Anupama Mallik · Santanu Chaudhury
Vijay Chandru · Sharada Srinivasan
Editors

Digital Hampi: Preserving Indian Cultural Heritage

Springer

Digital Hampi: Preserving Indian Cultural Heritage

Anupama Mallik   Santanu Chaudhury
Vijay Chandru   Sharada Srinivasan
Editors

Digital Hampi: Preserving Indian Cultural Heritage

Technical Contribution from K.R. Murali Mohan,
Department of Science and Technology (DST)

Springer

ditors

Anupama Mallik
Department of Electrical Engineering
Indian Institute of Technology, Hauz Khas
New Delhi
India

Vijay Chandru
International Institute for Art, Culture and
  Democracy
Bengaluru, Karnataka
India

Santanu Chaudhury
CSIR-CEERI, Pilani and Department of
  Electrical Engineering
Indian Institute of Technology, Hauz Khas
New Delhi
India

Sharada Srinivasan
National Institute of Advanced Studies, IISc
  Campus
Bengaluru, Karnataka
India

Foreword

India is a nation with a rich heritage. Even after centuries of deterioration, its historical sites bear several unique archaeological features and are major repositories of cultural knowledge and artefacts. The inhabitants of such sites and people living around them have kept alive several living traditions. Yet, rapid urbanisation and developmental activities associated with modern technology threaten to quickly erode the rich legacy of past generation. However, at the same time, modern technology has given us tools and techniques which can be harnessed for the preservation of heritage for posterity. We are at the edge of a precipice since heritage once lost is lost forever'. Even within another decade or two, our living traditions will be signi cantly diminished. We must act today. This book is the outcome of an attempt to bring alive the tangible and intangible heritage by 3D digital recreation and use of virtual reality towards visualisation with an effective and immersive communication and interaction.

Department of Science and Technology (DST), Government of India, supported a team of scientists, art historians and cultural heritage scholars in different institutions around the country to lead this task of not only developing the required technologies, but also applying them to the World Heritage Site of Hampi, as a case study, with attempts at graphical, pictorial and digital reconstruction. These scientists belong to the computer graphics and vision community. The heritage practitioners and art historians involved had decades of experience in cultural studies. In addition, experts in architectural design were roped in to lead the representations of restored and recreated heritage. The biggest challenge was that such a team with diverse expertise had no experience of working together. DST had, in the past, worked only with science and technology personnel. But the need and the opportunities were overwhelming. Taking together the passion of the scientists and cultural experts who worked on this project and the mentoring and encouragement of the DST's programme monitoring committee, the project came together successfully. The outcome was the Making of Digital Hampi', as an attempt to bridge culture and technology.

Hampi is one of the most spectacular sites in India with a range of religious and secular architecture. Indian artistic and temple architectural traditions give us rare

insights into design, construction, proportion and scale. The architecture of Hampi and Dravidian temple styles forms an impeccable synergy between structural innovation and architectural expression. Anegundi, the capital of the Vijayanagara Empire from the fourteenth to sixteenth century, boasts not only of spectacular ruins but also of a rich living tradition. The granite ruins of Hampi have several damaged and broken sculptures. The murals and paintings are in various stages of deterioration. But this did not deter the team. In fact, the team took it as an opportunity. Starting with the acquisition of 3D data and followed by 3D scanning of the remnants, they used existing tools and developed new ones to bring alive the broken statues and artefacts. Rather than attempting an authentic regeneration down to the minute detail, the focus was on lling-in or reconstruction of the damaged and defaced region and cracks in heritage monuments with least reconstruction errors and verifying scenarios using on-ground information. They have carried out digital restoration and rendering of mural paintings offering interactive interfaces, and created personalised virtual avatars using evolving haptic technology, both as interaction proxies for real users and as virtual inhabitants for digitally recreated worlds.

This volume is an outcome of this extensive work. It is a documentation of diverse technologies that are available today for this purpose. But it is much more than a book on technology. There is also a signi cant intangible heritage component through the study of the present-day festivals and rituals and links with representation in old murals, paintings and inscriptions. Hampi architecture has been digitally interpreted for interaction and understanding by a common man and enhanced with digital walk-throughs to bring alive the cultural heritage objects through virtual touch. Temple elements have been captured graphically, pictorially and digitally, in turn, making the whole digital experience navigational and participatory to all users. Architectural and Social Life reconstructions of Bazaar Streets of Hampi in Vijayanagara period have been attempted. The aspects of digital rendition have been brought together with iconographic and iconometric studies to better document and enhance the understanding of Vijayanagara sculpture.

Finally, for those who are familiar with the site of Hampi, it provides a ceaseless exercise of imagination. The project represents a successful endeavour to create a digital and cultural knowledge bank with digital archives on the heritage of Hampi, whereby fresh data can be constantly updated by researchers, practitioners and the broader public.

New Delhi, India   Ashok Jhunjhunwala
                                                                                          Principal Advisor
                                                                            Minister of Power and MNRE
                                                                                     Government of India
                                                                      Professor IIT Madras (on sabbatical)

Preface

In the 1972 UNESCO convention on heritage, heritage was de ned as "encompassing tangible and intangible, natural and cultural, movable and immovable and documentary assets inherited from the past and transmitted to future generations by virtue of their irreplaceable value." While to most citizens, heritage refers to monumental remains of cultures, the concept of heritage has gradually been expanded to embrace living culture and contemporary expressions. The World Heritage Site at Hampi, recognized by UNESCO, is an outstanding example of the austere and grandiose monuments of India.

Department of Science & Technology (DST) Govt. of India initiated the Indian Digital Heritage (IDH) Research Project in 2010 with an aim to extend the power of digital technologies to digitally capture, preserve, and restore all forms of tangible and intangible cultural and historical knowledge. While archiving and disseminating digital representations of heritage artefacts and cultural traditions, the emerging multimedia technologies in computer vision and user interface design would make possible immersive experiences of heritage and possibly inspire young citizens to participate in similar projects around the country. India is rich in cultural heritage with hundreds of important archaeological sites and rich traditions that need to be digitally preserved. The recent advances in digital technologies open up the possibility of creating rich digital representations of the heritage sites which can

be preserved for perusal by world citizenry for the foreseeable future. In addition, digital restoration of damaged monuments, digitally conjured animations, and augmented reality representations of social life of past eras are intriguing creative possibilities today. The objective of DST was also to build capacity in academia to create analytic tools for the art historians, architects, and other scholars in the study of the heritage of India.

There was an early realization that the goals of the IDH project required teams with capacity in the "best in class" digital technologies ranging from laser scanning, 3D printing to mature and novel Information and Communication Technologies to come together and work shoulder to shoulder with social scientist like art historians, archaeologists, architects, anthropologists, and digital humanities.

The IDH project was successful in its main objective of creating a new synergy between the art and science communities for developing new frameworks and solutions to preserve heritage in digital space. The innovative techniques and research outcomes of the IDH project are detailed in the various chapters of this book. I am con dent that this book will become an invaluable resource for sci-entists, heritage scholars, culture historians and academicians, and indeed for anyone interested in learning about the heritage of Hampi and efforts undertaken for its digital preservation. The technologies outlined here are replicable and scalable for application in digital preservation of heritage artifacts anywhere in the world, and I wish success to all such efforts in future.

New Delhi, India   Prof. Ashutosh Sharma
Secretary, Department of Science and Technology
Ministry of Science and Technology and Earth Sciences
Government of India

Contents

Part VI Conclusion

About the Editors

Anupama Mallik received her B.Sc in Physics (1986) and her Master's in Computer Applications (1989) from Delhi University, subsequently completing her PhD (2012) in Electrical Engineering at the Indian Institute of Technology (IIT), Delhi. Her PhD thesis dealt with ontology-based exploration of multimedia contents. She is currently associated with the Multimedia Research Group of the Electrical Engineering Department, IIT Delhi and has worked as a Research Scientist in projects sponsored by the Department of Science and Technology, Government of India. She has been teaching courses related to semantic web technologies as visiting faculty at several universities of the Delhi government. She is a member of the Association for Computing Machinery (ACM) and a co-author of the book Multimedia Ontology: Representation and Applications' published by the CRC Press. She is the founder-Director of a technology start-up focusing on the application of technology in heritage preservation, currently being incubated at IIT Delhi.

Santanu Chaudhury completed his B.Tech (1984) in Electronics and Electrical Communication Engineering and his PhD (1989) in Computer Science and Engineering from Indian Institute of Technology (IIT), Kharagpur. He is a Professor at the Department of Electrical Engineering, Indian Institute of Technology (IIT), Delhi. He was also the Dean of Undergraduate Studies at the IIT Delhi. He was awarded the INSA medal for young scientists in 1993. He is a Fellow of the Indian National Academy of Engineering and National Academy of Sciences, India. He is a Fellow of International Association of Pattern Recognition (IAPR). He has more than 250 publications in international journals and conference proceedings, and has authored /edited several books. He is currently working as Director of CSIR-Central Electronics Engineering Research Institute (CSIR-CEERI) and is on lien from IIT Delhi.

Vijay Chandru is an academic and an entrepreneur. As an academic, he completed his PhD in the Mathematics of Decision Sciences at MIT in 1982 and has taught at Purdue University (1982 1993) and the Indian Institute of Science (IISc) since 1992. He has co-authored and co-edited several books and published over 90

peer-reviewed research papers. In 1996, he was elected as a Fellow of the Indian Academy of Sciences and subsequently of the Indian National Academy of Engineering. As co-convenor of the Perceptual Computing Laboratory (PerCoLat) at IISc for a decade (1995 2005), he participated in semantic web language development, in advances in visualisation for scienti c data and in development of the Simputer. The International Institute for Art, Culture and Democracy (IIACD) is a Bangalore-based research institute, which was found in 2008 by a multidisciplinary group of academics and practitioners including him. At IIACD, his work in digital humanities has been in the representation, organisation and investigation of knowledge-bearing artefacts, which have been documented, annotated and archived. As a technology entrepreneur, Professor Chandru serves as the Chairman of Strand Life Sciences, a spinoff from IISc. He was named a Technology Pioneer of the World Economic Forum in 2007 and for contributions to science and society, he was awarded the Hari Om Trust Award by University Grants Commission in 2003.

Sharada Srinivasan Professor and Dean, School of Humanities at National Institute of Advanced Studies, Bangalore, specialises in technical art history, archaeological sciences, archaeometallurgy, archaeometry and material culture studies. She has a PhD from the Institute of Archaeology, University College London (1996) on Archaeometallurgy of South Indian bronzes; MA from School of Oriental & African Studies, London (1989) and B.Tech in Engineering Physics from IIT, Mumbai (1987). She is Fellow of the Royal Asiatic Society of Great Britain and Ireland, World Academy of Art and Science, and has been Homi Bhabha Fellow, Charles Wallace and V&A Nehru Fellow and Smithsonian Forbes Fellow at Freer Gallery of Art. Her awards include the Dr. Kalpana Chawla Young Woman Scientist Award, the Indian Institute of Metals Certi cate, Materials Research Society of India Medal, Malti B. Nagar Ethnoarchaeology Award, DST-SERC Young Scientist Award and Flinders Petrie Medal. She was the recipient of British Council funded UKIERI grants for studies on Indian wootz steel and for split-site PhDs with Exeter University in Archaeology and Drama. She is a member of the Advisory Boards of Centre for South Asian Studies, University of Exeter, Institute of Archaeometallurgical Studies, London, Ethnoarchaeology Journal, USA and Standing Committee member of the Beginning of Use of Metals and Alloys (BUMA) conference series. She is the rst author of India's legendary wootz steel' and more than 70 papers and is also a reputed exponent of classical South Indian dance. She was featured as one of India's top ten women scientists in the Conde Nast magazine Make in India' issue of January 2017.

Part
ntroduction

Introduction

**Santanu Chaudhury Anupama Mallik Vijay Chandru
and Sharada Srinivasan**

1 Prologue: What Is Heritage?

Heritage represents everything of the past for which we can be proud of today. UNESCO categorizes heritage into two broad classes—Cultural heritage and Natural heritage [1].

According to the UNESCO model, cultural heritage covers the following:

(a) Monuments: architectural works, works of monumental sculpture and painting, elements or structures of an archaeological nature, inscriptions, cave dwellings and combinations of these features;

(b) Groups of buildings: groups of separate or connected buildings which, because of their architecture, their homogeneity or their place in the landscape are objects of distinguished interest;

(c) Sites: works of man or the combined works of nature and of man, and areas including archaeological sites which are important from the historical, aesthetic, ethnological or anthropological points of view. These also include underwater cultural heritage instances like shipwrecks, underwater ruins and cities.

S. Chaudhury   A. Mallik ( )
IIT Delhi, New Delhi, India
e-mail: anupamamallik@acm.org

S. Chaudhury
e-mail: santanuc@ee.iitd.ac.in

V. Chandru
IIACD, Bengaluru, India
e-mail: chandru@iiacd.org

S. Srinivasan
NIAS, Bengaluru, India
e-mail: sharasri@gmail.com

Categories (a), (b) and (c) together, are, in general, referred to as tangible cultural heritage. The last category is as follows:

(d) Intangible cultural heritage includes oral traditions, performing arts, rituals, culinary traditions, etc.

Natural heritage refers to

(a) natural features consisting of physical and biological formations or groups of such formations;
(b) geological and physiographical formations and precisely delineated areas which constitute the habitat of threatened species of animals and plant;
(c) natural sites or precisely delineated natural areas which are important from the point of view of science, conservation or natural beauty.

UNESCO also defines cultural landscapes which are illustrative of the evolution of the human society and settlements over time under the influence of the physical and natural constraints representing the "combined works of nature and of man" [1].

UNESCO classifies some of these heritage elements as having "Outstanding Universal Value". Those heritage elements are distinguished by their cultural and/or natural significance, which transcend national boundaries and are of common importance for present and future generations of all humanity. Such elements are candidate for inclusion in the World Heritage List.

The World Heritage Committee of UNESCO in its tenth session in 1986 included the Hampi group of monuments in the World Heritage list. Currently, this is considered as an endangered World Heritage of Outstanding Universal value. This book is concerned with the preservation of tangible and intangible cultural heritages of Hampi in the digital space.

2   Digital Preservation of Cultural Heritage

Affordances of the digital technologies have made digital media as the ideal choice for the storage, representation, management and communication of cultural heritage. Tangible cultural heritage components like monuments, temples, groups of buildings and sites of historical significance can be scanned, modelled and archived. 3D scanning is a technology for capturing spatial data in three dimensions. 3D scanned models provide the data for analysis and visualization in virtual environments. These contents can be accessed through powerful search engines and database management tools. As a consequence, we create the possibility of disseminating the content through the world wide web (www) to audiences who otherwise may never be able to access or visit the site. Sometimes, 3D content is brought back into the physical world via 3D printing so that audience can physically experience the heritage artefact. There can also be augmented reality-based presentation around the re-created physical models [2]. In addition to the structural models represented by the 3D scan data, specialized imaging techniques are

required for faithful digital recording of colour and texture of the surfaces. In particular, where there are paintings (murals, frescos, etc.) on the walls or ceilings, capturing images of these with true colour tones in complete details is a necessity for authentic representation of the content in the digital space. Further, this data has to be linked with the structural model for true modelling and rendering of the tangible heritage elements.

Digital preservation of intangible heritage involves not only capturing the data on digital media but also preservation of the knowledge and processes involved in it. Much of intangible heritage are actually form of embodied practice and its preservation in digital space means capturing the knowledge of the practice so that future enactment of the practice becomes a possibility. It is also important to link digital preservation initiatives for tangible and intangible to completely preserve the cultural heritage of a site. A tangible cultural artefact may be related to intangible knowledge and practices in such a way that preserving the artefact without preserving the associated intangible heritage may result in loss of intellectual context. Digital representation not only preserves a digital representation of intangible heritage but also provides unique opportunity to present the heritage in appropriately contextualized fashion through walk-throughs in virtual spatial and intellectual spaces.

The UNESCO World heritage site Hampi was the capital of the Vijayanagara empire situated on the banks of river Tungabhadra in the Indian state of Karnataka. Grandeur and beauty of the remains of the palaces, the temples and the royal buildings of Hampi deserve special efforts for physical and digital preservations. Further, the site is associated with the Hindu mythological landscapes of Pampakshetra and Kishkindha (territory of Bali, Sugreeva and Hanuman in the Ramayana). Besides historical and mythological signi cance, the environment at Hampi has a rich combination of natural context, cultural and social traditions. Hence, representation of the cultural heritage of Hampi requires innovations in techniques and processes for capturing both tangible and intangible heritages of Hampi.

3 Some Initiatives for Digital Preservation of Cultural Heritage

We are looking at different initiatives of project Digital Hampi, as described in this book, against the backdrop of increasing interest and attempts at digital preservation of endangered cultural heritage across the world.

With the availability of relatively cheap but ef cient depth sensing devices, there have been a number of research projects across the world for 3D digitization of tangible cultural heritage. Here, we describe some of the projects from around the world [3], which were a precursor to Digital Hampi, to contextualize the work presented in this book.

The Digital Michelangelo Project by Levoy et al. [4] developed the pipeline for accurate 3D digitization of 10 statues created by Michelangelo, two building

interiors and 1163 fragments of an ancient marble map in Italy. The project made use of triangulation laser scanners, time-of- ight laser scanners and digital cameras. The key challenge of this project was dealing with large number of polygons (order of two billion polygons for the largest single dataset) for mesh representation of the 3D scan data. Ikeuchi et al. [2] have developed a technology which deals with issues that arise due to an outdoor environment and due to the attempts for registering 3D data obtained by scans taken from different directions. This work has focussed on archival of three large Buddha statues in Japan, which measure 2.7, 13 and 15 m, respectively, using a time-of- ight laser scanner.

In another project in Italy done by Fontana et al. [5], a 3D model of the Minerva of Arezzo was created employing triangulation using a high-resolution 3D laser scanner. The Minerva of Arezzo is an ancient bronze statue approximately 1.55 m in height. The project focussed on keeping track of the variations during the restoration process of the statue. Bernardini et al. in their Pieta Project [6] worked on the digitization of a large marble sculpture, with a height of 2.26 m, created by Michelangelo, and also located in Italy. Key contribution of this work was a Bidirectional Re ectance Distribution Function (BRDF) to estimate the intrinsic properties of the surface. Digitization of the Cathedral of Saint-Pierre in Beauvais, France was the focus of Beauvais Cathedral project [7]. Grün et al. in [8] present the photogrammetric techniques used in the project—the Great Bamiyan Buddha—for digitization of two large Buddha statues, 38 and 53 m, respectively, in height, located at Bamiyan in Afghanistan.

Some of the other projects have looked at larger sites, going beyond a particular artefact. The Eternal Egypt Project [9] is an example of such a project. This project was aimed at creating a digital guide and a virtual museum of artefacts based upon data collected using a time-of- ight range sensor and a digital imaging system for high-quality colour images. There was an attempt [10] for 3D modelling of one thousand heritage structures in Angkor, Cambodia using aerial photogrammetric methods.

Prof. Katsushi Ikeuchi and his team [11] undertook a project on digitization of the Bayon temple, in Angkor ruins, Cambodia. The Bayon temple is a large and complicated structure. The temple required development of special range sensors for capturing data of many hidden areas of the temple. Through this project, it was possible to successfully model the Bayon temple, a 150 m by 150 m by 40 m structure, in a digital form with a 1 cm resolution. This was extended to other ruins of Angkor Wat. These modelling results have been also used for archaeological research. The Bayon temple has 173 carved smiling faces. For example, using clustering, these faces have been grouped into three classes—Deva, Devata and Ashura as described in the related mythologies. This project has been a landmark in digital preservation of tangible cultural heritage.

Digitization of the tangible cultural heritage in terms of 3D models has given rise to a relatively new branch of knowledge [12]—3D Cultural Heritage or Virtual Heritage that utilizes information technology to capture or represent the data studied by archaeologists and historians of art and architecture. In addition to representation and archiving, 3D model construction enables the following:

(1) Measurement of the existing objects automatically using 3D capture technologies such as laser scanning or photogrammetry.
(2) Reconstructive modelling of damaged or no-longer-extant objects by manual (using software like AutoCAD, 3D Studio Max, Maya) or algorithmic interventions.
(3) Combination of captured and reconstructed models to create hybrids which are hypothesized representations of damaged artefacts.

These representations are useful for cataloguing and documentation, public outreach and education, historical studies, experimental architectural and urban history. The work described in this book creates a foundation for Virtual Heritage studies for Hampi.

Initiatives for documentation of intangible cultural heritage require an integrated approach for designing the digitization scheme for the target domain, formulating the mechanism for extraction of the latent human creativity hidden in them and to study the importance of the spatial features in the process of their evolution.

Traditionally, digitization of oral traditions is done by audio capture, accompanying images and video record details of the context. Data capture protocol typically speci es creating an image of colour depth of at least 24 bits (8 bit/colour RGB) (Minerva project [5]), the minimum resolution 4064 pixel × 2704 pixel (2-Megapixel) and the appropriate storage format is the TIFF format. When creating an audio le, the bit depth should be 24-bit stereo sound, the sampling rate 48 or 96 kHz and the appropriate storage format is uncompressed WAV format. When creating a video le, for the image, the sampling rate should be, at least, 25 frames/s and the colour depth 24 bit, for the sound 48 or 96 kHz and 24-bit depth. The digital video camera should be a 3-CCD camera, rather than 1-CCD, in order to provide better resolution. The appropriate storage format is RAW AVI format without the use of any compression scheme. The analysis scheme of Oral Traditions is focused on the content. Motif-Index of Folk Literature by Thompson [13] is an inclusive system of classi cation of every narrative genre. Motifs are basic elements that compose the narratives. Geographic information provides the spatial context of the tradition. GIS-based system can provide a picture about geographical or spatial evolution of the heritage.

Typically, dance performances are recorded in 2D video format. However, a 3D capture, using Kinnect or photogrammetric methods, would offer precise depth information for each dance movement. Motion capture systems record coordinates of various body points in time providing documentation about the nature of the movements of body parts. Dance analysis consists of movement analysis which attempts to describe, disseminate and interpret every possible movement [14]. We need to segment the sequence in terms of individual constituents in order to describe the motion. There have been several attempts to extract these features from motion capture data [15–17]. The analysis scheme is completed with other information related to the performance and its social role [18]: (i) location of performance and (ii) time of performance. In Mallik et al. [19], an ontology-based framework has been presented for digital documentation of Indian classical dance.

Standard way of digital preservation of music is through digital recording. In addition, notation is extremely important for capturing music. Notation allows recreation of the music as per the formal speci cations captured by the notation. Oral traditions do not have such mechanisms to formally record their compositions. However, different musical traditions have different notational forms. Documenting interpretation scheme for each notation remains a challenge because notes, intervals and scales differ from community to community. One such scheme is Cantocore [20]. It classi es songs according to their acoustic characters related to their structure, performance style and instrumental accompaniment. Another information important for documenting musical tradition is the occasion like religious, secular, march, lullaby, etc. associated with the corresponding musical performance.

Gen-Fang Chen has reported in [21] research ndings regarding digital preservation of Chinese opera called Kunqu. Kunqu opera is one of the oldest Chinese traditional operas and has a history of more than 600 years. It is a form of performance art that fuses literature, dance, music, martial arts, art and drama. Digital preservation methodology involved digitization of historical literature, segmentation of libretto musical score image, musical information recognition, musical score information representation, musical score information storage and libretto reconstruction on the Web.

4   The Book

This book is an outcome of the IDH (India Digital Heritage) project supported by the DST (Department of Science and Technology), government of India. The project was conceived with the idea of drawing together a pool of scientists and technologists working in computer vision, computer graphics, architecture and design together with art and craft historians, experts in cultural studies, to attempt and to work towards the documentation, representation and interpretation of the tangible and intangible heritages using digital approaches. The UNESCO World Heritage Site of Hampi in Karnataka, the capital of the Vijayanagara Empire from the mid-fourteenth to mid-sixteenth century, seemed to provide an ideal canvas for this novel experiment with its rich, diverse and dazzling legacy of architectural ruins, temples, art and living traditions. Technologies employed have ranged from laser scanning, modelling and gesture-based interactive displays, haptic explorations, experiential multimedia renditions and ontological explorations.

The well-known artistic and architectural splendours have provided a fertile ground for this exercise ranging from the Narasimha colossus, the luminescent murals of Lepakshi and the Virupaksha temple, the Vitthala temple complex and some of its enigmatic colonnades which emit musical tones, whether intentionally or not, and the remains of the bazaar, which was once the topic of much admiration by travellers from the Mediterranean. Also studied are the portable Vijayanagara bronzes used as "utsava murti" or processional images and portrait sculpture and related digital interfaces in the study of their iconometry and modelling. The legacy

of the "Girija Kalyana" ritual concerning the marriage of Siva and Parvati are also documented and interpreted from digital approaches.

This has been an attempt to explore the tangible and intangible heritages of Hampi from innovative perspectives which could have special relevance in the digital age. This book does not claim to generate a de nitive narrative frozen in time and space. It attempts to show how multidisciplinary engagement across technology and humanities can provide signi cant impetus to the cause of the documentation, showcasing and interpretation of cultural heritage and in some situations productive explorations of historical debates using digital tools.

