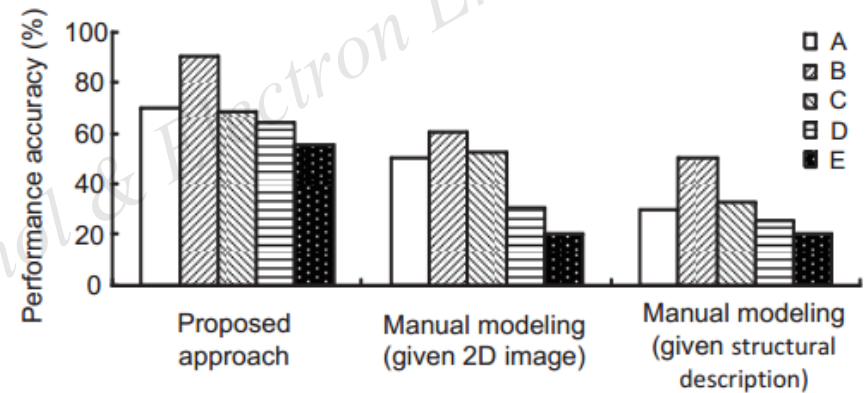


Fig. 13 Performance in terms of the accuracy of the modeling sessions of our user study. Each of the five bars represents a different user and the corresponding performance accuracy for the different sessions

Conclusions

- This approach is especially of value if just one image of an object and no other photographic information is available.
- The components generated to form a component library of a specific architectural style can be reused for reconstructing additional monuments of the same design style.
- The performance in terms of speed and accuracy was observed to be better in our proposed approach compared to manual modelling.