References

1. UNESCO (2003) Convention for the safeguarding of intangible cultural heritage. In: 32nd session of the general conference Paris, 29 September–17 October 2003
2. Ikeuchi K, Oishi T, Takamatsu J, Sagawa R, Nakazawa A, Kurazume R, Nishino K, Kamakura M, Okamoto Y (2007) The great buddha project: digitally archiving, restoring, and analyzing cultural heritage objects. J Comput Vis 75(1):189–208
3. Gomes L, Bellon ORP, Silva L (2014) 3D reconstruction methods for digital preservation of cultural heritage: a survey. Pattern Recognit Lett 50:3–14
4. Levoy M, Pulli K, Curless B, Rusinkiewicz S, Koller D, Pereira L, Ginzton M, Anderson S, Davis J, Ginsberg J, Shade J, Fulk D (2000) The digital Michelangelo project: 3D scanning of large statues. In: Proceedings of the conference on computer graphics and interactive, techniques, pp 131–144
5. Fontana R, Greco M, Materazzi M, Pampaloni E, Pezzati L, Rocchini C, Scopigno R (2002) Three-dimensional modelling of statues: the minerva of arezzo. J Cult Herit 3(4):325–331
6. Bernardini F, Rushmeier H, Martin IM, Mittleman J, Taubin G (2002) Building a digital model of Michelangelo s Florentine Pieta. IEEE Comput Graph Appl 22(1):59–67
7. Allen PK, Troccoli A, Smith B, Stamos I, Murray S (2003) The Beauvais cathedral project, vol 1
8. Grün A, Remondino F, Zhang L, für Geodäsie I, Photogrammetrie (Zürich) (2002) Reconstruction of the Great Buddha of Bamiyan, Afghanistan, ETH, Eidgenössische Technische Hochschule Zürich, Institute of Geodesy and Photogrammetry
9. Rushmeier H (2006) Eternal Egypt: experiences and research directions. In: International workshop on recording, modeling and visualization of cultural heritage, pp 22–27
10. Sonnemann T, Sauerbier M, Remondino F, Schrotter G, Campana S, Forte M (2006) Reality based 3d modeling of the angkorian temples using aerial images. British Archaeol Rep Int Ser 1568:573–579
11. Lu M, Kamakura M, Zheng B, Takamatsu J, Nishino K, Ikeuchi K (2011) Clustering Bayon face towers using restored 3D shape models. Culture Comput 39–44
12. Koller D, Frischer B, Humphreys G (2009) Research challenges for digital archives of 3D cultural heritage models. ACM J Comput Cult Herit 2, 3, Article 7 (December 2009), 17 pp
13. Thompson S (1955) Motif-index of folk literature Central trykkeriet, Copenhagen
14. Laban R (1966) Choreutics. MacDonald and Evans, London
15. Kahol K, Tripathi P, Panchanathan S (2004) Automated gesture segmentation from dance sequence. In: 6th IEEE international conference: automatic face and gesture recognition, Seoul, Korea, pp 883–888, 17–19 May 2004
16. Camurri A, Hashimoto S, Suzuki K, Trocca R (1999) KANSEI analysis of dance performance. IEEE system, man and cybernetics, Tokyo, Japan, 12–15 October 1999

17. Zhao L, Badler N (2001) Synthesis and acquisition of laban movement analysis parameters for communicative gestures. PhD thesis, Department of Computer Information Science, University of Pennsylvania
18. Rett J, Dias J, Ahuactzin J (2010) Bayesian reasoning for laban movement analysis used in human machine interaction. Int J Reasoning-Based Intell Syst 2(1):13–35
19. Mallik A, Chaudhury S, Ghosh H (2011) Nrityakosha: preserving the intangible heritage of Indian classical dance. JOCCH 4(3):11:1–11:25
20. Savage P, Merritt E, Rzeszutek T, Brown S (2012) Cantocore: a new cross-cultural song classi cation scheme. Anal Approaches World Music 2(1)
21. Chen G-F (2014) Intangible cultural heritage preservation: an exploratory study of digitization of the historical literature of Chinese Kunqu opera librettos. ACM J Comput. Cult Herit 7, 1, Article 4 (March 2014), 16 pp

Introducing Hampi: Insights from Literary and Historical Sources

S. Settar

1 Introduction

Celebrated in history, rooted in myths and now a tumbled mass of magni cent residues of an empire, Hampi, in Karnataka is probably the most renowned medieval Hindu metropolis in the history of the country. As the capital city of the erstwhile Vijayanagara Empire, from the fourteenth to sixteenth century, it was unparalleled in its time as suggested by the accounts of many visitors. Early writings include those of Col. Mackenzie 1800 AD), who prepared the rst sketch map of this city, and de Warren E. 1845 AD), who gave the rst eyewitness account of the city. Writings such as the enduring and authentic descriptions of A. H. Longhurst 1917 AD) have in recent times piqued the interests of serious scholars. Much research has been done on the site since then, all of which have aimed to piece together the pillaged past of this splendid city with several scholarly monographs and archaeological reports. However, a new direction has been added with explorations using the latest digital tools and exploratory processes which aim to make the glories of Vijayanagara accessible to the wider public.

2 Origins Sources and Evolutions

2.1 Pampapura Hampi

'Hampi' is an ancient name, rst given to a *Saivakshetra* eld of the Lord Siva) on the banks of river Tungabhadra in Hospet Taluka of the Bellary District, Karnataka

S. Settar   )
Bangalore, India
e-mail: settar@nias.iisc.ernet.in; ssettar@yahoo.co.in

11

State. The land of Pampa, *Pampakshetra,* is also referred as *Virupakshapura* in the records found during seventh to fourteenth century, covering the two banks of the Tungabhadra, which currently include Hampi and Anegondi. According to myth, Pampa, the daughter of Brahma, morti ed herself here to gain the hand of the Lord Siva as Virupaksha), which is the subject of the celebratory work *Girija Kalyana* by the late twelfth-century poet, Harihara. The *Tiruvannamalai Sthalapurana* states that Pampa, having provoked Siva to open his third eye, had to undergo penance at this place to exonerate herself from his curse. The *Skanda Purana* makes *Pampakshetra,* a *saktipitha,* seat of Goddess Parvati) and hence, an abode of the goddess Pampa. Yet, another legend talks about how the terrifying, dark-coloured goddess Kali gained a fair complexion by the grace of Pampa and settled here as *Saumya-Pampa,* shedding her fearsome form.

The *Pampasarovara*, where Pampa is believed to have performed penance, has been identi ed in Anegondi as a spot on the northern bank of the Tungabhadra between Kodandarama and Vijaya Vitthala temples), as per a record dated to the 1400 AD. The twelfth-century records generally identify it as *Virupaksha tritha.* The word *Hampe* also seems to have been in popular vogue during this period. The Kannada poet, Harihara, refers to *Pampakshetra, Pampapura* as well as *Hampe* and points to Virupaksha as its ruler arasa).* He continues to refer to his god, teacher and father as *Hampeya-Virupaksha, Hampeya-Sankaradeva* and *Hampeya-Mayideva.* The poet Kereya Padmaras refers to his predecessor, poet Harihara as *Hampeya-Ramanna.*

What was known as *Pampa* during the seventh century developed into the *Pampatirtha* and *Virupaksha tirtha* subsequently and further expanded into a town named *Pampapura* by about the eleventh to twelfth century. It came to be popularly known as *Hampi* by the beginning of the thirteenth century. As per the records of the late thirteenth and early fourteenth century, popular names such as *Pampapura* and *Hampi* had become somewhat obsolete and were replaced by new ones. These indicate the rise of a new township or extension of an existing settlement under new names. The Hoysala records of the late thirteenth- and early fourteenth-century mention names such as *Hosapattana, Basapattana, Hosa-Hampeyapauana, Virupakshapattana, Virupaksha-Hosadurga* and *Vijaya-Virupakshapattana*. Interestingly, Vira Virupaksha Ballala, the son of Hoysala Ballala III, the last ruler of the Hoysala Dynasty, seems to have been associated with this town, as a devotee of the Lord Virupaksha.

2.2 Capital of the New Kingdom

After the establishment of the Vijayanagara kingdom in the middle of fourteenth century, the *Saivakshetra* of *Pampa* began to grow into a large urban complex. The founders of this kingdom initially took suf cient caution to preserve the 'kshetra' character of Hampi. Hence, they situated their rst capital, Anegondi, which was its

twin city on the north bank of Tungabhadra, at a distance of about 5 km from *Virupakshakshetra.* Subsequently, when the imperial capital was shifted to the southern side of Tungabhadra, it was located at a respectable distance from the age-old Virupaksha complex.

The new capital across the river was rst called *Abhinava–Vijayanagara,* and from around 1368 AD only the term Vijayanagara continued. Variations such as *Vijeyanagara, Vijayanagari, Vijayanaagara, Vijanagar,* besides *Hastinapura-Vijayanagara, Vijayanagara-Hastinavat I* and *Vijayanagari* are found in the records in which Vijayanagara remained a standard name. The fteenth-century Virasaiva poet-minister Lakkanadandesa suggested that the capital city be renamed as *Vijayakalyana* under the reign of Sangama Devaraya II. In the sixteenth century, the Advaitins or followers of the teachings of the saint Sankaracharya from the monastery of Sringeri proposed the name *Vidyanagara*. However, these proposed names were con ned to the records and literature and never attained regular usage. In fact the local poets, Lakkanadandesa and Chandrasekhara, who wrote about Hampi in the fteenth century were unaware of the name Vidyanagara. Nearly, all the visitors referred to the empire's imperial city only as Vijayanagara. Subjected to the quirks of their local tongues, European visitors pronounced 'Vijayanagara' in a dozen different ways. Nearly, all the visitors to the empire seem to have known the imperial city, only as Vijayanagara. It was known as *Beejanugger* or *Beejnugger* to historian Ferishta from Persia, while the early Portuguese travellers named it *Bisnagar, Bisnaga, Bidjanagara* and *Bijanagher.* The well-travelled Portuguese visitors such as Domingo Paes and Duarte Barbosa referred to it as the kingdom of *Narsymga* or *Narasyngua*, and also mention Saluva Narasimha, the ruler of the time. Afanasy Nikitin, a Russian traveller, referred to this Hindu kingdom as *Chenudar, Benudar* and *Bichenegher.* The Italian Nicolo de Conti, the Italian, referred to it *Biumegalia.* To Cauto, it was *Visaja Nagar,* while he had also heard that it was locally known as the 'Kingdom of Canara' or 'Kingdom of Karnataka'. Only Fernao Nuniz appears to have come across the name *Vidyanagara,* for he had spelt it *Vydiajuna.* However, he also testi es that the more popular name of this city was *Bisnaga.*

3 The Vijayanagara Empire

3.1 Birth of a Kingdom

Contrary to a widespread myth, the vast Vijayanagara Empire was not founded by a stroke of luck or the whim of a spiritual guide. Before the rise of Vijayanagara, the waves of invasions by Malik Kafur 1308–1310 AD) and Khusru Khan 1318 AD) of the Khalji Dynasty, and Ulug Khan and Malik Zada 1323–1327 AD) of the Tughluk Dynasty made south India politically weak and fragmented. The fall of the foremost of the Hindu dynasties, the Hoysalas in 1342 AD, left South India

directionless and vulnerable. It was at this time that the efforts of the ve sons of a feudatory chieftain, named Sangama, metamorphosed into a kingdom, even to their own surprise. Harihara I, the founder–ruler, rose from the status of a *mahamandalesvara* to that of a monarch. The Vijayanagara kings seemed to have ruled their "kingdom of Karnataka" without a xed capital for a couple of decades. None of their records issued before 1357 AD identi es their capital city as Vijayanagara. Recent researches have indicated that their capital city was established sometime between 1357 AD and 1368 AD and that their continuous rule from the city of Vijayanagara may be traced only from 1368 AD.

3.2   Rulers of Vijayanagara

Established sometime between 1336 and 1346 AD, the kingdom of Vijayanagara lasted for about four centuries 1336–1726 AD). During this period, twenty-eight monarchs belonging to four feudatory dynasties ruled the empire. The founding of the empire is attributed to Harihara and his four brothers of the Sangama Dynasty, who ruled from 1336 to 1485 AD. It was under this dynasty that their political in uence spread rapidly, with intermittent clashes with the Bahamani sultans. In fact, fourteen Sangama kings and thirteen Bahamani Sultans fought eleven great battles during a period of one hundred and thirty years. The Bahamanis attacked the Vijayanagara capital six times but were never able to penetrate into its impregnable fortress. The Vijayanagara rulers, despite these disruptions, ourished in southern India. Under Devaraya II 1425–1446 AD), they controlled the region from Gulbarga to Ceylon and from Orissa to Calicut, bounded by the four oceans Fig. 1). It was also during this period, especially under Devaraya II, that the Muslim in uence came into vogue in the Empire, and that the Islamic practices gradually became an integral part of culture and architecture.

The Saluvas came to power when the provincial governor, Saluva Narasimha, overthrew the weak Sangama ruler Virupaksha III in 1486 AD. Ironically, his rule lasted only for 20 years, as they met with the same fate at the hands of their subordinate of cial of the Tuluva family in 1505 AD. The Tuluvas ruled for 70 years from 1506 to 1576 AD, and produced the greatest Vijayanagara king Krishnadevaraya I509–1576 AD), who was in the words of the traveller Domingo Paes, 'the most feared and perfect king that could possibly be'. The empire reached its zenith and also experienced ignominy under this dynasty. The glorious empire virtually ebbed away after the decisive Battle of Talikota in 1565, in which the victorious army of the Muslim Deccan Sultanates razed the metropolis leaving it a vast heap of rubble, blood and destruction. The Aravidu Dynasty ruled from 1570 to 1726 AD and kept the empire alive long after the fall of imperial capital, originally from Penugonda, and later from Chandragutti and several other centres.

Fig. 1 Map of the Vijayanagara Metropolis

4 Hampi-Vijayanagara The Capital and Around

4.1 *Puras and Suburbs*

Around the fourteenth century, Pampa or Hampi and the *Saiuakshetra* were integrated into the metropolis of Vijayanagara and formed a large urban metropolis. The new name of the capital city began to gain in popularity. Three distinct centres evolved within the larger Vijayanagara complex, forming the metropolis of about 30 km^2. Forming the sacred nucleus was Hampi, a *Saiuakshetra*, built around the Virupaksha temple, covering the entire area between the Tungabhadra and Hemakuta hillock. The mother capital, Anegondi, gradually developed into an important urban centre on the northern bank of the river. Vijayanagara, the imperial city, was built beyond the hillocks of Hemakuta. Townships or *puras'* along and beyond the river bank constituted part of a vast metropolis extending about 6 km

from Virupaksha to the Penugonda Gate and from Kamalapura tank to Talarighat). These were established through the initiative of the imperial authority in such a way that they lay contiguous to each other, demarcated by intervening hillocks, canals, river, tanks, orchards and ramparts. The suburban townships, with the names ending with '*gudi*' or '*pattana*' were located about 6 km southwest, on the Kamalapura–Hospet road. Interestingly, the visionary town planners of Vijayana-gara ensured the creation of a buffer between the imperial city and the suburban townships that was nurtured by agriculturists, horticulturists and mercantile groups. These townships and suburbs in and around the metropolis were established by royalty and named after the family deity, ruler or his family.

4.2   Virupakshapura

It is the most ancient and sacred Saiva pilgrimage along the Tungabhadra that evolved into a lavishly planned township. It comprised the Virupaksha temple and its Chariot Street Fig. 2) that extended up to the Matanga hill, a cluster of about forty temples on Hemakuta, south of the Virupaksha temple and a cluster of about sixteen monuments between the Virupaksha temple south) and the Tungabhadra river north).

The Virupaksha temple has a recorded history of about a thousand years. The Vijayanagara kings rebuilt the present temple of Virupaksha, in the early fteenth

Fig. 2 Hampi Bazaar street leading to the Virupaksha temple

century. Known by the name *Pampatirtha* 689 AD), it had evolved into *Pampapura* by 1018 AD, whilst still retaining its *tirtha status.* Between the seventh thirteenth century, it became revered by the names of *Pampa, Pampatirtha, Pampapura, Swami Pampasthala, Virupaksha tirtha* and *Hampi.*

   The twelfth-century poet Harihara, a Hoysala of cer from Hampi, hailed the area of the Hemakuta hill as the crowning glory of Pampapura. According to him, 'Sivapura' was a ourishing settlement of Sivasaranas attached to the Virupaksha temple. He further stated that along and beyond the two banks of the Tungabhadra were twenty-one temples, eleven *tirthas* and dozens of sacred spots. By this time, it was also known as *Hemakuta-pranta,* or province of Hemakuta. A record of 1119 AD attests that an almshouse chhatra) was established near the Virupaksha temple, which not only fed the touring Brahmins apurvis) but also seven hundred local residents. Another record of 1199 AD identi es a Nayaka ruler of this area and lists a cluster of ve temples in this area. The sixteenth-century Brahmin poet of the *Sri Virupaksha Vasantotsava Champu* also recalls that there was a settlement of Saivas there. This was under the rule of Sangama Devaraya II 1422–46 AD) who built the huge complex for the temple. The poet-minister, Lakkanadandesa, gave explicit details of the contributions made by Devaraya II. Domingos Paes 1520–22 AD) noticed that Virupaksha temple was the one 'which they hold in most veneration and to which they make great pilgrimages'. The main *gopura* or temple tower is called the *hiriyagopura* or the chief tower. Constructed by Devaraya's minister Proluganti Tippa, and later repaired by Tuluva Krishnadevaraya in 1510 AD, it was the largest of the *gopuras* raised by the Vijayanagara kings. The hall of the main temple is believed to have built under the patronage of Saluvamantri, a minister of Sangama Mallikarjuna 1447–1465 AD). Krishnadevaraya is also credited with the construction of the three-storeyed *gopura* and the open-hall *maha-rangamantapa* in the main temple premise, in the same year that he made the repairs to the main gopura.

   Lakkanadandesa mentions that the Chariot Street of the Virupaksha temple known today as the Virupaksha Bazaar Street) as well as the lofty chariots made for gods Virupaksha, Pampa and Ganesa were made by Devaraya II. Situated between the eastern gate of the Virupaksha and the northwestern foot of the Matanga hill, it was the second longest at 0.8 km. and one of the widest streets at 13.9 m. Paes described it as a very beautiful street with beautiful houses with balconies and arcades, sheltering pilgrims that come to it, and with houses for the upper classes. He added that the king had a palace in the same street in which he resided. Nicolo Conti, the rst European visitor to Vijayanagara 1420–1421 AD), refers to two chariots which carried idols through the city. Richly adorned young women or courtesans sang hymns and accompanied the procession. He noticed peculiar customs such as of people throwing themselves under the wheels of the moving chariots out of devotion. Poet Ahobala, the author of *Vasantotsava Champu,* also refers to the two chariots: one taken out by the Brahmins and the other by the merchants and shudras. Interestingly, the Virupaksha chariot festival

has been continued ever since it was introduced in the fourteenth century and neither the fall of the empire nor the destruction of the capital in 1565 AD seems to have affected its popularity or practice. To date, the largest gathering at Hampi is witnessed during the chariot festival of Virupaksha and Chandra Maulisvara, held every year in April.

4.3   Achyutapura

The centre of Achyutapura is the temple of Tiruvengalanatha, also called Venkatesa. It was built by Tirumalaraya Vodeyar Sr.), a nephew of the ruling king, Tuluva Achyutaraya and a brother of queen Varadambika, in 1534 AD. It is a typical late Vijayanagara structure. Locationally, tucked away in a valley, with the magni cent Matanga Hill forming a side-drop and a formidable granitic outcrop anking the main market, the temple is spectacularly isometric in perspective. The highlight of the temple is the grandeur of the ruined *maha-mantapa* as testi ed by its ornate pillars, which exhibit exceptional craftsmanship.

Flanking the temple is the market gallery, which was called *Achyutapete*, while the whole complex including the Matanga Hill was known as *Achyutapura*. The *pete* is 0.5 km long and 39.4 m wide and is lined with galleries. Broader than Virupakshapura's bazaar street, it must have been a vibrant chariot street. Strangely in the course of time, it developed a dubious reputation as *Sule-bazaar* Courtesans' Street). Interestingly, this is the only *pura* of Hampi that is laid on the south–north axis and therefore overlooks the Tungabhadra.

Matanga hill offers an unparalleled, magni cent view of its surroundings, particularly the scenery around the Virupaksha, so much so that Longhurst remarked, 'I do not think there is a more interesting or beautiful view than this, in the whole of South India'. Replete with myth and history, Matanga hill is one of the four scared mountains as referred to in medieval literature. The hill is named after the *muni* Matanga, who according to the lore of the Ramayana epic put Vali the monkey chieftain) permanently out of the bounds of this hill by spelling out a curse. Poet Chandrasekhara refers to the *tirthas, Vranamochana* and *Yakshamukha*, which surrounded this hill. In Harihara's *Girijakalyana,* the god of love choses this vantage point to disturb the concentration of meditating Siva and got burnt to ashes by his third eye. Historically as well, this hill has played an important role, both as a secular and as a religious centre. According to records, in 1199 AD, Madeyanayaka, a governor under the Sindas of Kurugodu, was headquartered here to 'protect the Virupaksha tirtha'. This hill also features the earliest temple of the Vijayanagara period, which was initially dedicated to god Sambhu by Madhavamantri in 1346 AD but later was established as a temple of Virabhadra or Viresa. In words of Lakkanadandesa   fteenth century), this great mountain symbolized the 'height of the glory of Pampavara'. The records of that period hail Virabhadra as Matangesvara.

4.4 Krishnapura

Krishnapura is the earliest Vaishnava township established at Hampi during the reign of Tuluva Krishnadevaraya. Located at a counterbalance to the Virupaksha complex of the Saivas, this *pura* was bounded by the Raghunatha temple of the *hiriya kaluve* or 'major canal' to the west at Kadiramapura Road) and the Virabhadra temple located at the extreme end of the valley to the east south of the Matanga hill) spanning an area of 1.2 km from west to east. It covered a distance of over 0.5 km distance, between Hemakuta and the 'Uddana-Virabhadra' temple from north to south. Krishnapura marked the celebration of the king's victory against the Adil Shah is of Bijapura and the Qutb Shahis of Golkonda, in general, and the Gajapatis of Orissa.

The Krishna temple was established between 1513 and 1515 AD, as part of Krishnadevaraya's visionary plan to convert his capital into as great a *kshetra.* To achieve this end, one of the ways was to capture one the cult deities from the areas he had vanquished during the military campaigns and to instal it in his capital. According to a famous record, on Friday, 16 February 1515, Krishnadevaraya installed an image of Krishna, which he had 'skillfully carried' from a temple at Udayagiri. It further states that image was installed in the *mani mantapa* or *jewelled pavilion,* located in the northeast corner of the courtyard. This temple has a large and ornate Eastern gateway, while its other mandapas have exquisite stuccowork Fig. 3), representing one of the nest structures by the Vijayanagara architects.

Fig. 3 Stucco work on the mantapa in Krishna temple

Flourishing between 1513 and 1533 AD, this Vaishnava township had a Brahmanical settlement under the name of *Madinelala Hunise Agrahara* or *Pratapadevarayapura,* located at *hiriya kaluve* West). The market was called *Krishnapurapete* and marked its southern limit of the *pura*. Visitors in the sixteenth century mention the presence of a seven hundred-pillared almshouse, behind the Tirumala temple. Further sources state that in 1533 AD, the market dealt with food grains *davasad-angadigal)*, and the shopkeepers' resided nearby. In the sixteenth century, the road between Krishnapura and the metropolis was referred as *Deveri-Vithi* or 'Saviour's Street', as it was dotted with temples and shrines of Muduviranna, Ganesa, Mallikarjuna and Prasanna Virupaksha.

The Chariot Street of Krishna temple is the most impressive part of Krishnapura. Placed in front of the temple East), 48 m in width and 0.6 km in length, it has cluster of boulders, which feature bold reliefs at the extreme end.

4.5 Vittalapura

The Vijaya Vitthala temple is a stupendous creation of the Vijayanagara artists and with few parallels in the architectural history of medieval India. The most outstanding components of the Vijaya Vitthala are the eastern pavilion of the main temple, originally called *Dolotsava Mantapa* and now popularly known as the 'Hall of Musical Pillars', the Stone Car, and the two open pavilions, the *Kalyana Mantapa* and *Bhoga Mantapa* which face the temple in the east. The original foundations of *Dolotsava Mantapa* were laid sometime during the reign of two Devarayas 1406–1446 AD) of the Sangama Dynasty. It was thoroughly renovated during the time of Tuluva Krishnadevaraya 1503–1528 AD). It experienced further expansion through grants from the two successors of Krishnadevaraya, i.e. Achyutadevaraya 1529–1546 AD) and Sadasivaraya 1542–1565 AD). The Stone Car Fig. 4) that faces the *Dolotsava Mantapa* has a shrine of *Garuda*, the vehicle of Vishnu Vitthala), which is another architectural marvel of this temple complex.

Interestingly, this Stone Car is not a monolithic monument and has components built of dressed and designed stones akin to that of any Dravidian temple. Inscriptions state that the entrance pavilions with *gopuras* were built in 1513 AD by the two chief queens of Krishnadevaraya. According to *Narasimhapurana,* an eighteenth-century Telugu work, Proluganti Tippa, is said to have built the *gopuras* during the reign of Devaraya II of the Sangama Dynasty. The 'Hundred-Pillared-Hall' has altogether hundred and eight pillars and pilasters and was built in 1516 AD by Krisnnadevaraya. The Kalyana Mantapa seems to have been constructed in about 1554 AD. This temple is unique, and unlike other temple complexes, it was repeatedly renovated until the capital was destroyed in 1565 AD.

Fig. 4 Stone Chariot of Vijaya Vitthala temple *Photo credit* Sharada Srinivasan)

The records of sixteenth century refer to this complex as 'Vitthala', rather than as 'Vitthalapura'. The temple complex extended over a distance of about a kilometre, with the main temple located within a high-walled enclosure. This temple was called the Vijaya Vitthala in most records and as *Kanada Vitthala* in one of the records. It is assumed that it was called *Vijaya* i.e. Vijaya-Virupaksha), perhaps to celebrate victory.

To the north, west and east of the Vijaya Vitthala temple were rows of galleries, of which only few survive now. The most impressive of the galleries faced the main gopura of the Vijaya Vitthala temple. The eastern bazaar of the Chariot Street of the Vijava Vitthala is about 40 m wide and a kilometre long. The boat festivals of the god and goddesses were held in the *teppotsava pond,* with its own gallery and gateway. The galleries served as shops, residential quarters, rest houses and camping centres for pilgrims.

4.6   The Suburban Townships

The suburban townships, with the names ending with *gudi'* or *pattana',* were located about 6 km to the southwest, on the Kamalapura–Hospet road. Interestingly, the town planners of Vijayanagara had the vision to create agricultural and commercial spaces between the imperial city and the suburban township. These townships and suburbs in and around the metropolis were established by the royalty and named after the family deity, ruler or his family.

   The suburb Malapanagudi is located about 6 km southwest of Kamalapura. As early as in 1412 AD, a rest house and a large well were established here according to one of the records, perhaps to serve the needs of travellers. In the middle of that century, the suburb of Malapanagudi evolved with a temple for god Mallayya or Mailara or Mallikarjuna. An octagonal well Fig. 5) on the outskirts of this village, called *Sulebavi* Courtesans' Well), is an early example of assimilation of Islamic style in secular structure.

Fig. 5 Sulebavi, the octagonal well exhibiting early Islamic in uences, Malpanagudi

Anantasyayanagudi is about 4 km to the southwest of Malapanagudi. Originally, it was called Sale Tirumalerayapattana, and its establishment is attributed to Krishnadevaraya in 1524 AD, in honour of his newborn son, Tirumaleraya. To the west of Anantasyayanagudi is the large town of Hospet. Now, these two have merged into a single entity.

Nagalapura and Kadiramapura are other suburbs of Vijayanagara era. The last of the suburbs, covering the extreme west end, was perhaps located near the present Tungabhadra Dam. In the sixteenth century, it was called Tirumaladevi pattana city of Tirumaladevi). Nearly, all the *puras* and *pattanas* extend from west to east, with exception of Achyutapura, Anantasyayanagudi and Malapanagudi.

5   Hampi-Vijayanagara   Capital City

5.1   *Literary References of the Metropolis*

The splendours of the Vijayanagara metropolis have swayed the imagination of chroniclers, historians, poets, researchers and archaeologists for many centuries. The Portuguese travellers who compared it with the best cities on their continent found even Milan and Rome only partially comparable to this city. They found it to be a second paradise. The Indian poets of the time found it equal or more than equal to the great cities of their gods, such as Indra's Amaravati, Siva's Alaka Gangadevi, mid-fourteenth century) or Nagendrapuri Chandrasekhara, early   fteenth century). The growth of the metropolis can be traced in three distinct stages, beginning from the reign of Bukka I, when the foundations of this capital were laid and the core area was build up. The core area culminated into a fully developed city under Devaraya II 1420–1440 AD) and   nally, it was the Tuluva's who made major changes to the city.

5.2   *Antiquity of the Site*

Interestingly, the site on which the metropolis was established had a far greater antiquity than that of Virupakshapura. A Prakrit record, found in the Darbar Enclosure, takes its history back to the second century AD. There was a thriving Saiva *matha* near the present ruins of the Mahanavami dibba, as early as the eleventh century. The residents of this *matha* were learned men imparting lessons on the *Puranas.* It was also a sacred centre for the Jains as one of the Jain saints had found it a suitable spot for terminating his life in the same century. This area is also linked with the episodes of the epic *Ramayana,* according to which the Malyavanta hill had allowed Rama and Lakshamana to spend a rainy season on it. The Madhuvana, where the monkeys are said to have celebrated the discovery of Sita by

Hanuman, lay a little distance away from this centre. In fact, the rst great temples for Rama Hazara Rama) and, perhaps for Hanuman, were constructed here in the centre of the site, which later evolved into a great metropolis.

5.3 Myths and Facts About the Foundation

There are several myths connected with the foundation of the metropolis and most of them revolve around the kings who founded it, the choice of the site, and the role of a spiritual guide. Nearly, all the myths credit the foundation to one of the Sangama brothers, Harihara I, who was helped and guided by the saint Vidyaranya, an Advaitin follower of Sankaracharya) of the Sringeri matha. Some also say that Vidyaranya, under some divine inspiration, established the capital and subsequently handed it over to a *kuruba* devotee. The *kuruba* was none other than Bukka. Bukka built a city named 'Visaja Nagar' or 'Bisnaga' which became the metropolis of the 'kingdom of Canara'. Legend goes that Harihara saw a hare challenging and injuring well-bred hounds that were pursuing it. When Harihara recounted this uncanny incident to a saint wandering on the bank of the Tungabhadra, the saint prophesied that the land could one day become an imperial city. In yet another version recorded by Keladi Basavarya in his *Sivatatvaratnakara,* the Sangama brothers as well as Vidyaranya were prompted in their dream by Lord Siva himself to found this metropolis. When they were counselling with each other, a shepherd appeared and informed them about the goddess of wealth, who lived on the hillock of Hemakuta and who could bless them with wealth to establish the city. In this myth, the *kuruba* shepherd) exacts a promise from the Sangamas to the effect that the descendants of the Sangama family would name themselves after him, in return for his being sacri ced to the goddess. The goddess blessed the future founders of the metropolis by causing a rain of gold for one and half hours over an area drawn up in the form of a *Srichakra.* This was the site on which the new capital was established.

However, recent researches have proved beyond doubt that Vidyaranya did not live anywhere near the metropolis during the time of Harihara I, and that the great city emerged only after the death of the rst Vijayanagara king. Nonetheless, the metropolis seems to have been laid on a site, roughly resembling the *Srichakra,* for even Abdur Razzak and Caeser Frederick stated that it was circular in shape.

5.4 Sancti cation of the Site

The metropolis was developed primarily to meet the demands of the state, and despite the distance from the Virupaksha complex, the founders consciously exerted themselves to make it an extension of the old Virupaksha complex. The metropolis was never an exclusive secular centre and royals made efforts to draw the parallels

with the cult of Virupaksha. They sancti ed it by bringing Lord Virupaksha into it. Geographically, the western end of the royal centre was demarcated by the Prasanna Virupaksha temple. The eastern end was demarcated by the Malyavanta Hill, in such a way that they lie in almost the same axis. The sancti cation process was intensi ed with the major landmarks in the new city being given old names. The *Matangadevara kottala* bastion), *Hampadeviya diddi* and *Jadeyasankaradevara diddi* gateways) were located on either side of the Saviour's Street a trunk road between the Vijayanagara metropolis and Virupakshapura). The *Hampeya kottala* and *Sankaradevara kottala* were names given to the watchtowers, located north of the Zenana. A record of 1378 AD states that Bukka, who built this splendid city of victory, ruled with 'the Tungabhadra as his foot-stool and Hemakuta as his throne', and seated 'like Virupaksha for the protection of the people of the earth'.

5.5   *Emergence of the Metropolis*

The *Madhuravijayam,* written by poetess Gangadevi, daughter-in-law of Bukka I who might have lived in the growing city, has many references and descriptions of the metropolis. She refers to it as *Vijaya* and *Vijayapura,* stating that it was 'beloved to the heart' of Bukka. She adds that Bukka was as 'pleasing as the full moon to the eyes of the people of Karnataka' 1,75). Between 1357 and 1368 AD, the city seems to have developed into a metropolis. The imperial title assumed by Bukka in 1368 connotes the mounting of 'the great jewelled-throne of lions' *maharatna-simhasana)* in the 'new city of Vijaya' *abhinava–Vijayanagara).* An inscription near the Elephant Stables states that 'this metropolis of Vijayanagara is of Sri Vira-Bukkaraya'. Its main gateway was the *Sringarada hebbagilu,* with the 'Bastion of Anegondi' to its north along the fort wall Fig. 6).

5.6   *The Mahanavami Dibba*

Vijayanagara was a military state and the monarchs of this kingdom took great pride in parading their military prowess at annual ceremonies such as *Dasara* commemorating the victory of Rama over demon Ravana. The massive podium, or the Mahanavami Dibba Fig. 7) located in the northeastern section of the Darbar enclosure, is the most ornate secular monument surviving in the metropolis. The king, his court and his *entourage* must have watched from here the ceremonial parades and festival sports, during the *Navaratri* Nine Day) festival. Paes describes a 'House of Victory', constructed by Krishnadevaraya after his victorious return from Orissa in 1513 AD. Foreign visitors to the metropolis give a vivid account of the round-the-clock ceremonies of this festival—the rituals, homages, sports and festivities.

Fig. 6 Fort wall inside the Royal Enclosure

5.7   Palaces and Palace Life

According to poet Chandrasekhara 1430 AD), the Vijayanagara capital was a city
of numerous forts, gateways, platforms and archwavs and that its mansions, with
neatly laid oors, were built of pale-red walls, adorned with cornices, niches,
ventilators vatavanas)* and *salabhanjikas* 40-v). Paes, who compared the palaces
to castles of Lisbon, noticed reliefs of beasts, acrobats and dancing girls, which are
still to be found in the carvings of the Mahanavami Dibba. Abdur Razzak in 1443
AD had noticed impressive structures such as the King's Audience Hall, the 'Palace
of Dannayaka' and the 'State Mint' and the 'Dewan Khaneh'. Queen Gangadevi,
the author of *Madhuravijayam,* wrote that the king Bukka I) loved to be lost in the
inner apartments of the palaces, which were lled with sweet-scented *aragu* fumes
and that he sported with his captivating consorts whose warm breasts were painted
with saffron paste V. 55).

The conventional Vijayanagara palaces followed xed patterns with stone base-
ments, but of timber and mud superstructures Fig. 8). They invariably faced either
east or north. They all rose in ascending order of two or more oors, each ascending
 oor being connected by a single or double stepway, anked by balustrades. The oor
was invariably plastered. The wooden pillars bore roofs which were decorated with
ivory, copper plates and even precious stones.

Fig. 7 Mahanavami Dibba in the Royal Enclosure *Photo credit* Sharada Srinivasan)

5.8 The Darbar Enclosure

The largest, densest and the most important of the enclosures in the metropolitan area is found to the west of 'the Dannayaka and the Mint' enclosures. It is also called the 'Darbar Enclosure', because the most eventful ceremonies of the state took place

Fig. 8 Aerial view of the Zenana Enclosure

here. The main entrance to this enclosure is on its northwest, making access from the Hazara Rama temple easy. The pomp and splendour of the Vijayanagara rulers, their civic sense, hydraulic skills and even their cosmopolitan lifestyle is best exempli ed by the Darbar enclosure. No other Hindu kingdom known to ancient and medieval history built its capital with as elaborate a network of water works as that of the Vijayanagaras, and nowhere are their hydraulic skill and civic sense mirrored better than in the Darbar enclosure. The Vijayanagara records refer to innumerable hydraulic schemes and hydraulic engineers. Krishnadevaraya is stated to have even invited Portuguese engineers to build dams and tanks.

6 The End of the Glorious Empire

6.1 After the Battle of Talikota

When the head of Ramaraya fell on the battle eld, a state of total anarchy ensued in 1565 AD. This confusion reigned around the ruined capital of Vijayanagara for about two centuries. The colossal empire collapsed, and one of the greatest of medieval metropolises was turned into a mound of ruins. Subjected to arson, loot, rape and dishonour, the proud citizens of Vijayanagara suffered untold misery. The rape of the undefended capital lasted for 6 months, till almost all its temples, palaces and thriving streets were desecrated and the wealth of the capital was fully drained out.

   The devastation was so complete that an Italian visitor, Caesaro Federici, could nd only wild animals in the ruined dwellings where only 2 years ago the most prosperous families had led an enviable life. Two of Ramaraya's brothers, Venkatadri and Tirumala, collected the royal treasure and together with the heir to the Tuluva throne hastened to Penukonda. With this, the Tuluva rule came to an end, and that of the Aravidu began 1570 AD). However, peace and prosperity continued to elude the usurpers. Tirumala had to strive for 6 years in order to stabilize himself, and this he could do by suppressing several Hindu feudatories. In fact, he could gain peace only after compromising with the Nizam Shahis of Golkonda and with the Nayaks of Madurai, Thanjavur and Gingee. After 1590, the Aravidu rulers ruled from Penukonda and also from Chandragiri, Srirangapattana and Vellore. Sriranga III 1642–1664 AD), the last of the effective rulers, had to take shelter in Mysore before getting lost in the woods at Ikkeri 1660 AD).

6.2 The Forgotten Empire

After 1565, the city of Hampi became a heap of dressed-stones, with dust, dirt, wild animals and thorny shrubs dominating rows of natural hillocks. The vast empire

was divided and shared by an ambitious group of Palegar, the Marathas, the Rajas of Tanjore and Gingee. Soon, they were ousted by the Moghuls. The Marathas and the Mysore Sultans Hyder and Tippu) occupied this territory. After the fall of Tippu in 1799, Harapanahalli came under the Nizam of Hyderabad. In 1800 AD, Hampi formed part of the ceded districts, whereby Bellary, Anantapura, Cuddapah and a part of Karnool District were ceded by the Nizam to the British. These ceded districts formed a political or administrative unit under the name *Rayalaseema* as early as the sixteenth to seventeenth century.

Sir Thomas Munro, who became the Principal Collector, had to face the Palegars, especially those of Rayadurga and Harapanahalli, who were indulging in reckless feudal sport, promoting only turbulence and lawlessness. Munro restored order, and for the rst time, after the fall of Vijayanagara in 1565 AD, the free-booters were brought back into the fold of law. With the exception of the Pindari invasion of Harapanahalli in 1816 AD, and the Revolt of Mundargi Bhima Rao in 1858 AD, the neighbourhood of Hampi remained in comparative peace. However, the ruined capital of the Vijayanagara, which had slipped away from the memory of the mass of people along with the 'Forgotten Empire'', lingered on only in the minds of pious pilgrims to Hampi, the original visitors of *Pampakshetra* and *Virupakshapura.*

Further Readings

1. Settar S 1990) Hampi, a medieval metropolis. Kala Yatra
2. Sewell R 1900) A forgotten empire—Vijayanagar. Asian Educational Services, New Delhi
3. Longhurst AH 2006) Hampi Ruins. Asian Educational Services, New Delhi original printed 1917 by Government Press, Madras)

Introducing Hampi: Landscapes and Legends

George Michell

1 Landscape Local Legends Hydraulics Natural Protection

Of all aspects of Hampi, it is perhaps its natural setting that most astonishes visitors. The landscape in which the ruins of Vijayanagara are scattered is a wilderness of granite ridges and boulders that gives the impression of violent, cataclysmic upheaval (Figs. 1 and 2). But this dramatic, rocky environment is actually the result of erosion, with countless millions of years of time and weather transforming the primaeval lava flows into the configurations seen all around the site. Running through this granitic wilderness is the Tungabhadra, a river fed by the monsoonal rains that fall in the ranges of the Western Ghats some distance away. The Tungabhadra gorge is of incomparable beauty, changing colours during the di erent seasons and times of the day. Along its waters, boatmen still direct the circular coracles that are the river's traditional crafts.

Hampi's landscape is charged with mythological associations that give meaning to the site, helping to explain why Vijayanagara came to be located here. To begin with, Hampi is an ancient *tirtha*, a holy spot located on the southern bank of the Tungabhadra, marked to this day by a small village with houses clustering around the Virupaksha temple. From ancient times Pampa, a river goddess, who has given her name to the village of Hampi, was venerated. Described in local legend as the daughter of the Hindu god Brahma, Pampa diligently performed penances on Hemakuta hill that rises above Hampi, thereby attracting the attention of the god Shiva who was seated in meditation nearby. As a result, Shiva became betrothed to Pampa and married her, whereupon she became his consort Parvati, and he became Pampapati, Pampa's Lord. The marriage of Pampa to Shiva, who is also known at

G. Michell ( )
Faculty of Architecture, Building and Planning, The University of Melbourne,
Melbourne, Australia
e-mail: georgemichell@aol.com

Fig. 1 Hampi landscape: scatterings of granite ridges and boulders

Fig. 2 Hampi landscape: the Tungabhadra gorge

Hampi as Virupaksha, He with Oblique Eyes, was the most important religious occasion at Hampi in the past and continues to be celebrated to this day. The annual chariot festival that takes place in the street in front of the Virupaksha temple is attended by huge crowds of devotees. The celestial union of Pampa with Viru-paksha is the central scene of the vividly toned painted composition on the ceiling of the pillared mandapa within the Virupaksha shrine, so expertly documented and technically analysed elsewhere in the present volume.

While the Pampa-Virupaksha cult just noticed is for the most part confined to the Virupaksha temple in Hampi village, the Hampi site as a whole is charged with a di erent mythology. According to local belief, this is Kishkindha, the legendary

monkey kingdom, where Rama and his brother Lakshmana arrived seeking for Rama's wife Sita who had been abducted by Ravana, the wicked king of Srilanka. As related in the Kishkindha chapter of the *Ramayana* epic, Rama befriends Sugriva, the dispossessed monkey king of Kishkindha, and together with Lakshmana helps him gain his rightful throne from Vali, his monkey brother. In turn, Sugriva introduces the brothers to Hanuman, the valiant monkey warrior, who sets o to find Sita. Episodes in the Kishkindha chapter of the *Ramayana* are identified with particular rocky caverns, rugged mountain summits and river banks and pools dotted around the Hampi site. That these associations continue to imbue the site with sacred meaning is evident from the pilgrims who make the journey to Hampi, not to marvel at the ancient ruins, but to experience the story of the *Ramayana* as a still-living narrative. For the emperors of Vijayanagara, Rama served as an image of an ideal ruler presiding over a righteous kingdom. Little wonder that they erected a splendid chapel dedicated to this deity for their private use in the middle of the Royal Centre of their capital, which they adorned with carved illustrations of the epic. Furthermore, they ensured that their Rama temple was related to the surrounding landscape. The north and doorways of the mandapa preceding the sanctuary frame distant views of Matanga and Malyavanta hills, two natural features linked with specific episodes in the *Ramayana*.

In addition to these legendary associations, Hampi's site also o ered significant practical benefits that made possible the growth of the capital. As already mentioned, Hampi's landscape is traversed by the Tungabhadra, the largest river in this part of the country. Due to the rocky terrain, the river loses height as it flows across the site, permitting channels to be run o   at higher levels so as to conduct water down into the surrounding fields. This, then, is the basis of the sophisticated hydraulic system, which made possible the city's growth. Only in this way could sufficient food be produced to sustain an urban population estimated as several hundreds of thousands. The network of channels that runs through Hampi's site is also partly fed by extensive tanks with massive earthen embankments to trap monsoonal rains. Present-day agriculture in the Hampi region still relies on these ancient channels and tanks.

In addition to its source of essential water, Hampi's rugged landscape also o ered another crucial advantage: that of defence, especially from the north, the direction from where most conquerors of the region before and after the establishment of Vijayanagara arrived. This aspect of the city's granitic environment helps explain how Vijayanagara came to be located in what might seem at first a highly unsuitable site for an imperial capital. In fact, Hampi's landscape functioned as a natural citadel that was exploited by local rulers. And indeed this is what happened in the years before the foundation of the city. Tucked away in the rocks some 15 km northwest of Hampi is the fort of Kumatgi, headquarters of a local line of rulers who governed the Tungabhadra valley, and who resisted the conquering troops of the Delhi sultan at the very end of the thirteenth century. Though succumbing eventually to the invaders, it was a set of brothers, whose father Sangama had served as a military commander of Kumatgi chief, who successfully expelled the intruders from the region. By the middle of the fourteenth century, Hakka and Bukka, the first two Sangama brothers,

had laid the foundations of what was to become the greatest and wealthiest city of its time in southern India. Among their first tasks was to fortify the City of Victory with massive ramparts that took advantage of the surrounding landscape, by running ramparts up and over boulders and along rocky ridges so as to create an impregnable citadel. And throughout its almost 200-year long history as capital of a vast empire, the city was besieged by its enemies but never captured until the final catastrophe of 1565, when it was abandoned and totally destroyed.

2   Foundation Growth Destruction Rediscovery

Having considered the extraordinary qualities of Hampi's natural setting it is now necessary to survey the equally extraordinary circumstances of Vijayanagara's history. Both the foundation and devastation of Hampi are best understood as consequences of profound political disturbance. The first of these occurred at the very end of the thirteenth century, when peninsular India succumbed to the depredations of the troops of the Delhi sultans. Though the invaders succeeded in extinguishing the previous Hindu lineages that had governed the region, they were unable to hold onto their conquests for any length of time. By the third decade of the century, they had retreated, leaving what was in effect a power vacuum. This situation provided opportunities for newly emerging leaders, none more successful than the Sangama brothers of the Tungabhadra valley. Within merely a few years, Hakka and Bukka had galvanised the leaders of southern India, forging their territories into a rapidly expanding polity that assumed the name of Vijayanagara, the city that they established at Hampi on the bank of the Tungabhadra. Under their rule, Hampi became the headquarters of a domain that rapidly became a veritable empire, encompassing all the diverse peoples and regions of peninsular India south of the Tungabhadra. Hakka and Bukka, together with their descendants, constituted the first dynasty of Vijayanagara. The Sangama rulers were outstanding military leaders, capable of compelling all the subordinate chiefs of southern India to proclaim their allegiance. Adapting to the innovations in warfare introduced into peninsular India by the troops of the Delhi rulers, they also proved capable of resisting the forces of the newly founded kingdom of the Bahmani sultans that lay immediately to the north. These innovations included an emphasis on cavalry contingents, which meant that the Sangama army employed Muslim officers who were skilled in the training of horses and the deployment of these animals, most of which had to be imported, in military tactics.

By the time of Devaraya I and II, two emperors of the same name who reigned successively in the early fifteenth century, Vijayanagara was capital of the most powerful kingdom in peninsular India. The immense affluence of its rulers derived from taxation and agriculture, as well as from valuable deposits of minerals and even precious diamonds. The lucrative Arabian Sea ports that came under Vijayanagara control gave access to the trans-ocean trade with the Middle East, while ports on the Bay of Bengal profited from lucrative textile exports. Through the course of the fifteenth century, the Sangamas exploited the wealth of their outlying provinces,

ensuring that Vijayanagara was developed into an unparalleled showcase of imperial magnificence. The wealth that they displayed at the capital was underpinned by the treasure that they seized in the military campaigns against their Bahmani neighbours to the north, and also the Hindu rulers of Orissa to the northwest, as well as by the taxes paid and military personnel donated by subordinate chiefs throughout their domains.

To confirm the loyalties of their commanders and governors, the Vijayanagara emperors devised a spectacular event at the capital, to which all these figures were compelled to attend. Each September–October, the capital witnessed the Mahanavami, or Nine Nights, festival, marked by magnificent parades of royal animals, military contingents and courtly women, displays of fireworks and martial contests and huge feasts. The central rite of this occasion was the worship of the goddess Durga by the Vijayanagara emperors in order to empower his troops and weapons. Only after this was concluded could they set o on warring campaigns. Foreign visitors were especially welcome at Vijayanagara during the Mahanavami, and accounts of the festival have come down, including that of Abdul Razzaq, an envoy of the Timurid ruler of Herat in Central Asia, who was at the capital in 1443. After the turn of the sixteenth century, when the Portuguese captured the Arabian Sea horse trade from the Arabs, it was Europeans who made the journey to Vijayanagara. Their chronicles o er tantalising glimpses into the daily life of Hampi, especially its teeming markets.

Towards the end of the fifteenth century, the influence of the Sangamas and the wealth of their empire declined. And so it came to pass that military leaders usurped the Vijayanagara throne on two occasions, in the process overthrowing the Sangamas. It was Vira Narasimha Tuluva who headed the second of these coups in 1505, thereby inaugurating the third dynasty of Vijayanagara. During the reigns of his two half-brothers Krishnadevaraya and Achyutaraya who succeeded him, the fortunes of the city and empire were restored. Hampi was developed on an unprecedented scale with huge temple complexes being laid out along the Tungabhadra financed by the ever-expanding provinces and the military successes against the newly established Muslim kingdoms of Bijapur and Golconda to the north. And it is against this background of prosperity and power that the circumstances that led inexorably to the final catastrophe unfolded. Sadashiva, the last of the Tuluva emperors, was unable to enjoy his rule for any length of time since Ramaraya, commander of the Vijayanagara army, seized control of the throne, becoming in e ect supreme ruler. But deteriorating relationships with his Muslim neighbours marred his rule. So aggressive were his dealings with the sultans that they temporarily laid aside their rivalries to create a military alliance of colossal proportions. And it was these forces that confronted those of Ramaraya at the battle of Talikota, fought in January 1565 at a site some 100 km away from Hampi. Soon after the hostilities commenced, Ramaraya was killed and his foot soldiers and cavalry deserted. In the short time, it took for the troops of the sultans to reach Hampi, Sadashiva and the Vijayanagara court escaped for safety, bearing the imperial treasury and whatever precious items could be carried. Once they learned of the defeat, the population must also have fled, leaving a virtually empty and defenceless city for the conquerors to pillage, demolish and burn.

The plundering must have been thorough since all the timbers of the palaces and pavilions were burnt, and the sanctuaries of the temples and shrines systematically robbed. Evidence of these destructive acts is seen to this day in upturned floor slabs, scorched basement mouldings and smashed wall slabs.

Once this orgy of pillage had run its course over several months the jubilant troops departed, leaving the city wrecked and depopulated. Sadashiva and his successors, now relocated for safety at Penukonda and then Chandragiri in what is now the southern part of Andhra Pradesh, several hundreds of kilometres to the southeast, made little attempt to return to Hampi. Perhaps the location of the city insufficiently distant from Bijapur and Golconda dictated the decision to permanently relinquish any claim on the capital. And so it came to pass that Hampi came to be permanently abandoned. Occasional visitors to the site towards the end of the sixteenth century describe a wasteland inhabited only by wild animals. And in time nature took its toll, with stone structures succumbing to bushes and trees, and earthen terraces on rocky hillsides being flushed down by monsoonal rains into valleys, altogether burying roads and buildings. And it is this wilderness that is recorded in the first map of the site prepared in 1800 by Captain Colin Mackenzie, which records what could be observed at the time of the city's fortifications and visible monuments. For Alexander Greenlaw, the pioneer photographer who was at the site in 1856, almost exactly 400 years after Talikota, it was not the documentation of the city's ruins that was the primary attraction, but rather the romantic vision of picturesque decay.

During the years after 1565, few visitors made the journey to the site, since this had become dangerously malarial. Even so, worship at Hampi seems to have continued somehow and the Virupaksha temple came to be refurbished on several occasions in the eighteenth and early nineteenth centuries. The vandalised basement of the east gateway in front of the monument was topped with a newly constructed pyramidal brick tower, presumably replicating the original tower that had been dismantled in 1565, and the paintings that cloaked the ceiling of the hall in front of the main shrine were entirely reworked. At the same time, several pavilions for spectators were added to the colonnades lining the street so as to better view the chariot festivals that once again took place. The agents of these improvements remain for the most part unknown, though there is record of a British District officer ordering the completion of the north gopura of the temple in 1832. These and other such efforts testify to a revival of religious ceremonies within the temple as well as the festivals celebrating the betrothal and marriage of Virupaksha and Pampa in the street outside. This contributed to Hampi's fame as the premier Hindu pilgrimage destination in this part of southern India, a reputation it continues to enjoy to this day, thereby explaining why the ruins are better known today as Hampi rather than Vijayanagara.

As for the vestiges of Vijayanagara's past glory, from the end of the nineteenth century, these came under the protection of the archaeological authorities, who began clearing the various structures of clinging vegetation and accumulated dirt so as to rescue them from further decay, a task that is still to be completed. Since the 1970s, excavation work by archaeologists in the palace zone of the city has exposed the buried basements of royal palaces, audience halls and other features yet to be identified. One of the most spectacular discoveries was the unearthing of a

magnificent well encased in steps and landings. These investigations, together with the mapping project of the overall site by an international team of architects and archaeologists, co-directed by the author of this chapter, stimulated considerable interest in Vijayanagara both within India and internationally. In 1986, the words 'Hampi Monuments' were inscribed on the UNESCOs prestigious World Heritage list, and since then the site has come under the Hampi World Heritage Area Management Authority.

3   Urban Core Royal Centre Irrigated Valley Sacred Centre

No appreciation of the overall Hampi site is possible without acknowledging its past splendour. Quite simply, the ruins constitute the earliest and best preserved architectural and archaeological record of an imperial Hindu city in India. Unlike previous dynastic capitals in the country, like Thanjavur of the Cholas or Halebid of the Hoysalas, for instance, of which only grand religious edifices still stand, Hampi preserves an abundance of fortifications, defensive gateways and watchtowers; audience halls, royal residences, pleasure pavilions, stables and stores; and countless temples and shrines consecrated to di erent Hindu divinities, in addition to the occasional mosque and Muslim tomb. These diverse features, in varying stages of decay, are scattered all over the site, in accordance with the local topography. However, Hampi is not a planned city in the usual sense, since it exhibits no regulating, geometric configuration. Instead, the urban area is divided into a series of separate zones ingeniously integrated into the landscape, exploiting wherever possible the natural defences o ered by the rocky ridges and the river. And it is this sequence of zones that give the best idea of how Hampi functioned over the approximately 200 years of its career as the City of Victory.

   The core of Vijayanagara is an irregular zone contained within a complete circuit of fortifications, extending more than 5 km along its greater, southwest–northeast axis. On its northern flank, the ramparts climb up and run along the rocky ridges, resulting in an irregular line of walls; on the south, where the ground is level, they are more linear. Defensive gateways set into the walls indicate a radial system of roads that proceed into the middle of the city, converging on the palace area which we term the Royal Centre. The gateways are preceded by protective barbican enclosures and are topped by lookouts and watchtowers. The walls themselves are faced with massive granite blocks, ingeniously fitted without any mortar, reinforced with regularly spaced, rectangular bastions. Within the walls is a profusion of stone, religious structures that include temples and shrines consecrated to Hindu divinities and even Jain saviours, as well as mosques and tombs in what we term the Muslim Quarter. Numerous wells and tanks, as well as pottery fragments and stone block mortars for grinding grains dotted throughout the Urban Core testify to an extensive population.

Fig. 3 The Vitthala temple complex

Fig. 4 Hazara Rama temple in the Royal Centre

But of the actual dwellings of these inhabitants, presumably built of ephemeral materials like rammed earth, bamboo and thatch, nothing survives (Fig. 3).

The Royal Centre of Vijayanagara is situated within the Urban Core, but displaced towards its south-western end. This part of the city comprises a number of high-walled compounds that contain a variety of ceremonial, residential and recreational structures. It is here that the emperors, their retinues and household lived, worked and entertained. In the middle of the Royal Centre is the king's private temple known as Hazara Rama temple (Fig. 4) dedicated to Rama, which has already been referred to. That Rama had a crucial role to play in mediating between the different activities of the Royal Centre is suggested by the disposition of the various structures around the temple. The walled compounds to the east are termed the Zone of Royal Performance, since here are located the structures linked with the more

public activities of the emperor and his representatives, such as the multistoreyed platform associated with the Mahanavami festival, hundred-columned audience hall, ceremonial bathing tanks, and elephant stables overlooking the parade ground, all guarded by lofty watchtowers. The walled compounds to the west of the Rama temple we call the Zone of Private Performance, since here are found most of the residential complexes, many with complicated entryways for guards to ensure privacy. Here, too, is a shrine that replicates the cult of Virupaksha at Hampi, presumably for the exclusive use of the king's household.

Immediately north of the Urban Core is a valley through which flows a tributary of the Tungabhadra river, which supplies the still functioning network of water channels that has already been referred to. The absence of any buildings or pottery fragments indicates that this valley was always reserved for irrigation. Furthermore, it served as a transition between the Royal Centre and what we term the Sacred Centre. This latter zone is composed of a number of discrete urban units, referred to in ancient inscriptions as '*puras*', including the present-day village of Hampi, located on or near the Tungabhadra in the most rugged part of the Hampi site. Each pura focuses on an impressive religious monument surrounded by high walls and entered through one or more monumental gateways with soaring pyramidal towers known as gopuras. The largest of these temples are consecrated to Virupaksha, Balakrishna, Tiruvengalanatha and Vitthala (Fig. 3). While the cult of Virupaksha, as already noticed, is indigenous to Hampi, the cults of the other three deities, all aspects of Vishnu, were introduced to Hampi by the Vijayanagara emperors, no doubt as part of a conscious policy of creating an imperial pantheon in their capital. All these temples are approached along paved roads lined with colonnades used for annual chariot festivals. While the entrance gopuras and pillared halls within the temples are typical of southern Indian religious architecture, as found at innumerable religious centres throughout the region, the pillared roads that run through the Sacred Centre appear to be a unique, Vijayanagara invention. The huge scale of the temple compounds, hemmed in by high walls, and the considerable length of these streets, some extending almost 1,000 m, together with the facilities for priests, devotees and attendants, such as bathing tanks, wells, rest houses, kitchens and stores, still convey an idea of Hampi's urban grandeur (Fig. 3).

South of the Urban Core and Sacred Centre stretches an uninterrupted, level territory, in which a number of settlements grew up that like Hampi itself, continue to be inhabited. These villages and towns are strung along the various routes that lead into the Urban Core of the capital. At Malpannagudi, for instance, the gateways through which one of these roads passed can still be seen. Temples and shrines, wells and even a pair of Muslim tombs at Kadirampuracan still be seen, unlike the dwellings of the original inhabitants that have disappeared beneath recent developments. This is even true of the imperial residence of Krishnadevaraya, who shifted his headquarters to Hospet, the modern town that serves as the nearest railhead for Hampi, some 12 km away (Figs. 5, 6 and 7).

Fig. 5 Stepped bathing well in the Royal Centre

Fig. 6 Pushkarni water tank at the Krishna temple complex

Fig. 7   Inside the Vitthala temple complex

4   Walls Gateways Palaces Pavilions Temples

Coming now to Hampi's unrivalled architectural heritage, which, in spite of the devastation of 1565, still preserves the broadest possible range of military, courtly and religious buildings, matched by an equally broad spectrum of constructional techniques and building styles. The huge stone blocks with which the earthen ramparts of the Urban Core are faced have already been mentioned. The blocks are cut and fitted with remarkable precision to create irregularly jointed walls, without any mortar. The same is true for the defensive gateways provided with broad passageways roofed with corbelled beams. The compound walls of temples in the Sacred Centre and enclosures in the Royal Centre exhibit the same type of jointing on both external and internal faces. But the walls in the latter zone are unusually slender and tapering, more for privacy than protection.

Since solid granite blocks were employed for walls, columns and roofs of temples and shrines, it is hardly surprising that religious institutions are among Hampi's most completely preserved structures. The earliest shrines, as on Hemakuta hill, have pyramidal stone roofs rising over the sanctuaries, a typical feature of earlier temple architecture in this part of India. However, from the fifteenth century on, temple towers at Hampi were built of brick and coated with plaster, a lighter and quicker method of construction. Similar brick and plaster towers supported on granite walls mark gateways in temple compound walls. Known as *gopuras*, these pyramidal superstructures have badly eroded, and preserve little of their intricate detail. Temple walls are composed of closely fitted blocks, their surfaces raised on basement mouldings, and divided horizontally into narrow bays by slender pilasters with lotus bud brackets. The mastery of stone cutting is most evident in the pillared mandapas that precede temple sanctuaries. Columns have their shafts divided into blocks sculpted with a profusion of images of gods and goddesses, semi-divine figures, saints and devotees. Sixteenth-century temple mandapas have columns fashioned as complicated

piers with cut-out colonettes as well as fantastic leonine beasts called *yalis*, often ridden by warriors. Such designs filter the external glare, as in the halls added to the Virupaksha temple at Hampi and the Vitthala temple in the Sacred Centre.

As already mentioned, the largest religious monuments at Hampi are conceived as complexes of sanctuaries and minor shrines contained in walled compounds entered through one or more gopuras. In the Tiruvengalanatha temple, better known as Achyutarays temple, the sanctuaries of the god Vishnu and his consort Lakshmi are contained within two concentric walled compounds, accessed through a pair of identical gopuras. Sanctuaries are approached through two mandapas: an inner hall encased in walls with side doorways; and an outer, open mandapa with elaborate pillars of the type already described. The Anatashayana temple on the outskirts of Hospet is the only religious monument with a rectangular sanctuary roofed with a lofty brick vault. This was intended to accommodate an image of reclining Vishnu, the god to whom the temple was dedicated. Composed of plaster-coated brickwork, the image has entirely disappeared, leaving only the long stone pedestal on which it was once displayed.

In striking contrast to these all-masonry religious buildings, many of the structures in the Royal Centres at Hampi were only partly stone built. While the excavators have exposed the masonry podiums of these buildings, their plaster clad brick walls, timber columns and tiled roofs have altogether disappeared. Such construction methods were generally selected for audience halls, palaces, barracks, stores and other buildings linked with imperial governance and courtly life that have yet to be identified.

The audience hall that is the largest structure of this type in the Royal Centre has exactly one hundred columns according to the lines of stone footing blocks set into its plaster floor, but its timber columns and lofty ceiling are lost, no doubt burnt in 1565. Palaces conform to a standard design, with ascending floor levels arranged in U-shaped formations. Steps flanked by yali balustrades climb to small square chambers at the uppermost levels, which were probably reserved for reception and sleeping. The designs of palace roofs, perhaps rising in sloping tiers and topped with gleaming metal finials, however, remain a matter of speculation.

An altogether di erent architectural tradition is also present in another series of better preserved buildings in the Royal Centre that are entirely constructed of plaster-covered rubble, featuring pointed arches and recesses, as well as domes and vaults. All these attributes derive from the buildings of the neighbouring Muslim kingdoms, as in the mosques, tombs and palaces in Gulbarga and Bidar. In spite of this obvious Islamic origin, this architecture was not in any way deemed inappropriate for the palace zone of a Hindu imperial city. Indeed, it seems that the emperors of Vijayanagara promoted this style for their Royal Centre, perhaps as part of a cosmopolitan outlook that embraced other peoples and cultures. And it is worth recalling once again that Muslims had been welcome at the Hindu capital from earliest times, and appear in sculpted representations on the Mahanavami platform and in reliefs on the Rama temple in the middle of the Royal Centre.

Sultanate influenced buildings at Hampi are mostly concentrated in the walled compounds of the innermost domain of the emperors. They include the Lotus Mahal

Fig. 8 The Lotus Mahal at Royal Centre

(Fig. 8) in the northeast of the Royal Centre. Though popularly understood as forming part of the zenana, or private women's quarter, of the Vijayanagara palace, this is more likely to have served as a meeting place for the emperors and their commanders, since the elephant stables and martial parade grounds are only a short distance away. The Lotus Mahal is laid out on a square, symmetrical plan, with prominent projections in the middle of each side. It rises vertically in two storeys, each with lobed arched openings surrounded by bands of richly encrusted plaster ornamentation, and sheltered by double-curved overhangs. A cluster of nine pyramidal, temple-like towers rises above the corners and midpoints of the pavilion. Watchtowers set into the surrounding compound walls, one with a square staircase shaft, the other octagonal, indicate that privacy was a concern in this part of the Royal Centre. Both watchtowers have arched windows and projecting balconies with balconies, sheltered by curving overhangs, as well as fanciful towers topped with temple-like finials.

By far the most majestic examples of sultanate influenced architecture at Hampi are the elephant stables (Fig. 9) facing onto a spacious parade ground (now planted inappropriately as a lawn). The stables present a long line of 11 square chambers, each with an arched opening, within which a richly decked animal could have been displayed on ceremonial occasions. The chambers are topped with domes, both with smooth and fluted surfaces, alternating with 12-sided, multi-tiered vaults disposed in strict symmetrical formation about a central upper chamber. This chamber was most likely reserved for musicians who accompanied the pageants of troops and animals place below. The northern boundary of the parade ground is occupied by a similarly long building with a raised arcaded verandah, from where the emperors and their

Fig. 9   The elephant stables in the Royal Centre

guests could observe these displays. The interior of this grandstand-like building has a ground level court, an ideal for the mock fights between soldiers, and even with animals, that were much enjoyed at the time.

Buildings at Hampi in the sultanate influenced style were also commissioned for the more private pleasures of the emperors and their nobles. They include the queens' bath, which, in spite of its present name, is unlikely to have been used by royal women since the building is located some distance from the private zone of the royal household. The building looks inwards to a square pool surrounded by arcaded corridor with ornate balconies projecting over the water, from where courtiers could gaze at the bathers. Other water structures on the fringes of the Royal Centre include an octagonal pavilion with a fountain, and an octagonal outdoor pool surrounded by a colonnade. That such octagonal schemes are a particular feature of this sultanate influenced idiom is confirmed by the two-storeyed pleasure pavilion in the western part of the Royal Centre, in the middle of the residences of the emperors' household.

Structures in this style are also occasionally found beyond the confines of the Royal Centre, like the gateway with four lofty pointed arches carrying an imposing dome, southeast of this zone. A chamber with typical arched windows and a parapet of ornamental battlements surmounts the gate leading to the river crossing of Talarighat to the northeast. Then there is the water structure beside the road linking Hospet to the capital, which has a deep octagonal well surrounded by arched recesses, accessed by a long flight of steps. A U-shaped pavilion with open arcades, hidden from view in the sugarcane fields and banana groves of the irrigated valley

beyond the walls of the Urban Core, probably formed part of a country estate, but whether belonging to the emperors or one of their courtiers, is not known.

This rapid survey of Vijayanagara's landscape, turbulent history, zonal urban layout and variant architectural styles is intended to convey the immense value of Hampi as one of India's premier heritage sites. While the investigation of its ruins over more than 100 years encompasses all available techniques, including the advanced digital technologies showcased in the present volume, the preservation of the monuments in their natural setting still lacks an e ective, comprehensive plan with adequate implementation. Only as an archaeological park, encompassing the entire site, can Hampi in its totality be protected. It is the fervent wish of the present author that all the agencies presently involved in managing the site will work in harmony so as to preserve Vijayanagara's past glory for future generations.

IDH Snippets

P. Anandan Vidya Natampally and Srinivasa Ranganathan

1 Indian Digital Heritage: The Way Forward

By P. Anandan and Vidya Natampally

India is known for her rich cultural and historical heritage. India's landscape is studded with historical sites and monuments that tell stories about vast kingdoms, great empires, their architecture, culture and civilization. India has had a continuous living culture that has generated a vast treasure of written, visual and performance art and craft forms that attest to the glory of the civilizations that have occupied her land. The continuously evolving Indian tradition of higher art has meant that much of the origins and history of this evolution is known only to a few people. The task of documenting, archiving, and sharing India's heritage is itself a monumental task, and despite dedicated efforts by many people and the government, it still remains a challenge.

We live in an age of digital storage and restoration. Digital technologies have the capacity to protect memories of the past against the destructive forces of time, nature and man-made events. They also have the power to bring both the tangible and intangible components of our heritage to life by offering vivid experiences and seamless interactions to everyone. The emerging technologies in computer vision, graphics, audio and video technologies and user interface design along with new hardware like the handheld devices, gesture-based devices, virtual reality and augmented reality infrastructure offer the prospect of creating interactive and real experiences of heritage for the users. Like all technologies, the pace of technical advancement increases when they are applied to interesting and challenging real-world scenarios.

P. Anandan   V. Natampally   )
Bangalore, India
e-mail: vnatampally2011@hotmail.com

S. Ranganathan
National Institute of Advanced Studies, Bangalore, India

At the time of the founding and inauguration of Microsoft Research India in 2005, it was realized that there was a unique opportunity to simultaneously advance the state of research in computing technologies with a key to capture, preserve and experience cultural heritage and to work on personally satisfying and rewarding scenarios related to India's vast heritage. The rst foray of Microsoft Research MSR) India was to create an interactive walk-through narrative of the Sri Andal temple in Srivilliputhur, Tamil Nadu, by combining multiple types of rich media, interactive 3D exploration and storytelling.

With the Srivilliputhur demo in hand, we approached the Department of Science and Technology DST) to ful ll the desire to advance the state of research in the key technologies while working on challenges related to Digital Heritage. The DST strongly supported these interests and joined us in organizing a planning workshop at IIT Delhi, where the key researchers working in related technology areas assembled and developed a common vision for the project. At this event, the impetus was obtained from the then Secretary of DST, Dr. Ramasami to initiate a unique research project that took a holistic approach to bringing together the culture and technology components of the research. We reached out to some of the leading experts on Indian history, art, and culture such as Prof. S. Settar at the National Institute of Advanced Studies and Prof. S. Ranganathan, a leading scientist also at NIAS. The Secretary of DST requested a proposal for a project and directed Maj. Gen. Dr. Siva Kumar and Dr. K. M. Murali Mohan at the NRDMS to oversee the effort. The result was the Indian Digital Heritage IDH) research project.

1 1   The IDH Research Project

The basic objective of the IDH project was to bring the power of GeoICT, a ne synergy of geospatial technologies with matured information communication technologies and development of technology tools to help preserve, use and experience India's vast heritage in digital form. With this background, DST constituted a Programme Advisory and Monitoring Committee PAMC) on GeoICT under the Chairmanship of Prof Ashok Jhunjhunwla, IIT Madras. IDH was one of the R and D tracks of broader GeoICT programme launched by DST NRDMS. A Project Advisory Committee was constituted under the chairmanship of Dr. Gautam Sengupta, the Director General of the Archaeological Survey of India ASI) to monitor the progress of the IDH project. Hampi in Karnataka, a UNESCO World Heritage Site was chosen as the ASI site for this project. Hampi, with an area of approximately 20 km^2, offered diverse and exciting challenges to the research community both on the culture and the technology aspects.

1.1.1   Goals of IDH

The goals of the IDH research project were formulated as follows:

1. Investigation of the basic requirements related to digital heritage archiving and usage: Encourage collaboration between the culture and the technology

communities, and to identify the nature and form of heritage that is most suitable for digital capture, storage and usage.

2. Data collection and archiving: Create a community-based effort—data, sources, standards, schemas as well as the necessary technology to enable the storage and processing of relevant media.

3. Research and creation of tools and technologies: Identify and address technological research issues that arise during the course of the project and further the state of the art by innovating new technologies.

4. User Experience: Create compelling and holistic end user experiences that highlight the historical and artistic signi cance of various monuments and other types of heritage content.

In addition to these technical and technological goals, the project had a few key societal goals:

1. The project was to bring together diverse groups from the technical and cultural communities to work together in a key showcase effort.

2. The project was to create a set of technologies and tools that can be used for similar efforts and be made available to researchers.

3. The project aimed at the establishment of close collaborations between national and international agencies to work towards delivery-based end results. This would allow both—an important and useful exchange of ideas and also prepare these agencies for further exploitation of the technologies after the project is completed.

The project grew and matured over years and culminated with a showcase of its outcomes at a workshop and exhibition held at an important convention centre in New Delhi in 2014. The rest of this book records its progress and innovations. We feel privileged and grateful for this opportunity to have been able to dream and initiate an effort that uniquely combines technology and culture research on a topic close to the heart of every Indian—namely the rich and wonderful cultural heritage of our country.

2 Indian Digital Heritage: Reimagining Vijayanagara

By S. Ranganathan NIAS

In the spring of 2011, while releasing the iPad2, Steve Jobs voiced his vision of the intersection between liberal arts and technology, saying It is in Apple's DNA that technology alone is not enough—it's technology married with liberal arts, married with the humanities, that yields us the results that make our heart sing'. The Indian Digital Heritage project—an initiative of the Department of Science and Technology DST), Government of India, is a stellar example of similar synthesis.

In this section, we would like to start with mentioning the pivotal role played by the National Institute of Studies NIAS), Bengaluru, which was a base for the start of this project. The NIAS journey to Vijayanagara commenced from the experiences drawn from the International Conference on Tangible and Intangible Heritage of Hampi' organized by the Friends of Hampi and NIAS at the Institute 16–18 January in 2009. This conference explored several aspects of the tangible and intangible heritage of Hampi, such as art, architecture and material culture including performance art, as well as the issues related to heritage management and sustainable development through local activism as represented by the formation of the interest group, Friends of Hampi. Our co-ordinating team from NIAS included Prof Settar, Prof Sharada Srinivasan and myself who were then also drawn into co-ordination of the IDH project at the behest of DST.

The above event furthered the interest of the Department of Science and Technology to hold a workshop-cum-project development meeting on Indian Digital Heritage on 7 January 2009, wherein 40 delegates from different parts of the country were invited and made presentations on technology and cultural heritage aspects. Subsequently, a proposal was made to DST to focus the digital heritage effort on the UNESCO World Heritage site of Hampi, the seat of the rulers of the erstwhile Vijayanagara Empire.

As a preliminary step of the mega project, I visited Hampi along with Prof. Settar to document the current day celebration of the mythological wedding of goddess Pampa Devi to Lord Shiva in December 2010. Prof. Uma V. Chandru of International Institute of Art, Culture and Democracy IIACD), Bengaluru, also joined us in this expedition.

On 29 January 2011, DST sanctioned 22 projects under IDH, with phase I consisting of 9 technology projects, and phase II consisting of 3 technology projects and 10 cultural and heritage projects in a multi-institutional and multidisciplinary mode for a period of 3 years. This was a groundbreaking step in the sense that Science and Humanities were to be working together.

Subsequently, NIAS organized consultation meetings where I met with Prof. Sharat Chandran of IIT Bombay and with Dr. K. R. Murali Mohan of DST during the rst 2 months of 2010 in order to plan the modalities of this mega project and to design the launch programme as well as a workshop combined with a 2 days visit to Hospet and Hampi. Inauguration of the progamme was held on 25 March 2010 at NIAS and a workshop was held at Hospet on 26 and 27 March, attended by all the Project Investigators PI), Co-PIs and associates including PAC and PAMC members. The IDH project was nally launched on 25 March 2011 by Prof. V. S. Ramamurthy, the then Director of NIAS and former Secretary of DST.

At this point of time we may step back and review contemporary developments during the past 6 years. During a visit to Australia, Prof. Settar came across an exhibit called Place Hampi'. It was created by Prof. Sarah Kenderdine,[1] a world-renowned maritime archaeologist now working with immersive and new

---

[1]http://www.niea.unsw.edu.au/people/professor-sarah-kenderdine.

digital media. This led to interactions with Jindal Foundation and Kaladham at Jindal South West JSW) Steel, which features the   rst Digital Museum in India.

Interactions with Taiwan and Japan also deserve special mention. As chairperson of a six-member delegation, I went to Taiwan and presented papers on Indian Heritage preservation at the National Science Centre Taiwan. Prof. S. Settar elaborated on IDH and efforts being undertaken at Hampi. To mark 60 years of diplomacy between India and Japan, Tokyo University of the Arts organized India Day', where I was invited to present a special lecture on Hampi. The links with Prof. Kastushi Ikeuchi[2] were further strengthened by visiting him at the University of Tokyo and learning about his seminal work on Angkor Wat.

The recent unrest and consequent destruction of the precious heritage monuments in the Middle East has given further impetus to the   eld recognized as Digital Meets Culture'. Thanks to LIDAR and laser scanning, it is possible to create 3D images of the monuments. A synergy with 3D printing enables re-creation in physical space. The IDH Workshop and Exhibition held at New Delhi in November 2014, helmed by Prof. Santanu Chaudhury of IIT Delhi, presented the remarkable innovations of IDH to the public. This was followed by the Science Express train travelling through India carrying the exhibits showcasing the technologies developed as part of IDH. Unlike its physical predecessors, these new digital avatars overcome the barriers of space and time and their ravages. We welcome everyone to join us in this digital Vijayanagara yatra presented through chapters in this book.

---

[2]http://www.cvl.iis.u-tokyo.ac.jp/~ki/.

Part II
Modeling and Representing
Tangible eritage

Making of Hampi—An Attempt to Bridge Culture and Technology

Meera Natampally

1 Background

The Hindu temple architecture is a typical trabeated style of construction based on precise grid design and symmetry. The Vijayanagara architecture is said to be culmination of Dravidian temple building tradition, a foremost temple building style of southern India. The marvelous architectural design of the temple complex and its elements with the structural system using stone is an impeccable synergy between structural innovation and architectural expressions.

Hampi, a picturesque village in Bellary district, Karnataka, rich in heritage and culture, is surrounded by lush green valley fed by Tungabhadra River. Being the capital city of Vijayanagara empire, Hampi was the most developed city during its glory. The city had fallen into destruction after Islamic invasion in 1565 and was rediscovered in 1800s by the British in the Madras presidency. From then onwards till today, Hampi has been promoted as important archaeological site, also recognized by UNESCO.

The city is divided into royal and sacred centers as per the rituals and functioning of its spaces as shown in Fig. 1 (a). The focus is on the Vittala temple complex which is located in the sacred center of the Vijayanagara Empire.

The orientation of the temple proper is to the normative east with Bazaar Street towards east and north. See Fig. 1 (b) The Vittala temple—processional path in the east, the bank of the river in the north and west, hills in the south. The influence of Srivaishnava religion is prevalent at this temple complex as revealed by the iconography of the minor shrines to the south, west and north. The study forms an understanding of the social impact of the sect on the temple complex [1].

M. Natampally( )
National Institute of Advanced Studies, Bangalore, India
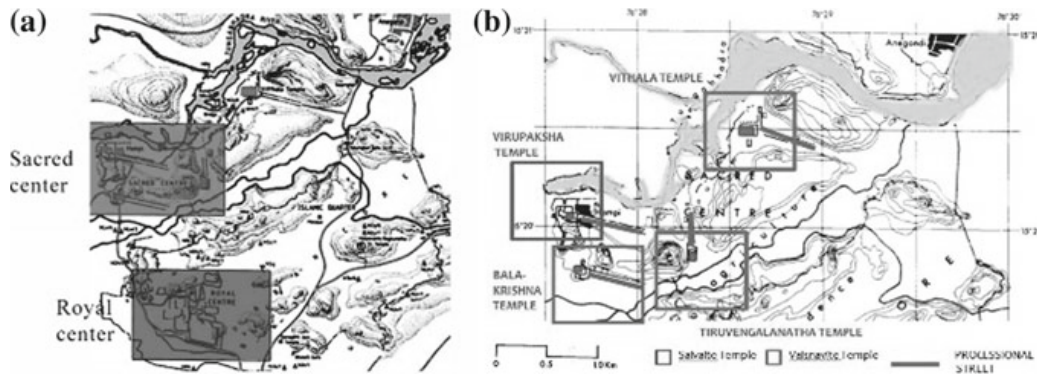e-mail: mn.architect@gmail.com

Fig. 1 **a** Plan of Hampi showing royal and sacred centers, [2] **b** Four puras or districts in the sacred center of Vijayanagara showing Vittala temple and the bazaar street [3]

2 Objective

Objective is to visually reconstruct the missing parts of the temple complex in detail which is an attempt to restore it to its original glory digitally.

Objective of the research is to digitally interpret Hampi/Vijayanagar architecture for interaction and understanding by commoners. To create an open axis digital cultural knowledge bank with digital archives on the evolution of temples architecture of Hampi, where data can be constantly updated by other interested researchers practitioners and broader public. Represent the temple complexes and temple elements graphically, pictorially, and digitally and in turn making it navigational and participatory to all users.

3 Methods

The readings on Vittala temple complex were done in various scales and approaches. The conceptualization of such a genre could be observed through a thorough documentation of the temple, literature studies, and comparative analysis.

The temple complex and its elements are documented through systematic pictorial documentation, measure drawing, and sketches. Two-dimensional AutoCAD drawings are generated and verified with the actual measurements on site. 3D models are created based on the 2D drawings and details.

Research done at various stages involves through study of the Vijayanagara style, deriving proportions of the elements of the temple complex, study of the photographs tracing the past history, and proportion determination in consultation with sthapathis.

For the visual reconstruction, investigation involves a detailed literature study, photographic study, and comparative analysis of di erent monuments of the same period also in comparison with the pre- and post-Vijayanagara. The study of the Vijayanagara contemporary temples and its predecessors helps in understanding the
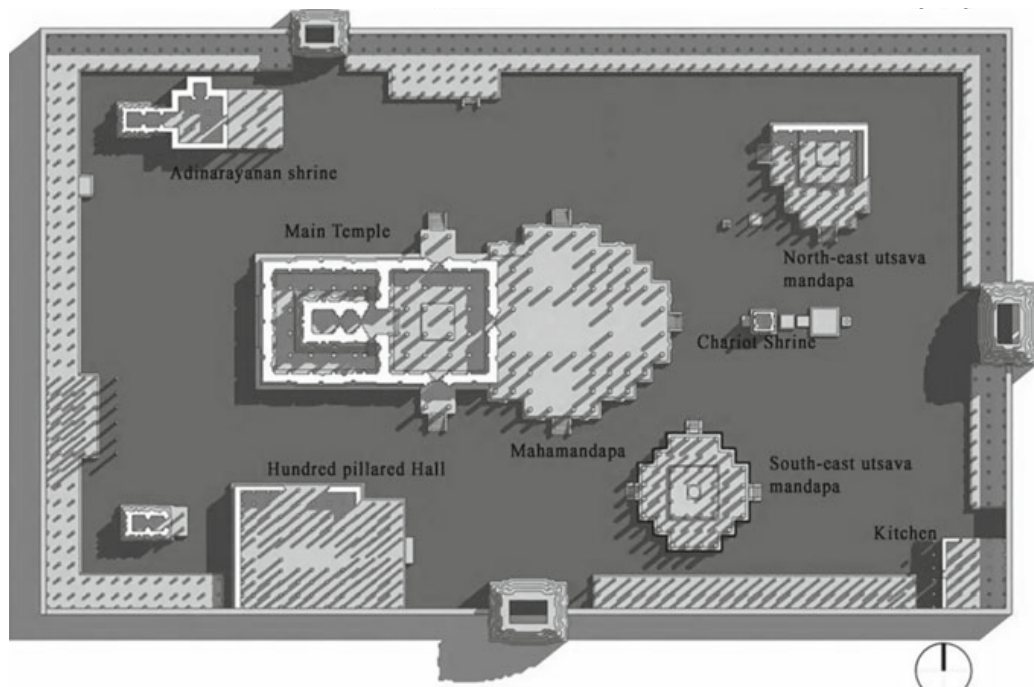
Fig. 2 Plan of Vittala temple, Hampi

proportions which in turn supports visual reconstruction of the missing parts. Visually reconstructed 3D models are generated based on the analysis. The walk-throughs for the temple parts and temple complex are constructed with sketchup models. Narratives are added later in consultation with art historians. Texture is applied to the model which attempts to make the model closer to reality (Fig. 2).

4 Research

Vittala temple has 16 structures within the temple complex. Outline description for all the structures is as follows:

Mahamantapa

The main shrine is approached through an elaborately decorated Mahamantapa. This hall probably being the last major addition to the temple complex is the most architecturally detailed part of the Vittala temple in both artistic/aesthetic and technical aspects. Magnificent and ornately worked mahamantapa adjacent to the main temple, where Vittala and his consort were ceremonially seated in a swing and enriched with music and dance. The mantapa has many musical columns which produce different percussion instrument sounds.

Laid out on a staggered plan, Mahamantapa has 56 columns. The columns and aisles are arranged to create three rectangular open halls and a larger hall in the

middle delineated with a slightly raised floor. The richly decorated plinth consists of friezes of horses, attendants, hamsas, Padma with miniature seated lions and seating yalis, jeweled kumuda, kantha with miniature figures. The mantapa has entrance from three sides with steps flanked on both sides by striding elephants in the east, yalis with fighting warriors in the south, and cutout yalis subduing makaras on the north. There are various kinds of complex columns which hold the roof with massive slabs, some more than 10 m tall supported on multiple brackets. The monolithic columns display complex arrangements of core column shafts, attached collonades and projecting animals or deities or figural columns. The column shafts have flattened tripartite Adhishthana moldings, and triple blocks of carvings separated by 16-sided sections with octagonal bands and prominent Nagabandas. The brackets are of the PushpaPotika type. Over the north and south halls, the brackets rise in triple tiers projecting successively inward. Raised ceilings receive further support from the beams, treated as petalled Kapotas with ribs and pendant buds on the underside and figural panels above. The ceilings are richly carved with stylized flowers, figures, and animals [1].

Missing/dilapidated elements: Roof, doubly curved sunshade, turrets, balustrades, and a few columns.

Main Shrine

The east-facing main shrine stands in the middle of a spacious compound (164 m X 95 m) and has four spaces. The garbhagriha where the deity is placed, and the antarala, intermediate foyer are surrounded on three sides by pradakshinapatha at a lower level. The ardhamantapa, which is the transition space between antarala and rangamatapa, has two stairways on north and south side leading to a sunken pradakshinapatha or circumambulatory space which is lit from small skylight openings around the Vimana or the pinnacle. Rangamantapa is a 16-pillared hall which can be approached either from mahamantapa or the two single-bayed porches on north and south sides. The plinth consists of upaana, Padma, pattika, kantha, inverted Padma, gala with raised bands, occasionally decorated with creeper designs, and petalled kapota with foliated kudus. The walls that rise above have detailed pilasters with pushpapotika brackets, niches, and kumbhapanjara pilasters. A frieze of yalis marks the ends of roof slabs. The single-bayed porches are flanked on both sides by yali balustrades. Balcony seating is provided on the north porch. The ceilings show ornate medallions with pendant lotus buds, surrounded by seated lions, dancing maiden, and dikpalas. The garbhagriha has a square brick tower which is triple storeyed. The lower two storeys have pilaster positions capped with fully modeled kutas and panjaras flanking central larger and higher projections topped with shalas. The topmost storey consists of a circular griva with pilastered niches on four sides framing seated figures [1].

Missing/dilapidated elements: Shikhara, turrets, antarala interiors, columns, pilasters, and doorways.

Uyyalamantapa / Utsavamantapa

This structure occupies the NE section of the temple complex built between 1530 and 1554 AD. It is unusual with asymmetrical plan being open on the south and west accessed by steps and with colonnades projecting outward but walled on east and north. At the north, there is a slightly raised dais attached to the northern wall approached by south and west. The southwest corner of the hall houses two sub-shrines.

The plinth has striding elephant sculptures, dancers, musicians, and female figures. These steps are flanked by elephants. The columns have double capitals with pushpapotikas (banana flower shaped). Columns are with Yalis, fluted miniature pillars, and sculpted bands.

The overhanging eaves are doubly curved sunshade with band of creeper motif. The rituals like uyyala / Diety on swing with music and other performances were conducted on the raised dais [1].

Missing/dilapidated elements: Turrets and doubly curved sunshade.

Sabhamantapa / Hundred Pillared Hall

Dated 1516 AD, sabhamantapa built against the south prakara, originally, was a free standing structure. This pavilion has three levels ascending from east to west. All the other mantapas have concentric arrangements and regular spacing of columns whereas here the spacing of columns is irregular creating wider spaces for assembly.

The structure is open on the east and north. The columns are simpler in their structure and decoration, though they have the common features such as sculpted panels, Yalis and pushpapotikas. It has no elaborate raised plinth and has shallow plinth with elephant, dancers, and miniature sculptures.

The walls have regularly spaced pilasters with octagonal shafts and double capitals. The columns at the uppermost levels are shorter and have seated lions [1].

Missing/dilapidated elements: Turrets and columns.

Kalyanamantapa

It is a pillared festival pavilion which is positioned in the southeast of the temple complex built between 1530 and 1554 AD. It is symmetrical in plan with a central square dais formed by 12 columns.

The plinth or the adhishtana has alternating three projecting tiers with carved elephants. It is approached with the flight of steps on east, north, and west. The other features of the columns, plinth, and sunshade are very similar to the northeast mantapa.

The parapet has turrets built with brick and lime plaster which have ornate decorations. The ceiling panels of the bays display variety of lotus designs surrounded by rings of hamsa, dancers, and musicians. The deity is brought in procession and marriage rituals are performed on the raised dais (Fig. 3).

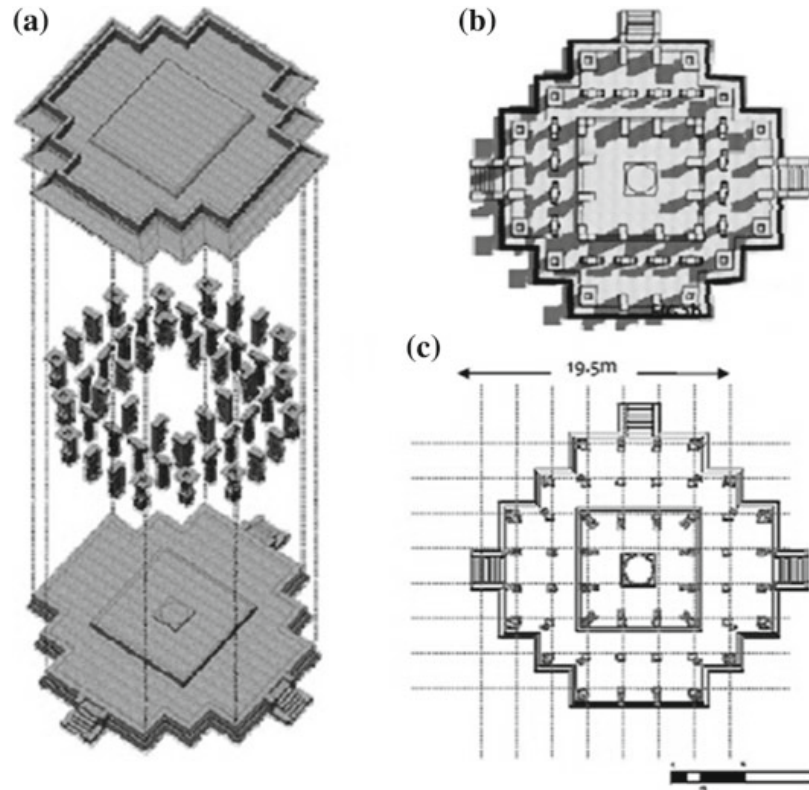Missing/dilapidated elements: Balustrades and turrets.

Fig. 3 Plans and view of the Southeast Kalyanamantapa

Shrines on Northwest and Southwest

Northwest subsidiary shrine is dedicated to Lord Adinarayana built in the fifteenth century. Squares sanctuary, antarala opening into a porch with two freestanding columns originally which was later added with a square Rangamantapa with a small shrine extending outward and an open mantapa with 16 columns. The niches on the outer walls with pilasters have octagonal shafts, double capitals, and PushpaPotikas. Some portions are occupied with Kumbapanjara pilasters. Over the Sanctum Sanctorum is a brick-constructed shikhara which is now dilapidated. The Rangamantapa columns are more ornate whereas the Mahamantapa are very roughly fashioned.

Southwest subsidiary shrine dedicated to Lord Narasimha, built in fifteenth century, has only two columns as freestanding porch in front of the antarala leading to the garbhagriha. This structure is very simple with decorative walls. Exterior is decorative with niches flanked by octagonal shafted double capital pilasters, Kumbapanjaras and well-defined plinth with moldings, figures, and sculptures.

The shikhara over the Sanctum Sanctorum is missing which originally existed [1]. Missing/dilapidated elements: Turrets, wall, and shikharas.

East North and South Gopurams

Constructed in 1513 AD, east gopuram has five Bhumis or tiers crowned with barrel-shaped structure having Kirthimukha at the ends. The structure has stonework up to

a height and then the tapering part constructed with brick and lime mortar. It has rich carvings and sculptures on the structure.

The north gopuram is smaller and simpler, but it is very similar to the east gopuram. It is a triple-storied tower, the tapered portion is constructed with brick and lime mortar.

The south Gopuram is the most elaborate and largest Gopuram in the temple complex. It was constructed in 1538 AD. This is also similar to the east Gopuram, but more rich in sculpture and carvings [1].

Missing/dilapidated elements: Tapering portion in brick.

Lakshmi Narayanashrine

This shrine is located at the west side of the temple complex built into the colonnade. This has a raised platform with columns. The antarala and garbhagriha have doorways, built in 1545 AD. There is a platform extension on the north side. The inscriptions mention a presence of a Yogavarada Narasimha shrine here [1].

Missing/dilapidated elements: Roof and sanctum sanctorum.

Colonnade/Prakara

The colonnade attached to the Prakara has very simple columns with no sculptures but only with the PushpaPotika capitals which carry flat roof and a singly curved sunshade all around.

The columns are in three rows and in addition, a row of half column is attached to the Prakara. Each row of column is different in their design which makes the structure interesting [1].

Missing elements: Columns, column capitals, roof, and sunshade.

Garuda Shrine

This shrine is housed inside a stone chariot which has four wheels very ornately carved. The structure is aligned to the axis of garbhagriha. This is a very unique structure which marks the glory of Hampi. It is square in plan with steps on east flanked by elephants.

The piers are monolithic and columns are tapering octagonal shafted with PushpaPotika above as capitals [1].

The shikhara is built with brick and lime mortar is missing.

Kitchen

It is located in the southeast corner of the temple complex, set into the colonnade. It is partially walled and the roof of the structure is slightly raised from the colonnade roof level to form a clearstory [1].

The visual reconstruction is done to the missing parts/elements by not only studying the structure of the temple complex(explained in Sect. 4) but also by analyzing various aspects of architecture and structure which are explained below.

4 1   Influence of Rituals on Architecture and Vice Versa

The temple complex at Hampi, Srirangam, and Kanchipuram of Vijayanagara style of architecture has minor shrines of the Alvars of Srivaishnava sect. Apart from the Alvar shrines, they included feeding houses and endowments to support utsavams.

The influence of temple rituals can be observed in the spatial organization of the temple complex. The utsavamurthis were taken on processional path during special occasions that circumvents the temple complex apart from the axial one. This influences architecture directly. The kalyanamantapa in Vittala temple has double plinth, where the deity's marriage is performed at the center of the mantapa [4].

The Uyyalamantapa at the northeast of the temple complex again has a second plinth attached to the northern part of the mantapa.

The cloisters around the temple provide shelter to the pilgrims when the rituals take place in the temple complex.

4 2   Strong Role of Geometry in Indian Temple Architecture

The Indian temple architecture inoculates high level of geometric proportions. Different types of proportions can be analyzed from the plan and the elevation of the temple complex. In 1505 AD, the main shrine and the prakara around were built, such that the center of the GarbhaGriha (Sanctum Sanctorum) falls at the center of a square.

The Shrine dedicated to garuda (stone chariot) is at the center of a rectangular portion adjacent to the square mandala. The garuda fall is at the central axis of the garbhagriha, the northern gopuram is also along the central axis of the garbhagriha on the other side, and the southern gopuram is along the axis of the center of the rectangular temple complex enclosure.

With garbhagriha as the center, the nine-square mandala is inscribed as shown in Fig. 4, the Amman shrine and the mantapa (100 pillared hall) fall are outside the mandala.

The 2.4 m x 2.4 m grid is taken from inside the temple complex, when it is extended outside the temple complex, we see that the other mantapas and structures fall with same grid.

The connectivity between di erent temple complex and other important structures is by the processional path of the festivals. The bazaar streets are developed along these paths. Unlike the temple of Madurai and other Chola temples, the development pattern is not concentric, it is more linear in Vijayanagara Empire.
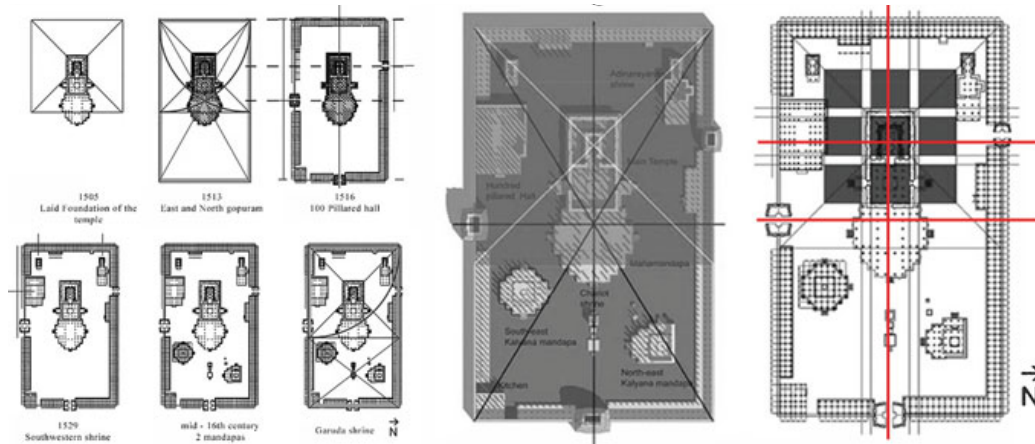
Fig. 4 Development of the Vittala temple complex over the years showing the proportion, Vittala temple complex showing the nine-square mandala

4 3 Spatial Organization of Parts of the Temple

The temple has an elongated plan with the longer direction along principal axis as shown in Fig. 4. The platform on which the deity is installed is at the higher level than the prakara/circumambulatory levels. The other mantapas like kalyanamantapa, uyyalamantapa, and hundred-pillared hall are kept at a little lower level than the sanctum sanctorum. The ceiling height at the center of the mantapa is raised and also the plinth at the center.

Compared to Chalukya, Chola, and Nayaka temples in the Vijayanagara style, the height of the base (plinth/Adhisthana) is very predominant. The scale and visual penetrability in this style increases by reducing the heights of mantapas in elevations without compromising on majesty, but the height of the Vimana is kept as the tallest in the complex.

At the east entry of the temple complex, a stone chariot is built in the form of a temple Vimana. It was originally enshrined with a Garuda the vehicle of Lord Vishnu. The original pictures show the cell with shikhara and the kalasha. It is a monolithic structure in giant granite block with four wheels carved in stone at the base and sculpted panels on the vertical surface.

Mahamantapa along the axis of the main temple has a pillared hall with three entrances. The pillars are of four types (earlier mentioned). In addition, there are exceptionally beautiful musical pillars which give the sound of musical notes and musical beats. Also, the huge sculpture panels are depicting the story of Mahabharatha.

4 4 Structural Systems

The structure of the mantapas is highly developed having different types of orna-mented columns to support the roof slab. The doubly curved sunshade in monolithic stone forms a transitional element between column and roof slab.

The highly carved plinth of the mantapa is in two levels having two sets of columns of varying heights supporting the ceiling. Innovation of T-beams corbelling of brackets and development of complete columns are contribution of Vijayanagara style to the structural system. Supporting the roof beams are characteristic features of Vijayanagara style. The cusped arches and development of parapet are the integral part in elevation which again structurally hold the roof in place.

4 5 Grammar of Columns and Design Elements

Basically, we have four types of columns. One with Yali, other with sculpture pan-els, and yet another with miniature musical pillars, etc. as shown in Fig. 5. The Yali column is a development from the thirteenth-century Tamil tradition temple archi-tecture. Basically, the Vijayanagara columns have two parts. One is the core shaft and the other is the figural column.

The columns play a major role in Vijayanagara architecture bringing out the characteristic of Vijayanagara style. The columns are usually monolithic granite stone with a single base, sculpture/pillared shaft with a vallapoo capital (resem-bling banana flower). The columns are made complex by the addition of miniature columns on two sides and three sides to the core. Because of this character, the man-tapa looks filled with crop of columns. But the cloisters in the temple complex and bazaar colonnade are kept simple (Fig. 6).
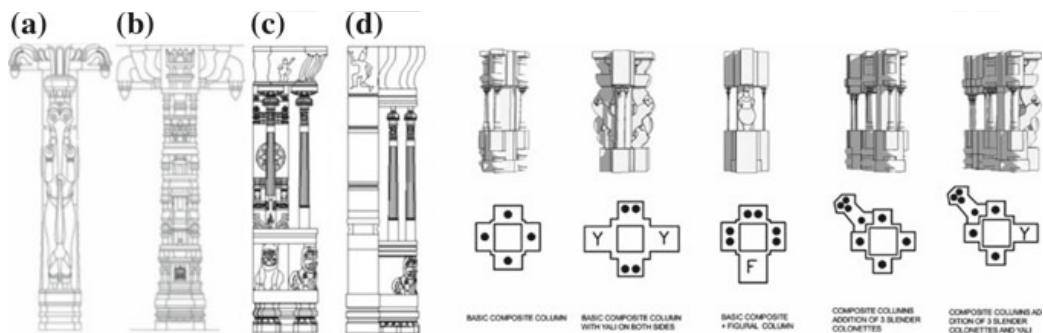

Fig. 5 Column types **a** Yali, **b** Sculpture, **c** Sculpture core and miniature **d** Core and miniature, additive grammar of composite columns

Sl. no	Name of the temple	Place/ Period of inception of the temple	Patron/ Zone	Chief deity	Entrances to temple complex	Location w.r.t Main shrine	No. of entrances to Mantapa	Shape of the plan	Nature of arrangement / Plan arrangement	Degree of enclosure
1	LAKSHMI NARASIMHA SWAMI (fig 1)	Kadhiri, Ananthapura district/ 1352 A.D	A Nayaka (Telugu zone)	Lord Narasimha, Avatar of Vishnu	4- N, S, E, W. Orientation of main shrine- E	SW	1 (from E)	Rectangular	free standing -Has 2 sections -Front part at lower level -Elevated part reached by a flight of steps guarded by elephants.	Pillared on all sides
2.	MALLIKARJUNA Temple (fig 2)	Mallavvakonda Hill Candragiri, Chittoor district/ In existence before 1405 AD. Presumed 1375 AD	Maha Navankacharya Dorappanavam Srinrarinav aka (Telugu zone)	Mallikarjuna form of Shiva	Open on all sides. Orientation of main shrine- E	NW	1 (from E)	50' x 30' rectangular	free standing -Has 2 sections -Front part at ground level -Elevated part reached by a flight of steps	Originally open -Pillared on all sides and walled on all sides excepting the front - Walled on 3 sides subsequently
3	RAMALINGESHWA RA TEMPLE (fig 3)	Tadpatri Anantapura district/ 1450 AD	Telugu zone	Swayambu, form of Shiva	2- N, S. Orientation of main shrine- W	SW	-	Rectangular	abutting temple enclosure -Has 2 sections -Front part at ground level -Hind part elevated	-
4	VITTALA TEMPLE (fig 4)	Hampi, Hospet/ Late 15" century and early 16" century	Tuluva period foundation (Kannada zone)	Vittala. form of Vishnu	3- N, S, E. Orientation of main shrine- E	SW	3 from N, W, E	Star shaped	free standing - Ornate basement with raised dias in the centre to house the Lord and his consort.	Pillared on all sides
5	VENKATRAMANA TEMPLE (fig 5)	TadpatriAnantapu ra district/ Earliest in scription 1551AD, maybe ascribed to about the first quarter of 16" century AD	Telugu zone	Form of Vishnu	3- N, S, E. Orientation of main shrine- E	SW	1 from E	Rectangular	abutting temple enclosure -has 2 sections -front part at ground level -Hind part elevated	Pillared on all sides

Sl. no	Name of the temple	No. of columns	Type of column	Type of capital	Kapota/ cornice	Roof edge decoration	Parapet of Mantapa	Ceiling of Mantapa
1	LAKSHMI NARASIMHA SWAMI (fig 1)	- Lower level: row of 6 pillars - Upper level: 2 rows of 6 pillars each	- 2 types of pillars: a) With pillared projection b) Figure at the base of the pillaret	Cola capital	-	Row of brick built arches with deity figures in stucco	-	-
2.	MALLIKARJUNA Temple (fig 2)	- Lower level: row of 4 pillars - Upper level: pillars with rectangular base, alternating hexagonal and square part.	- Lower pillars with fluted shafts - Upper level: pillars with rectangular base, alternating hexagonal and square part.	Cola capital with vertical and horizontal line | drawing on each arm of the capital.	Concave upper, convex lower portion above pillars, running around mantana	-	-	-
3	RAMALINGESHWA RA TEMPLE (fig 3)	- Lower level: 2 row of 3 tall pillars - Elevated section with 7 pillars at E, 2 in N edge and pavilion with 4 pillars in centre.	A main shaft and 8 pillarets each in 2 sections, 1 above the other, projecting from shaft.	Lower level pillars with roll and patta capitals	Cornice of double flexure	-	-	Pavilion ceiling is a lotus inside a square.
4	VITTALA TEMPLE (fig 4)	-Outer piers have pairs of rearing beasts in the middle of each side - cluster of colonetts at the corners	-	-	Cornice of double flexure	Corner of eaves have leaf like protuberances and stone rings	Brick parapet with pilaster framing stucco sculptures and miniature temple towers	Elaborate lotus in the middle of the ceiling
5	VENKATRAMANA TEMPLE (fig 5)	-Front selection: row of 13 pillars in the centre- in ruins now -UPPER LEVEL: row of 4 pillars on E edge ahnd 3 on S edge; 4 in the centre	-	-Pillars of the pavilion: Viiavanagara capital -2 pillars on E edge: roll and patta capital and other 2: Viiavanagara - N edge pillars: roll capital - Front section: second capital above the pillars containprojecting arms in the shape of an inverted lotus with a potika at the end.	-	-	-	-

Fig. 6 Comparative analysis of temple complexes [5]

5   Results

The 3D Virtual reconstruction is done for the temple complex in order to get deeper understanding of the monument in their original form. The visual reconstruction is done through study, documentation of the monument in ruins, and comparing them to similar monuments of this period also with the help of ancient texts of temple architecture, discussion with traditional craftsmen (Fig. 7).

Based on the study of proportions of existing building, all structures have been virtually reconstructed (Fig. 8). The sculpted details have been understood from literature study from ancient treatise on temple architecture. These 3D models help us to do visual tours and also visualize the cultural and ritual activities along with temple structure, giving a holistic picture of the Vijayanagara Empire in all its former glory.

5 1   Texture Mapping in Kinect

Contribution from IIT Delhi using Google SketchUp models: Texture mapping is a graphic design process in which a two-dimensional (2D) surface called a texture
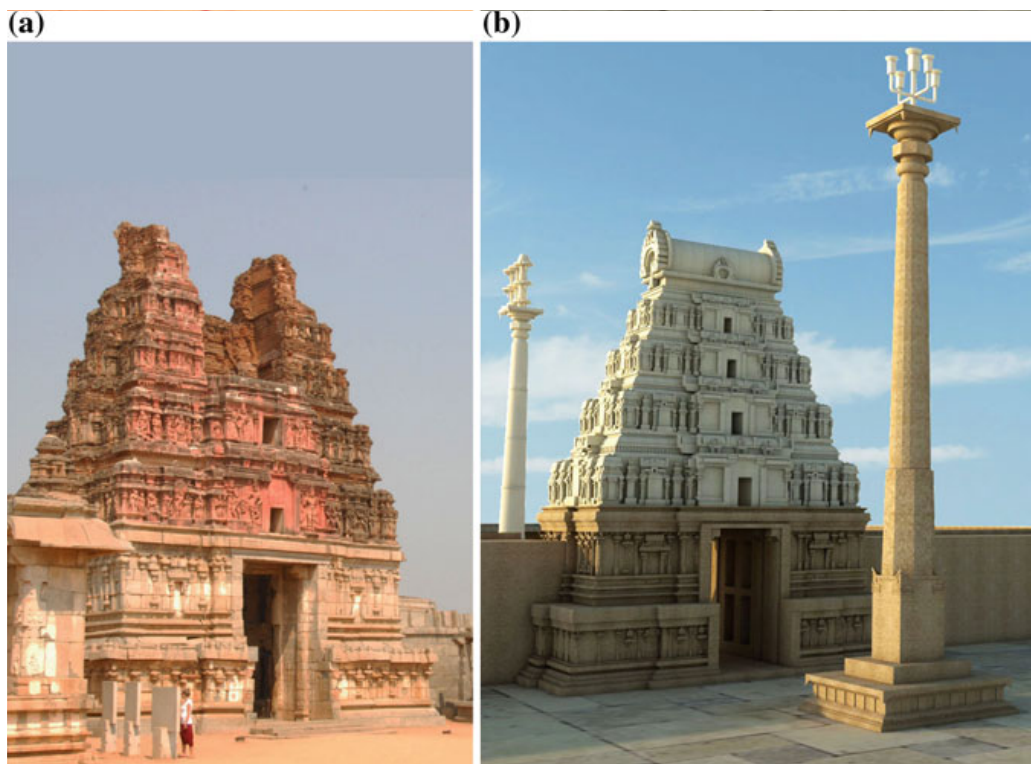

Fig. 7   **a** Photo image of existing east gopuram **b** Reconstructed view of east gopuram
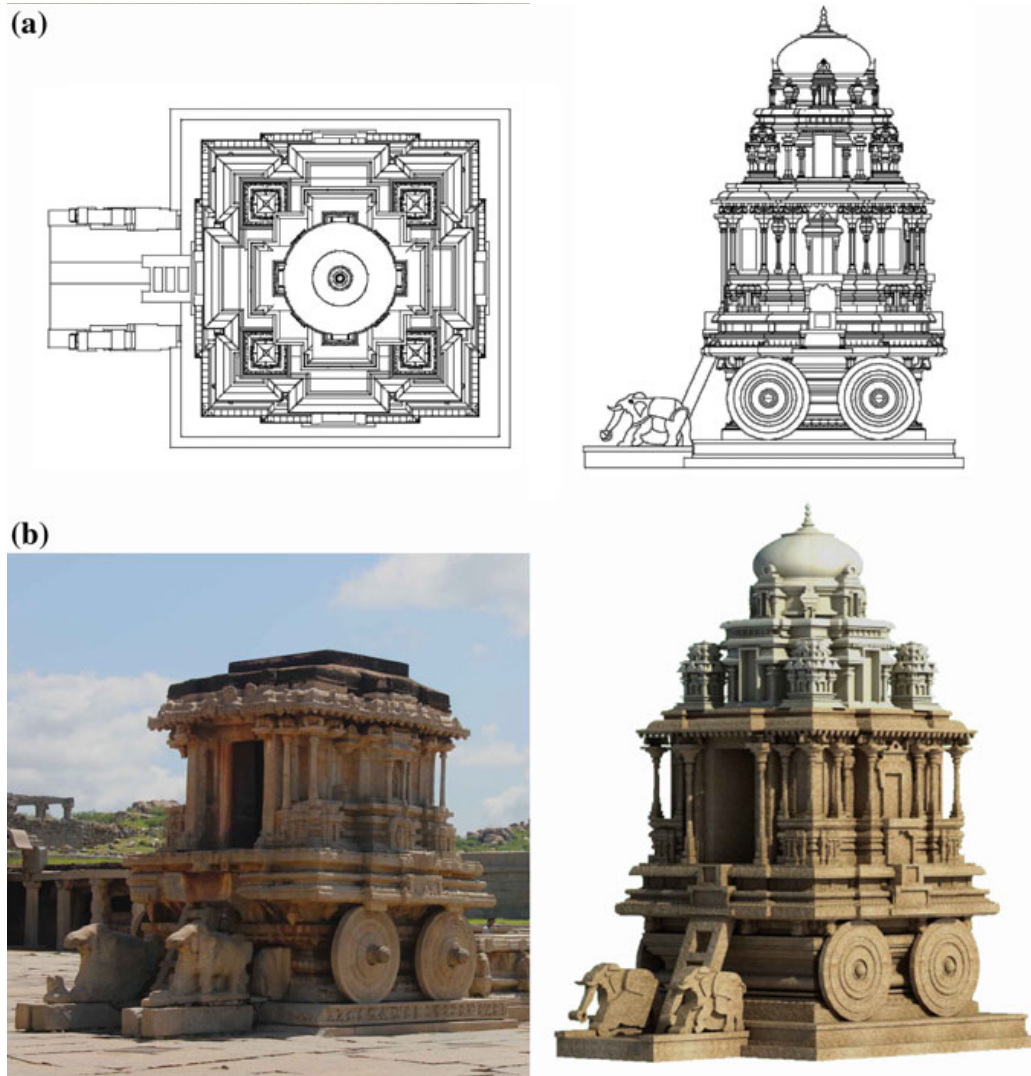
Fig. 8 **a** Google sketchup view of stone chariot, Vittala temple complex, **b** Reconstructed stone chariot, Vittala temple complex (an attempt)

map or image is mapped to a polygon or wrapped around a three-dimensional (3D) object. Thus, the object acquires a surface texture similar to that of the 2D image.

5 2 *Kinect Model: Course to Fine 3D Reconstruction*

The registration of the coarse and ne level 3D models is done in Autodesk 3ds Max software interactively. This is carried out by overlaying coarse 3D model on the ne 3D model. Figure 9 shows fine and coarse-level modes which need to be registered.

The corresponding points in the fine and coarse-level models are given in 3Ds Max. The figure shows the registered models. During the process of ne—level 3D
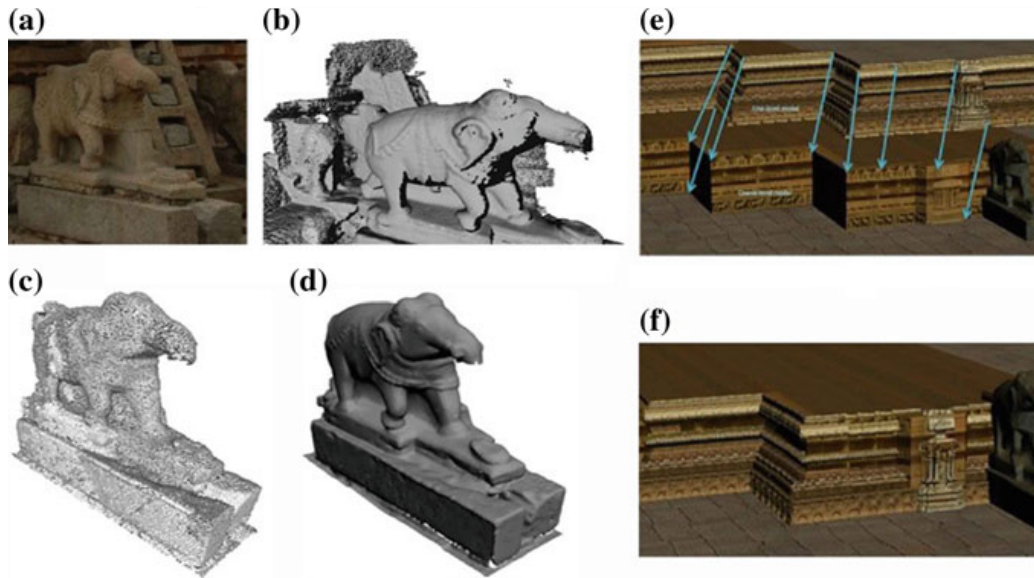
Fig. 9 **a** The elephant part of the chariot, **b** The point cloud, **c** Filtered and subsampled point cloud **d** Surface reconstructed 3D model. **e** The arrows show the corresponding points. **f** Registered ne and coarse-level models

reconstruction with Kinect sensor, it is not possible to reconstruct the entire monument at a time because of performance issues. Hence, the monument is reconstructed part by part and registered in 3Ds Max interactively. For further details and information on kinect model, refer to the journal article [6].

6 Conclusion

In the span of two centuries, Vijayanagara emperors have built thousands of temples. The temple complex is a complicated structure with di erent types of mantapas, variety of columns, evolved parapet details and high level of scale and proportion. Rituals and cultural activities have high influence on architecture. The characteristics of Vijayanagara style such as doubly curved sunshade, di erent types of columns, double plinth, and turrets are very evident.

The study and analysis have been concluded with 2D AutoCAD drawings and 3D Google SketchUp models of the visually reconstructed elements of the temple complex.

References

1. Michell G, Wagoner, PB (2001) Vijayanagar- architectural inventory of the sacred centre. Vol 1. Manohar Publishers, New Delhi, pp 217–229
2. Fritz, John and others (2006) Vijayanagara: Archeological Exploration, 1990–2000, Papers in memory of Channabasappa Patil S: Part 1. Manohar and American Institute of Indian Studies, New Delhi, p 15
3. Michell G, Wagoner PB (2001) Vijayanagara: Architectural Inventory of the Sacred Centre, Vol 1: Text and Maps. Manohar and American Institute of Indian Studies, New Delhi, p 3
4. Dallapiccola AL, Verghese A (2006). The Ramanuja Temple in Vithalapura Vijayanagara: Archeological Exploration, 1990–2000, Papers in memory of Channabasappa Patil S: Part 1. Manohar and American Institute of Indian Studies, New Delhi, pp 267–286
5. Kameswara Rao V, Select Vijayanagara Temples of Rayalaseema, Archeaological Series No 47 ; Michell G, The New Cambridge History of India: Architecture and Art of Southern India ; Michael, Meisterand W, Dhaky MA, Encyclopedia of Indian Temple Architecture: South Indian, lower Dravidadesa; Fritzand JM, Michell G, Hampi; CRISPIN BRANFOOT, Expanding Form: The Architectural Sculpture of the South Indian Temple
6. From report: 3D reconstruction and rendering using GPU, Prof. Uma Mudenagudi, Department of Computer Science and Engineering, B.V.B.C.E.T., Hubli

Acquisition Representation Processing and Display of Digital Heritage Sites

Prem Kalra Subodh Kumar and Subhashis Banerjee

1 Introduction

3D recreation of heritage sites with present technology o ers a powerful tool to communicate archaeological features and cultural knowledge not only to experts but also to the broad audience for the purpose of education and virtual tourism. With the advancements in the field of virtual reality over the years, it provides coupling of 3D digital recreation and visualisation with an e ective and immersive communication with the contained information. The 3D visualisation with interaction o ers additional degree of freedom beyond the mere presentation of static visualisations. It can allow real-time interaction with the environment and give users a sense of immersive experience. In this chapter, we present techniques and implementations of arriving at demonstrable interactive application that can be used in public spaces o ering such an experience with 3D recreations and reconstructions of a digital heritage site. The application is developed for the Vittala Temple Complex, within Hampi, a UNESCO-declared heritage site. It is an outcome of the project 'Acquisition, Representation, Processing and Display of Digital Heritage Sites'. The project has greatly benefitted from the interdisciplinary collaboration with other partner institutes of Indian Digital Heritage project especially BVBCET, NID, NIAS, IIT Bombay, IIT Madras, and IIIT Hyderabad and with IISc Bangalore.

P. Kalra ( ) S. Kumar S. Banerjee
Indian Institute of Technology Delhi, New Delhi, India
e-mail: pkalra@cse.iitd.ac.in

S. Kumar
e-mail: subodh@cse.iitd.ac.in

S. Banerjee
e-mail: suban@cse.iitd.ac.in

There have been attempts in numerous projects involving the modelling and rendering of digital cultural heritage in 3D virtual environments [44]. Some of the examples are reconstruction of ancient fresco paintings for a revival of life in ancient Pompeii [30], the reconstruction of Peranakans and their culture [32] and the reconstruction of nineteenth-century Koblenz (Germany) as a case study for a 4D navigable movie [23]. In [26] 3D, real-time virtual simulations of the populated ancient sites of Aspendos are done combining virtual reality and simulated dynamic virtual humans for an immersive experience. In this chapter, we give highlights of the research activity undertaken in the project. In brief, the chapter contains the following. We present techniques which have been developed through collaboration of research scholars, students and others. Part of these are published in different forums with corresponding references included in appropriate sections.

1. We give different modalities of acquisition of data which help in the processing of 3D recreation. The acquisition was done visiting the site several times with the help of other partner and collaborating institutes. Subsequently, we provide ways to combine different modalities for the purpose of recreation.
2. We give techniques of 3D reconstruction of existing parts of the scene at the site of interest. These techniques are based on the type of acquisition of the information—for example, one which relies on one image, some which are based on large number of images and use combination of acquisition from different sensors. It may be noted that the word reconstruction in our context is capturing or recreating 3D models as they exist, and not relating to restoring or reconstructing what it may have been originally. Some work on reconstruction and restoration of 3D models has been done by others which have also been included as necessary. We, however, have worked on some aspects of image restoration which we present briefly in a separate section.
3. One of the main outcomes from the project was to offer techniques and tools to design walk-through which are useful for performing virtual tours. To this end, we have developed tools on standard platforms such as Unity 3D. This provides real-time interaction an exploration of the recreated 3D models. In addition, we provide a unique interface mixing real (physical) models and virtual models to give a sense of immersion and a novel method of exploration with physical models augmented by virtual models.
4. In addition to using standard rendering and visualisation tools, we provide specialised rendering of artefacts which helps enhancing aspects relevant for artistic and non-photorealistic rendering.

The rest of the chapter is organised as follows. Section 2 includes brief discussion on acquisition. Section 3 provides techniques developed for 3D reconstruction. Section 4 reports the development of an interactive virtual tour system, and Sects. 5 and 6 present some efforts in the direction of artistic rendering and restoration of painting respectively. Section 7 concludes the chapter.
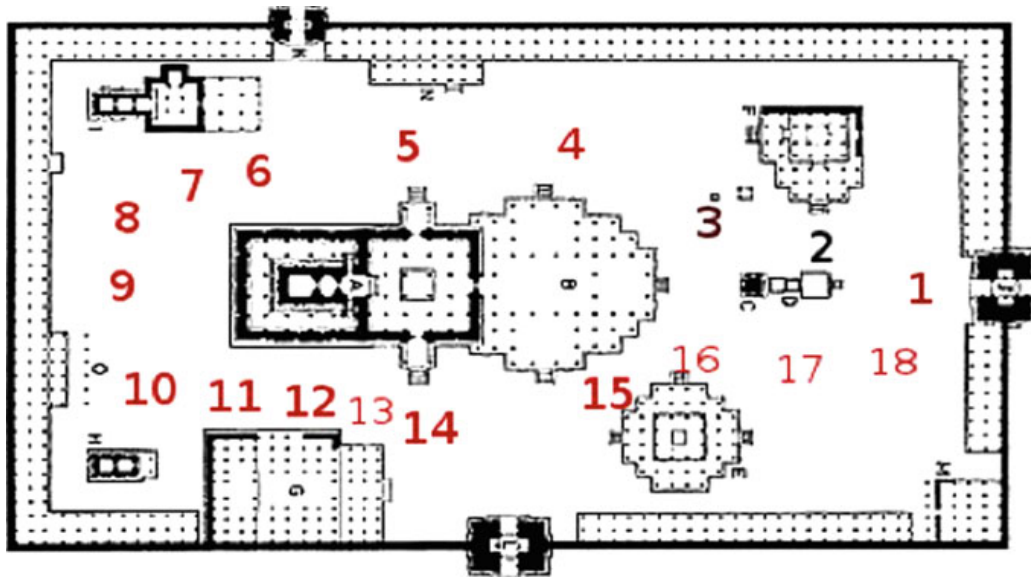
Fig. 1 The Vittala Temple Complex Plan with vantage points for the video walk-through

2 Acquisition

For acquisition, we use di erent modalities such as high definition cameras for acquiring images and videos, Kinect for obtaining simultaneous RGB and depth maps, and laser scanners. We provide algorithms for fusing the acquisitions from di erent modalities, which require alignment and registration of data obtained from multiple modalities. The fusion helps in object completion, and in some cases augmenting the resolution of acquisition. In the first visit to Hampi, about 2500 images were taken considering di erent vantage points, in addition images using Kinect and videos were also acquired. Figure 1 shows plan of Vittala Temple Complex. In subsequent visits, more imaging and video data was acquired considering the needs for the respective techniques for reconstruction.

3 3D Reconstruction

As mentioned earlier, with 3D reconstruction, we imply recreation of 3D geometric structure with surface details of artefacts as they are present. 3D reconstruction primarily relies on research in computer vision which is based on recovery of 3D structure from images. We present some of the research e orts done in this direction by the vision and graphics group at Department of Computer Science and Engineering IIT Delhi.

The final models combine 3D reconstructions obtained from a variety of techniques.

3.1   Single View Reconstruction [20]

Single-view reconstruction is carried out using orthogonal planes in the scene using a single image. A grammar-based approach is developed to obtain simultaneous reconstruction of all the planes collectively. The method also allows generation of surface of revolution and surface of translation. It is also possible to merge separately created structures from di erent (single) view reconstructions.

Single-view reconstruction has emerged as a useful technique for piecewise modelling of a scene with since Tour into the Picture [18]. The technique is particularly applicable to scenes with regular geometric structures consisting primarily of blocks and faces, as can be found in many heritage monuments in India and elsewhere. Subsequently, several systems for single-view reconstruction have been reported [11, 21, 24, 28, 43] which use user-provided constraints on vanishing points and lines, parallelism and orthogonality, coplanarity and incidence relationships to calibrate the camera and determine the scene geometry. Criminisi et al. [11] show that the vanishing line of a reference plane and a vanishing direction is the minimal information necessary for camera calibration and interactive affine measurements in the image. Sturm and Maybank [43] use coplanarity of points and lines, and parallelism and orthogonality of planes and directions to derive constraints for single-view reconstruction of sets of connected planes. The basic geometric techniques have also been extended for automatic reconstruction from single views in some situations [17]. Incremental computation of single-view reconstruction from user-provided constraints often results in error accumulation during the sequential computation steps. We address the problem of first capturing all user-provided constraints through a graphical user interface in a symbolic representation and then compiling the symbolic information to derive a set of equations which can be solved in one go in a least square sense [20]. One shot computation of a tightly coupled system ensures robustness and prevents error accumulation. We derive a set of constraints that can be captured symbolically using a simple data structure and programming semantic and describe a method for subsequent compilation of the symbolic constraints in to a set of equations. As an added benefit, we can also process the symbolically captured information to determine whether the constraints provided are adequate for the complete reconstruction. For details, readers are encouraged to refer to Khurana et al. [20]. The tool has also been used by other institutions. Figure 2 shows result of a 3D reconstruction using single view of the frontal part of the Kalyan Mandapa.

3.2   Multi view Large Scale Reconstruction [5]

Given many images, structure from motion (SfM) solves for both the 3D scene being viewed and the cameras parameters. Many large-scale SfM methods [1, 9, 40, 41, 47] use the bundle adjustment method [45] which simultaneously optimises for both structure and camera parameters using point correspondences in images by

Fig. 2 Single-view reconstruction of the frontal part of the Kalyan Mandapa (courtesy BVBCET)

minimising a global cost function. This, however, fails for a large data of images due to an accumulation of error in an incremental reconstruction or when cameras are weakly connected to 3D feature points. In addition, owing to the very large number of variables involved, bundle adjustment is also very computationally demanding and time consuming. As a collaborative work, we adopt a divide-and-conquer strategy that is designed to mitigate these problems (see Bhowmick et al. [5]). In essence, the approach partitions the full image data set into smaller sets that can each be independently reconstructed using a standard approach to bundle adjustment. Subsequently, by utilising available geometric relationships between cameras across the individual partitions, it solves a global registration problem that correctly and accurately places each individual 3D reconstructed component into a single global frame of reference.

Using an image match graph based on matching features, we partition the image data set into smaller sets or components which are reconstructed independently. Following such reconstructions, we utilise the available epipolar relationships that connect images across components to correctly align the individual reconstructions in a global frame of reference. This results in both a significant speed up of at least one order of magnitude and also mitigates the problems of reconstruction failures with a marginal loss in accuracy. The e ectiveness of our approach is demonstrated on some large-scale real data sets of Vittala Temple Complex in Hampi.

In what follows, we show that this approach is not only more robust with respect to failures in reconstruction but also gives significant improvements over the state-of-the-art techniques in terms of computational speed. The main contributions are as follows:

1. A principled method based on normalised cuts [37] to partition the match graph of a large collection of images into disjoint connected components which can be independently and reliably reconstructed. This process also automatically identifies a set of connecting images between the components which can be used to register the independent reconstructions. Specifically, these are the image pairs specified by the cut edges in the graph.
2. A method for registering the point clouds corresponding to the independent connected components using pairwise epipolar geometry relationships. The epipolar-based registration technique proposed in this paper is more robust than the standard techniques for registering point clouds using 3D–3D or 3D–2D correspondences. Registration methods based on 3D point correspondences do not use all available information (image correspondences) and may fail when the point clouds do not have sufficient number of 3D points in common. 3D–2D based methods, such as a sequential bundler [1, 47, 48], often result in broken reconstructions when the number of points available is inadequate for resectioning or when common 3D points are removed at the outlier rejection stage [47]. The proposed registration algorithm using pairwise epipolar geometry alleviates this problem as is shown in Fig. 3. Considered as an independent approach, the epipolar-based algorithm can also be used to register independently reconstructed point clouds by introducing a few connecting images.

Matching all pairs of images in an iterative bundler is computationally expensive, especially when the number of images in the collection is large. There have been several attempts to reduce the number of images to be matched. Frahm et al. [14, 33] try to find some representative 'iconic images' from the image data set and then partition

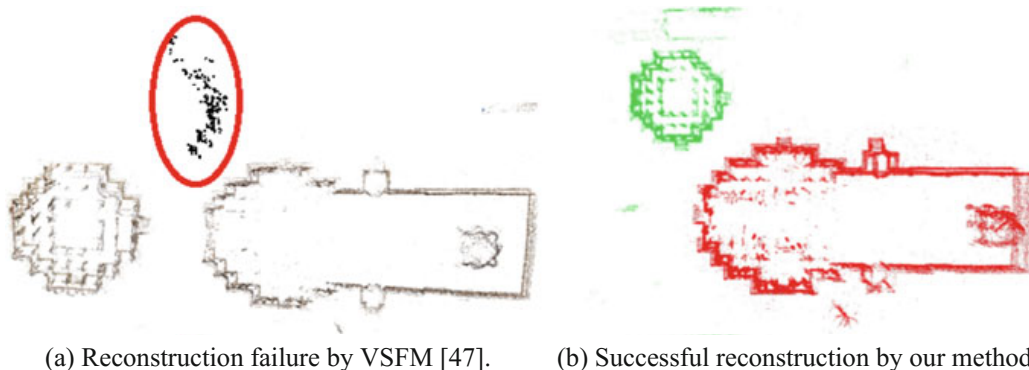(a) Reconstruction failure by VSFM [47].　　(b) Successful reconstruction by our method.

Fig. 3 Plan view of reconstruction of two temples at the Hampi site in India: **a** illustrates the failure of VSFM [47] due to inadequate points during resectioning (circled in red) whereas **b** our approach correctly solves the reconstruction problem. Please view this figure in colour

the iconic scene graph, reconstruct each cluster and register them using 3D similarity transformations. Snavely et al. [2, 42] and Havlena et al. [16] compute skeletal sets from the match graph to reduce image matching. All these methods reduce the set of images on which they run SfM. Moreover, incremental bundle adjustment is also known to su er from drift due to accumulation of errors which increase as the number of images increase [9, 10, 47]. Crandall et al. [9, 10] propose an MRF-based discrete formulation coupled with continuous Levenberg–Marquardt refinement for large-scale SfM to mitigate this problem. To reduce the matching time, Wu [47] (henceforth VSFM) proposed preemptive matching to reduce the number of pairs to be matched. Moreover, all cameras and 3D points are optimised only after a certain number of new cameras are incorporated into the iterative bundler. Although VSFM demonstrates approximately linear running time, sometimes it fails for large data sets when the accumulated errors of iterative bundler become large [47]. Although there have been some recent global methods [19, 27], to be able to solve large-scale SfM problems, global methods need to be exceedingly robust. Farenzena et al. [13] also propose to merge smaller reconstructions in a bottom-up dendrogram. However, their largest datasets are of only 380 images and their use of reprojection errors of common 3D points for merging is unsuitable for very large datasets. In our approach, we propose to decompose the image set into smaller components so that the match graph of each component is densely connected. This is likely to yield correct 3D reconstructions, since fewer problems are encountered during the resectioning stage of a standard iterative bundler and the reconstruction is robust. Restricting pairwise image matching to within each component also yields a significant reduction in computation time. Moreover, SfM-based reconstruction of each component can be carried out in parallel. Our approach is conceptually depicted in Fig. 4.

Images used for bundle adjustment can either be acquired from a site or aggregated from various sources on the Internet. When the images are acquired from a site in an organised manner, the problem of decomposition into smaller sets becomes simpler. In what follows, we provide an illustration. Figure 5 shows the Google Earth
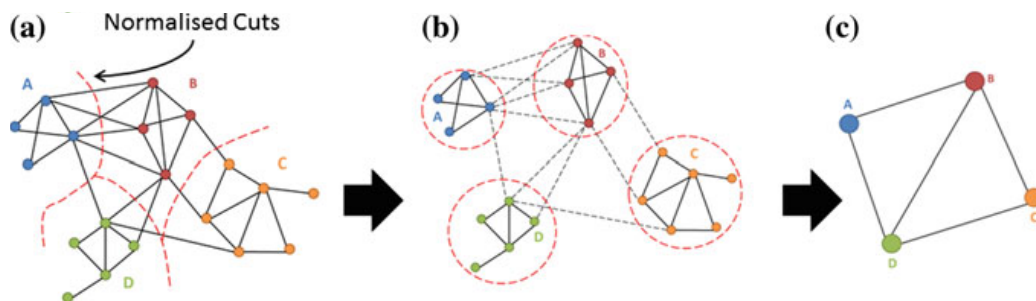

Fig. 4 **a** Original match graph where images (nodes) are connected by edges having similar image features. The edge weights represent similarity scores. **b** Normalised cut partitions the full image set into connected components which can be reconstructed independently. The 'connecting images' across components are defined by the cut edges. **c** The individual cuts are now equivalent to individual nodes that represent independent rigid 3D reconstructions which are registered using pairwise epipolar relationship of the connecting images
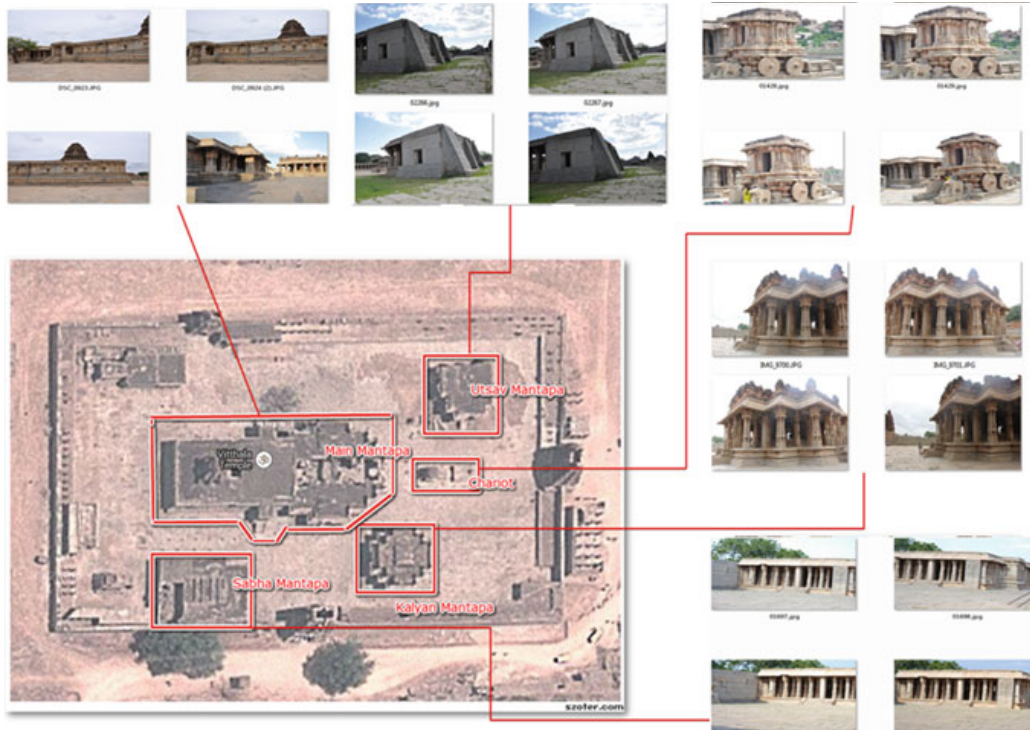
Fig. 5 Google Earth view of the Vittala temple, Hampi, Karnataka, India. The red boxes denote different buildings of the temple. Images for each building were captured separately

view of the Vittala Temple. Figure 6 shows a typical example where images of two buildings are captured separately and it also shows a typical *connecting* image which has parts of both the buildings. We call such data sets *organised*.

In case such planned acquisition is not possible, the collection of images needs to be automatically partitioned into smaller components. Unorganised data sets downloaded from the Internet are typical examples. In such cases, a method for automatically grouping into visually similar sets and finding connecting images needs to be established. To this end, we train a vocabulary tree [29] using all image features (SIFTGPU [46]) and extract top p (typically p = 80) similar images for each image in the set. We form a match graph where each node is an image and the edge weights between two nodes are the similarity values obtained from the vocabulary tree. We aim to partition the set of images such that each partition is densely connected. The partitions only capture dense connectivity of matched image features and need not represent a single physical structure. Here, the dense connectivity ensures that SfM reconstruction is less likely to fail due to the paucity of reliable matches or accumulated error or drift.

We use the multiway extension [37] of the normalised cut (NC) formulation to automatically partition the match graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ into individual clusters. Since, in our case, edge weights are based on visual similarity, the normalised cut would yield those connected components in which connected images are visibly similar. We use the images that belong to the cut as candidate connecting images. In Fig. 7, we show
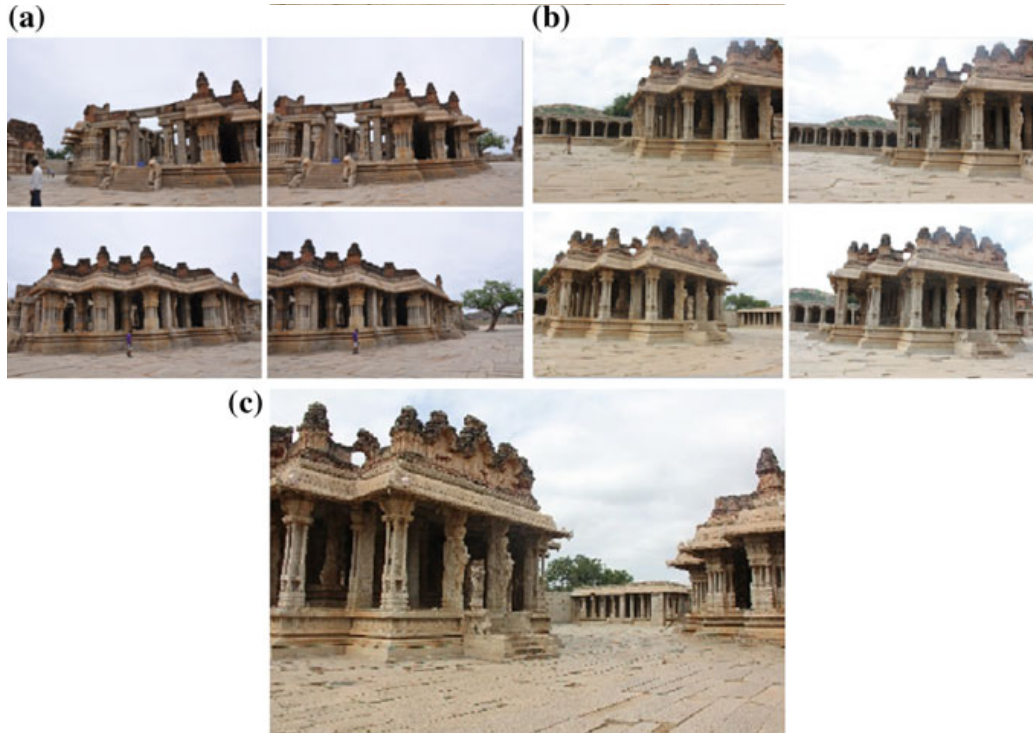
Fig. 6 **a**, **b** Two buildings of the Hampi temple complex, and **c** a typical connecting image

the result of our estimation upon applying the normalised cut to the set of images collected at the Hampi site illustrated in Fig. 5, i.e. when we treat the images as an unorganised dataset. Figure 7a shows the cameras partitioned into connected components in different colours. Figure 7b shows the plan view of the 3D reconstructions obtained for each component marked in corresponding colours. It should be noted that in this example, the graph weights are based only on pairwise image feature similarity scores. We can improve the quality of the graph by incorporating geometric information such as the robustness of computation of pairwise epipolar geometries of connected images. Such a scheme would not only ensure that the connected pairs of images can be reliably matched but would also ensure that the pairwise epipolar geometries can be robustly estimated.

The number of candidate connecting images is often very large. Reducing the number of connecting images will reduce the time for estimation of pairwise epipolar geometry. The connecting image extraction process takes this into account [5]. Further, an approach is adopted to register the individually reconstructed groups of cameras to a single frame of reference. To register a pair of 3D reconstructions, we estimate the relative transformation between them, i.e. we estimate relative rotation, translation and scale between a pair of reconstructions using epipolar relationships between the reconstructed cameras and the connecting cameras. If A and B are two components and if R_{Al}, and C_{Al} be the rotation and position of camera l in the local coordinate system of A, and R_{Bk}, C_{Bk} be the rotation and position of camera k in the local coordinate system of B, and R_{lk}, t_{lk} be the pairwise rotation and translation

(a) Cameras partitioned into connected compo-
nents. Each component is shown in a different
colour.

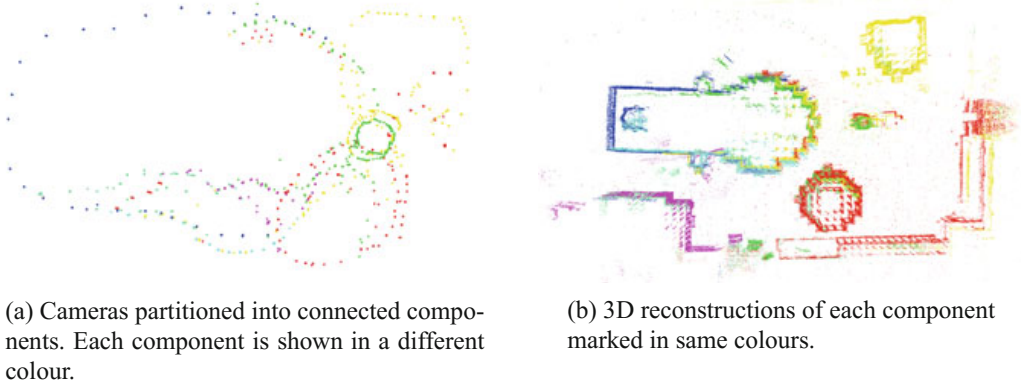(b) 3D reconstructions of each component
marked in same colours.

Fig. 7 SfM results on the Hampi dataset (unorganised data) illustrating the e ect of graph
partitioning

between the camera l and k obtained from the epipolar relationship, then the rotation
and translation between A and B (R_{AB}, T_{AB}) are computed as

$$R_{AB} = R_B R_A^T = R_{Bk}^T R_{lk} R_{Al}$$
$$T_{AB}^k = C_{Bk} - R_{AB} C_{Ak}$$

Scale can be obtained from the relationship $t_{lk} \quad R_{Ak}C_{Al} - R_{Ak}C_{Ak} = 0$. In our
implementation, we start the process of global registration using the largest recon-
struction (with maximum number of images) as the seed and register all other recon-
structions which are connected to this seed and merge them into a single model. We
also remark that the motion models required for registering individual reconstruc-
tions connected to the current model can be estimated in parallel.

For our experiments, we used an Intel i7 quad core machine with 16GB RAM
and GTX 580 graphics card. We first present our result on an *organised* image set
acquired from Vittala Temple Complex (see Fig. 5). The data set consists of 2337
images covering 4 temple buildings. The physical footprint of these four buildings
covers an area of approximately 160×94 m. For reconstructing the images in each
individual set, we use VSFM [47] as the iterative bundler. We merge each of these
reconstructions into a common frame of reference. Figure 8a shows our reconstruc-
tion after registration superimposed on a view from Google Earth. As we do not
have ground truth for such real-world data to analyse the quality of our reconstruc-
tion, we use the output of VSFM applied on the entire data set using all pairs match-
ing as our baseline reconstruction. We note that all pairs matching is necessitated
here as the scheme of preemptive matching suggested in [47] fails on this data set.
Figure 8b shows the comparison where the red point cloud is obtained from VSFM
and the green points are obtained using our method. VSFM took 5760 min to recon-
struct the data set using all pairs matching. In contrast, our method takes 2578 min
(using all pairs matching) to reconstruct the same data set. The computation time
of our method is calculated by considering the time required for reconstruction of
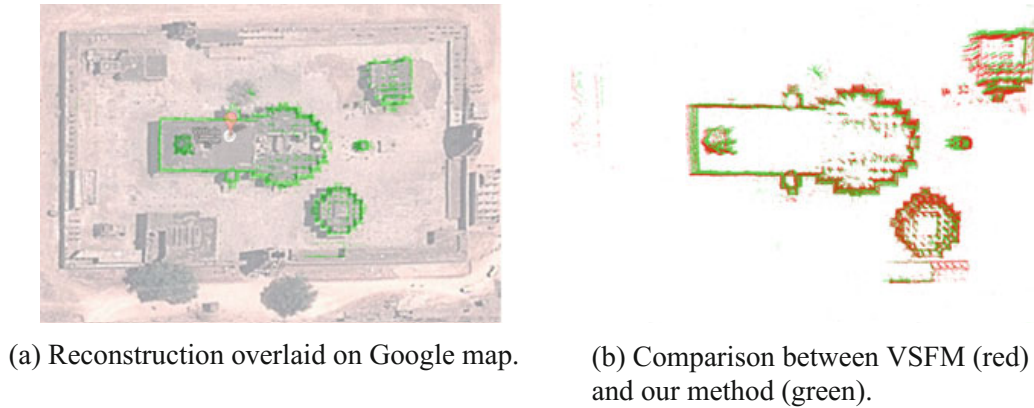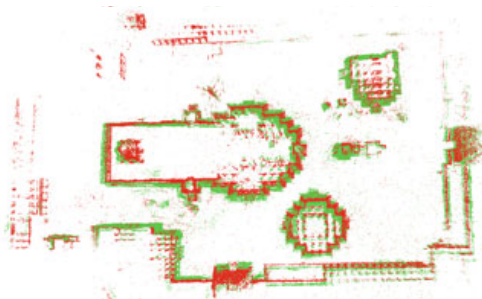
(a) Reconstruction overlaid on Google map.   (b) Comparison between VSFM (red) and our method (green).

Fig. 8 Validation of reconstruction of Vittala Temple Complex data set (organised data)

the largest component and the total time for registration, since the reconstruction of each component is done in parallel. We also compare the 3D camera rotations and positions (i.e. translations) obtained by our method against the 'ground truth' provided by VSFM. As the two camera estimates are in diﬀerent frames of reference and may also diﬀer in scale, we align them in a common Euclidean reference frame by computing the best similarity (Euclidean transformation and a global scale) transformation between them. Apart from being much faster than VSFM, our result is qualitatively better or similar to that obtained by VSFM.
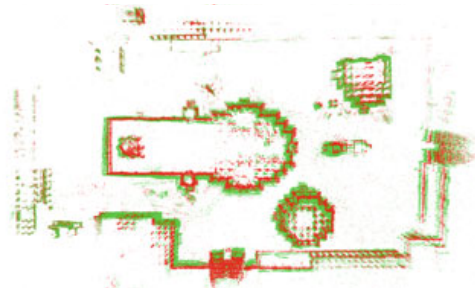
For experimenting with unorganised image datasets, we consider a total of 3017 images from the data set. We train a vocabulary tree [29] using SIFT [25, 46] features and take 80 most similar images from vocabulary tree for each image in the set to construct a match graph. Normalised cut is applied on this match graph and connected components are obtained. In our experiments, the expected number of connected components is decided intuitively and is used as an input parameter for the number of components needed using normalised cut. We find the connecting images and then run VSFM on individual connected components and merge them into a single coordinate frame. Figure 9 shows a frontal view of the reconstruction by our method. Figure 7b shows the 3D reconstructions corresponding to each of the connected components registered and in diﬀerent colours. To validate our result, we overlay our reconstruction on the corresponding site map from Google Earth and Fig. 10c shows that the registration is accurate. We also run VSFM with all 3017 images and compare the results. Figure 10a shows the comparison results where the VSFM output is marked in red and the output obtained using our method is marked in green. Figure 10b shows the corresponding results using a measure of robustness of epipolar estimation as edge weights in normalised cut.

Fig. 9 Frontal view of the Hampi reconstruction (considered as unorganised data)

(a) Comparison with VSFM (red) and our method (green).

(b) Comparison with VSFM (red) and our method with epipolar robustness (green).

(c) Overlaid on Google map.

Fig. 10 Reconstruction of the Hampi data set (considered as unorganised data) validated against VSFM reconstruction and Google Earth

3.3 Multiple View Dense Reconstruction

For some parts, we have used available toolkit of CVMS (Clustering Views for Multi-view Stereo) [15] from Yasutaka Furukawa. However, this has been used primarily by the partner institute BVBCET and has been detailed in another chapter of this book. That chapter also discusses the techniques involving acquisitions from depth images of Kinect for the purpose of creating 3D models. We also combine the methods of dense reconstruction from inputs from low-resolution sensor, e.g. Kinect with high-resolution images from high definition cameras, which we present in the next section.

3.4 Multimodal Reconstruction [31]

We present a methodology which combines high definition (HD) resolution images with the low-resolution Kinect to produce high-resolution dense point cloud using graph cut [31]. First, Kinect and HD cameras are registered to transfer Kinect point cloud to the HD camera for obtaining high-resolution point cloud space. Then, we discretise the point cloud in voxel space and formulate a graph cut formulation which takes care of the neighbour smoothness factor. A dense 3D reconstruction is carried out in the registered frame using two basic criteria: (i) photo-consistency [22], and (ii) rough agreement with Kinect. The reconstructed point cloud is at least ten times denser in comparison to the initial point cloud. In this process, we also fill up the holes of the initial Kinect point cloud. This methodology produces a good high-resolution image with the help of low-resolution Kinect point cloud which could be useful in building a high-resolution model using Kinect.

There has been considerable work with time-of-flight (ToF) cameras which capture depth scans of the scene by measuring the travel time of an IR wave emitted from the device towards the object and reflected back from the object [34]. A much cheaper range sensor has been introduced by Microsoft called the Kinect [39] which has an inbuilt camera, an IR emitter and a receiver. The emitter projects a predetermined pattern whose reflection o the object provides the depth cues for 3D reconstruction. Kinect produces range data in VGA resolution, this data can be very useful as an initial estimate for subsequent resolution enhancement. There have been several approaches to enhance the resolution of a point cloud obtained from range scanners or ToF cameras, using interpolation or graph-based techniques [34, 35]. Diebel et.al. [12] used a MRF-based approach whose basic assumption is that depth discontinuities in scene often co-occur with intensity or brightness changes in the scene, or in other words regions of similar intensity in a neighbourhood have similar depth. Yang et.al. [49] make the same assumption and use a bilateral filter to enhance the resolution in depth. However, the assumption is not universally true and may result in over smoothing of the solution.

Normalised cross correlation (NCC) method, which tries to find point correspondences in an image pair by computing statistical correlation between the window centred at the candidate point, is an inadequate tool for finding dense point correspondences. Projecting the sparse Kinect point on to an HD image leaves most pixels without depth labels, and one can attempt to establish correspondence for these pixels using normalised cross correlation along rectified epipolar lines. Once the correspondence is found, we can obtain the 3D point for this correspondent pair using stereo triangulation technique. Using NCC, the reconstruction may give rise to many holes due to ambiguous cross correlation results and incorrect depth labels.

The voxel labelling problem can be represented as one of minimising an energy function of the form

$$E(L) = \sum_{p \in \mathscr{P}} D_p(L_p) + \sum_{(p,q) \in} V_{p,q}(L_p, Lq)$$

where $\mathscr{P}$ is the set of voxels to be labelled, $L = \{L_p | p \in \mathscr{P}\}$ is a 0–1 labelling of the voxel P, $D_p(.)$ is data term measuring the consistency of the label assignment with the available data, defines a neighbourhood system for the voxel space and each $V_{p,q}(.)$ is a smoothness term that measures the consistency of labelling at neighbouring voxels.

When the above energy minimisation problem is represented in graphical form [6], we get a two terminal graph with one source and one sink nodes representing the two possible labels for each voxel. Each voxel is represented as a node in the graph and each node is connected to both source and sink nodes with edge weights defined according to the data term of the energy function. In addition, the voxel nodes are also connected to each other with edges, with edge strengths defined according to the neighbourhood interaction term. A minimum cut through this graph gives us a minimum energy of the configuration. Photo-consistency [22] is one of the most frequently used measures for inter-image consistency. However, in real situations, several voxels in a close neighbourhood in depth satisfy the photo-consistency constraint resulting in a thick surface. In view of this, we use closeness to initial Kinect data as an additional measure to resolve this problem of thickness in the output high-resolution point cloud.

We define the data term based on two criteria: adaptive photo-consistency measure for each voxel, and distance of each voxel from its approximate surface. We use the photo-consistency measure suggested by Slabaugh et. al. [38]. We project each voxel i on to the N HD images and calculate the following two measures:

1. $S(i)$, the standard deviation of the intensity values in the projection neighbourhoods calculated over all N images.
2. $s(i)$, the average of the standard deviation in the projection neighbourhoods for each image projection.

The voxel i is photo-consistent over the N images if the following condition is satisfied
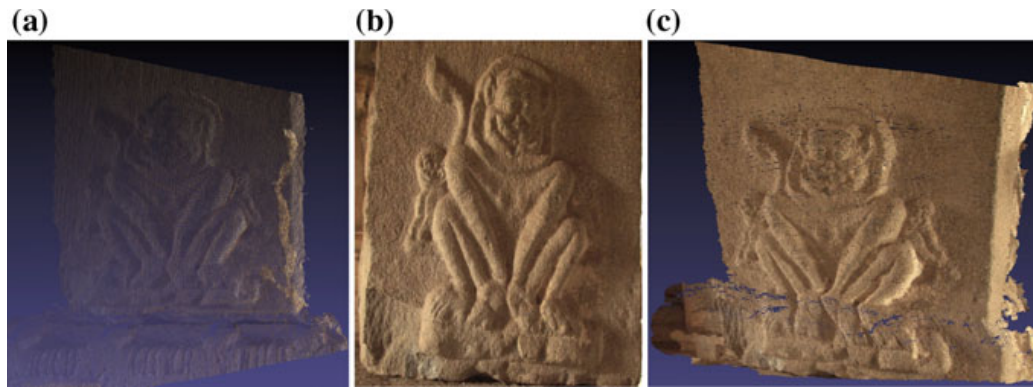
$$S(i) < \tau_1 + \tau_2 * s(i)$$

Fig. 11 A sculpture depicting a monkey on a pillar. **a** Initial low-resolution point cloud from Kinect, **b**, **c** front and side view of the high-resolution point cloud generated by our method

where τ_1 and τ_2 are global and local thresholds to be suitably defined depending on the scene. The overall threshold specified by the right-hand side of the above inequality changes adaptively for each voxel.

Figure 11 shows the result of resolution enhancement on a sculpture in an outdoor scene in the Vittala Temple Complex (Hampi) done using one Kinect and two HD cameras. The Kinect point cloud was taken in low light during early morning so that IR from sunlight does not a ect the depth map. Here also, the point cloud was at least 10 times denser than the initial point cloud. Here, the number of voxels taken was $300 \times 300 \times 100$ and the values of τ_1 and τ_2 were chosen to be 80 and 0.5, respectively. This is because of the fact that the texture was coarser and hence the threshold value has to be larger in order to account for all photo-consistent voxels. For more details, readers are encouraged to refer to the paper [31].

4 Interactive Walk-through and User Experience [7 8]

We provide a tool to design guided walk-throughs in terms of navigation path, gaze direction and speed. The approach that we use is optimisation based on constraints on importance or visibility. The features of the walk-through that we take into consideration are path planning, gaze direction and speed control, multilevel rendering and multimodal rendering.

In addition, we have designed and developed a virtual tour system for the Vittala Temple, which can also be used for other heritage sites. Our system uses a unique technique to blend physical models with virtual models. This technique is then used to create a simple user interface which o ers a high degree of immersion without compromising on the richness of experience. In our system, we have created a tabletop replica of the entire temple complex by employing 3D printing technology. Two consumer grade LCD projectors are then used to simultaneously project two

Fig. 12 The interface using mixed mode of models

di erent views of the temple structure on a large screen and over the 3D-printed objects.

The user interaction device to control this system is a single light emitting device, like a laser pointer or a flashlight. The user shines a path on the 3D temple structure to create a virtual tour, with the walk-through being displayed on the secondary screen. During the walk, the user changes his viewing direction and elevation using an intuitive guide handle. Additionally, the user can also shine the light on a monument to learn more about it through audiovisual content shown on the secondary screen. We have verified our system based on various parameters such as usability, scalability, robustness and cost. The usability of our system is verified by conducting a user survey while robustness is tested by independently running it in the real world under harsh conditions. The system cost is compared with other commercially available CAVE-based systems and its scalability is demonstrated by using it for a di erent layout (a layout other than Vittala Temple). Figure 12 shows the user interaction interface using both the 3D printed models and virtual models.

Our main contribution in this work is the design, development and evaluation of an innovative virtual tour system which is simple, cheap and user-friendly. It is also scalable in nature and thus it can be used in a variety of other walk-through scenarios.

5 Artistic Rendering

Rendering techniques for architectural illustrations have been developed using both ambient occlusion rendering and di use lighting [36]. This enables us to depict the

features of the 3D architectural objects with high emphasis on the important features. We develop methods to extract prominent features from artefacts of archaeological importance at the expense of other less prominent details using non-photorealistic rendering techniques. We apply various feature defining techniques like ridges, valleys, suggestive contours, etc., individually to show various areas of prominence, and show features with the help of line drawing, shading, etc.

6 Restoration of Paintings

In one of the methods for restoration of old paintings, we use interleaving of inpainting and denoising techniques [3]. We also develop methods for changing the appearance of an old painting to that of a chemically cleaned painting by using a sample clean painting that has the same colour distribution as the old painting [4]. To be able to do so, we move data points of old painting by scaling, rotation and translation to fill data points' cluster of the sample cleaned image in RGB colour space. Image inpainting and restoration provide image correction in terms of completion, colour correction, crack detection and filling. The approach that we use is an extension of exemplar-based method and colour space transformation. Figure 13 shows one example of inpainting performed on an image. We have dealt with this problem in a limited sense and used the inpainted images for texture mapping of the 3D restored parts. A separate chapter provides a more comprehensive treatment to this problem.

Fig. 13 The inpainting result, the left is the original image and right is the inpainted image

7 Conclusion

This chapter presents some of our research e orts towards the project. They cover di erent aspects related to—acquiring the data in the form images, videos, depth maps and laser scans, representation and recreation of 3D virtual models of various mandapas of Vittala Temple Complex, display and exploration of these models through walk-through and real-time interaction, and processing the data for the purpose of specialised rendering and restoration. Considering the limited space, the chapter has provided an overview of these e orts. However, many of these are available as separate publications which are included in the reference where readers can get more information.

References

1. Agarwal S, Snavely N, Seitz S, Szeliski R (2010) Bundle adjustment in the large. In: Proceedings of the European conference on computer vision, pp 29–42
2. Agarwal S, Snavely N, Simon I, Seitz S, Szeliski R (2009) Building rome in a day. In: Proceedings of the international conference on computer vision, pp 72–79
3. Arora N, Kalra P (2011) Interactive image restoration using inpainting and denoising. In: NCVPRIPG 2011—Proceedings of national conference on computer vision, pattern recognition, image processing and graphics, Hubli, 15–17 December, 2011, pp 219–222
4. Arora N, Kumar A, Kalra P (2013) Digital restoration of old paintings. In: WSCG 2013—Proceedings of conference on computer graphics, visualization and computer vision plzen, Czech Republic 24–27 June, 2013, p F89
5. Bhowmick B, Patra S, Chaterjee A, Govindu VM, Banerjee S (2014) Divide and conquer: efficient large scalre structure from motion using graph partitioning. Proc ACCV 2014:273–287
6. Boykov Y, Kolmogorov V (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. In: IEEE transactions on PAMI, vol 26, no 9, September 2004, pp 1124–1137
7. Bugalia N (2016) Immersive environment system for an efficient human computer interaction. MSR thesis, Amar Nath and Shashi Khosla School of Information Technology, IIT Delhi
8. Bugalia N, Kumar S, Kalra P, Choudhary S (2016) Mixed reality based interaction system for digital heritage. In: Proceedings of the 15th ACM SIGGRAPH conference on virtual-reality

continuum and its applications in industry, VRCAI 2016, Zhuhai, China, December 3–4, 2016, pp 31–37

9. Crandall DJ, Owens A, Snavely N, Huttenlocher DP (2011) Discrete-continuous optimization for large-scale structure from motion. In: Proceedings of IEEE conference on computer vision and pattern recognition, pp 3001–3008

10. Crandall DJ, Owens A, Snavely N, Huttenlocher DP (2013) SfM with MRFs: discrete-continuous optimization for large-scale reconstruction. IEEE Trans Pattern Anal Mach Intell 35(12):2841–2853

11. Criminisi A, Reid I, Zisserman A (2000) Single view metrology. Int J Comput Vis 40(2): 123–148

12. Diebel J, Thrun S (2006) An application of markov random fields to range sensing. In: Advances in neural information processing systems, pp 291–298

13. Farenzena M, Fusiello A, Gherardi R (2009) Structure-and-motion pipeline on a hierarchical cluster tree. In: Proceedings of IEEE international conference on computer vision workshop on 3-D digital imaging and modeling, pp 1489–1496

14. Frahm J, Georgel P, Gallup D, Johnson T, Raguram R, Wu C, Jen Y, Dunn E, Clipp B, Lazebnik S, Pollefeys M (2010) Building Rome on a cloudless day. In: Proceedings of the European conference on computer vision: Part IV, pp 368–381

15. Furukawa Y (2016) Clustering views for multi-view stereo. http://www.di.ens.fr/cmvs. Accessed 26 July 26 2016

16. Havlena M, Torii A, Pajdla T (2010) Efficient structure from motion by graph optimization. In: Proceedings of the European conference on computer vision, Lecture notes in computer science, vol 6312, pp 100–113

17. Hoiem D, Efros AA, Hebert M (2005) Automatic photo pop-up. Proc SIGGRAPH 2005: 577–584

18. Horry Y, Anjyo KI, Arai K (1997) Tour into the picture: using a spidery mesh interface to make animation from a single image. Proc SIGGRAPH 97:225–232

19. Jiang N, Cui Z, Tan P (2013) A global linear method for camera pose registration. In: Proceedings of IEEE international conference on computer vision, pp 481–488

20. Khurana D, Sankhla S, Shukla A, Varshney R, Kalra P, Banerjee S (2012) A grammar-based gui for single view reconstruction. In: Proceedings of ICVGIP 2012, p 14

21. Kushal A, Chanda G, Srivastava K, Gupta M, Sanyal S, Sriram T, Kalra P, Banerjee S (2003) Multilevel modelling and rendering of architectural scenes. In: Proceedings of EUROGRAPH-ICS 2003 short presentations

22. Kutulakos K, Seitz S (1999) A theory of shape by space carving. In: 7th IEEE international conference on computer vision (ICCV-99), vol I, pp 307–314

23. Laycock RG, Drinkwater D, Day AM (2008) Exploring cultural heritage sites through space and time. J Comput Cult Heritage 1(2):1–15

24. Lourakis M, Argyros A (2007) Enforcing scene constraints in single view reconstruction. In: Proceedings of EUROGRAPHICS 2007 short papers, pp 45–48

25. Lowe D (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110

26. Magnenat-Thalmann N, Foni AE, Cadi-Yazli N (2006) Real-time animation of ancient roman sites. In: Proceedings of 4th international conference on computer graphics and interactive techniques, GRAPHITE, pp 19–30

27. Moulon P, Monasse P, Marlet R (2013) Global fusion of relative motions for robust, accurate and scalable structure from motion. In: Proceedings of IEEE international conference on computer vision, pp 3248–3255

28. Muller P, Zeng G, Wonka P, Gool LV (2007) Image-based procedural modeling of facades. Proc SIGGRAPH 2007:181–184

29. Nister D, Stewenius H (2006) Scalable recognition with a vocabulary tree. Proc IEEE Conf Comput Vis Pattern Recogn 2:2161–2168

30. Papagiannakis G, Schertenleib S, O'Kennedy B, Arevalo-Pozat M, Magnenat-Thalmann N, Thalmann D (2005) Mixing virtual and real scenes in the site of ancient pompeii. Comput Anim Virtual Worlds 16(1):11–24

31. Patra S, Bhowmick B, Kalra P, Banerjee S (2012) Kinect. In: VISIAPP 2012
32. Petridis P, White M, Mourkousis N, Liarokapis F, Sifiniotis M, Gatzidis C (2009) Exploring and interacting with virtual museums. J Comput Cult Heritage 2(1):1–20
33. Raghuram R, Wu C, Frahm J, Lazebnik S (2011) Modeling and recognition of landmark image collections using iconic scene graphs. Int J Comput Vis 95(3):213–239
34. Schuon S, Theobalt C, Davis J, Thrun S (2008) High-quality scanning using time-of-flight depth superresolution. In: IEEE Computer society conference on computer vision and pattern recognition workshops 2008
35. Schuon S, Theobalt C, Davis J, Thrun S (2009) Lidarboost depth superresolution for tof 3d shape scanning. In: CVPR 2009
36. Sharma A, Kumar S (2014) User-guided modulation of rendering techniques for detail inspection. In: GRAPP 2014—Proceedings of the 9th international conference on computer graphics theory and applications, Lisbon, Portugal, 5–8 January, 2014, pp 247–254
37. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905
38. Slabaugh G, Schafer R (2004) Methods for volumetric reconstruction of visual scenes. Int J Comput Vis 179–199
39. Smisek J, Jancosek M, Pajdla T (2011) 3d with kinect. In: IEEE workshop on consumer depth cameras for computer vision
40. Snavely N, Seitz S, Szeliski R (2006) Photo tourism: exploring photo collections in 3d. In: Proceedings of ACM SIGGRAPH, pp 835–846
41. Snavely N, Seitz S, Szeliski R (2008) Modeling the world from internet photo collections. Int J Comput Vis 80(2):189–210
42. Snavely N, Seitz S, Szeliski R (2008) Skeletal graphs for efficient structure from motion. In: Proceedings of IEEE conference on computer vision and pattern recognition, pp 1–8
43. Sturm PF, Maybank SJ (1999) A method for interactive 3d reconstruction of piecewise planar objects from single images. Proc BMVC 1999:265–274
44. Trapp M, Semmo A, Pokorski R, Hermann CD, Dollner J, Eichhorn M, Heinzelmann M (2010) Communication of digtial cutltural heritage in public spaces by the example of roman cologne. In: Proceedings of Euro Mediterranean conference, pp 262–276
45. Triggs B, Mclauchlan P, Hartley R, Fitzgibbon A (2000) Bundle adjustment a modern synthesis. In: Vision algorithms: theory and practice, LNCS, pp 298–372
46. Wu C (2007) SiftGPU: a GPU implementation of scale invariant feature transform (SIFT). http://cs.unc.edu/~ccwu/siftgpu
47. Wu C (2013) Towards linear-time incremental structure from motion. In: Proceedings of the international conference on 3D vision, 3DV '13, pp 127–134
48. Wu C, Agarwal S, Curless B, Seitz S (2011) Multicore bundle adjustment. In: Proceedings of IEEE conference on computer vision and pattern recognition, pp 3057–3064
49. Yang Q, Yang R, Davis J, Nistr D (2007) Spatial-depth super resolution for range images. In: CVPR 2007

Robust Feature Matching for Architectural Scenes

Prashanth Balasubramanian Vinay Kumar Verma
Moitreya Chatterjee and Anurag Mittal

1 A Performance Comparison of Feature Descriptors for Matching in Architectural Images

1 1 Review of Past Work

Identification of point correspondences between images is an important problem that finds application in many tasks such as Registration, Stitching, Disparity Matching, 3-D Reconstruction, Tracking, Object Identification and Classification. As the transformations between the images are seldom known a priori, the practice is to localize on distinctive regions of images (called as *keypoints*) and match them under different transformations. Matching of keypoints across 2 images is done by building *feature descriptors* that express the visual characteristics of the regions around the keypoints, and correspond them using a suitable distance metric. The descriptors are expected to be su ciently distinctive so as to represent the keypoint and be robust to geometric transformations, illumination variations, different blurs and artefacts due to sampling and compression.

P. Balasubramanian ( )   M. Chatterjee   A. Mittal
Indian Institute of Technology Madras, Chennai, India
e-mail: bprash@cse.iitm.ac.in

M. Chatterjee
e-mail: metro.smiles@gmail.com

A. Mittal
e-mail: amittal@cse.iitm.ac.in

V. K. Verma
Indian Institute of Technology Kanpur, Kanpur, India
e-mail: vkverma@cse.iitk.ac.in

Many interesting attempts have been made to design descriptors which satisfy these said characteristics. Early work used the raw pixels of the regions around the keypoints and studied their correlation measure. As correlation measures do not consider geometric information, such measures cannot tolerate localization errors of keypoints, and so are good when the regions are exactly registered. Further, these measures can only handle linear changes in intensities while it is well known that nonlinear variations in illuminations are commonplace occurrences, especially in the under-saturation and over-saturation regions.

Gradient-based methods have proposed effective strategies to handle many of these challenges. The popular SIFT [1] algorithm captures the local gradient distributions around the keypoints. Bay et al. [2] propose a faster variant of SIFT called as *SURF*, by computing Haar wavelet responses using integral images. It is also compact(64 dimensions) and uses the sign of the Laplacian to perform faster indexing. The *GLOH* descriptor [3] improves the robustness and distinctiveness of *SIFT*. It divides the region into a log-polar network of 17 spatial bins, on each of which is a 16-dimensional orientation histogram built. PCA is used to reduce the 272 dimensions to 128 which are used in matching. Ke and Sukthankar [4] propose a dimensionally reduced descriptor *PCA-SIFT* by vectorizing the x and y gradients of the pixels of the normalized patch and linearly projecting the vectors onto a much lower dimensional (30) eigenspace. They argue that an eigen projection is su cient to model the variations in the 3D-scene and viewpoints, although the evaluation in [3] shows other descriptors to perform better. Shape context [5] is another method that bins the orientations of pixels into a log-polar grid. Although the authors applied it only for edge point locations and not orientations, it can be used as a region descriptor as well [3]. Apart from these, there are also other modifications of gradient histograms such as those in [6–8].

Order-based descriptors that are constructed based on the sorting of pixels are an alternative strategy to gradient-based descriptors. Zabih and Woodfill [9] proposed two techniques—*rank* and *census transforms*—that are based on the order of intensities of neighbours of a pixel and the count of flipped point pairs. Such order-based methods are inherently invariant to monotonic changes in illumination. However, they fail in the presence of pixel noise as a single salt-and-pepper flip can change the counts, which is alleviated to a certain extent by Bhat and Nayar [10]. Mittal and Ramesh [11] improve the latter by penalizing an order-flip in proportion to the change in the intensities of the pixels that underwent the flip. This helps to prevent the movement of pixels due to Gaussian noise. Tang et al. [12] propose the *OSID* descriptor that builds a histogram of orders computed on the entire patch. Though invariant to monotonic illumination variations, it can fail on a patch having many pixels of similar intensities as these tend to shift under Gaussian noise. Gupta and Mittal [13] alleviate this problem by designing a histogram of relative intensities whose bins are adaptively designed for the saturated and the non-saturated regions. Wang et al. [14] improve upon this in their *LIOP* descriptor by inducing rotation invariance to it. The motivation is based on their study [15] that identifies estimation of keypoint orientation as a major source of localization error.

There are other variants of order-based descriptors that are bit strings of comparisons of pixels. These are attractive because of their minimal storage requirements and their ability to be compared fast. *Local Binary Patterns* (*LBP*) [16], first applied for face recognition and texture classification, are formed by the comparison of a pixel with its neighbours and constructing a histogram of these patterns. Since these patterns are rather high-dimensional, variants such as [13, 17] compare only certain pixels in the neighbourhood without sacrificing the discriminative ability of the *LBP* patterns. Calonder et al. [18] propose the *BRIEF* descriptor that randomly samples 128 or 256 pixel pairs from the smoothed patch and forms a bit string based on the outputs of their comparisons. The bit string turns out to be, surprisingly, discriminative. Because of the manner in which it is constructed, *BRIEF* is not rotation-invariant and Rublee et al. [19] propose the *ORB* descriptor that makes *BRIEF* rotation-invariant. Leutenegger et al. [20] design a variant of *BRIEF* called as *BRISK* [20] that is formed by the comparisons of pixels placed uniformly on concentric circles. The region is rotation-normalized according to the orientation estimated from the pixels on the circles. To avoid aliasing while sampling points from the circles, each point is smoothed by a Gaussian window of width that is su cient to not distort the information content of close-by points. They also propose a fast keypoint detector. The *FREAK* descriptor by Alahi et al. [21], is another binary descriptor that compares intensities of pixels sampled in a pattern as observed in the human retinal system. They also outline the reason behind why such comparison-based binary descriptors work, based on studies of the human visual system. Mikolajczyk and Schmid [3] provide an extensive comparison of many keypoint descriptors including *SIFT SURF shape context SIFT-PCA GLOH cross-correlation and steerable filters* and observe that, although *SIFT* performs well in many scenarios, there is no one particular descriptor which works for all cases. A comparison of the modern descriptors has been made independently by Miksik and Mikolajczyk [22] and Heinly et al. [23].

In the first part of this chapter, we aim to study the performance of four descriptors—*SIFT*, *LIOP*, *HRI* and *HRI-CSLTP*—for matching keypoints in the applications of stitching of images of architectural scenes. This chapter is an elaboration of the work described in [24]. Such images are characterized by well-structured and textured monuments that can be varying in depth, may have large areas of homogeneous regions especially when shot for a panoramic mosaic and can have varying illumination levels. Accordingly, we test these descriptors on four kinds of images from a dataset of archaeological sites and historical monuments: (1) well structured with su cient depth variation (2) partly structured and partly homogeneous (3) nearly homogeneous with a few structured regions and (4) illumination change on a dataset. We aim to study the scope of application of these descriptors by testing them on the said challenges. To that end, we plot their response graphs for matching, compare their performances and draw conclusions.
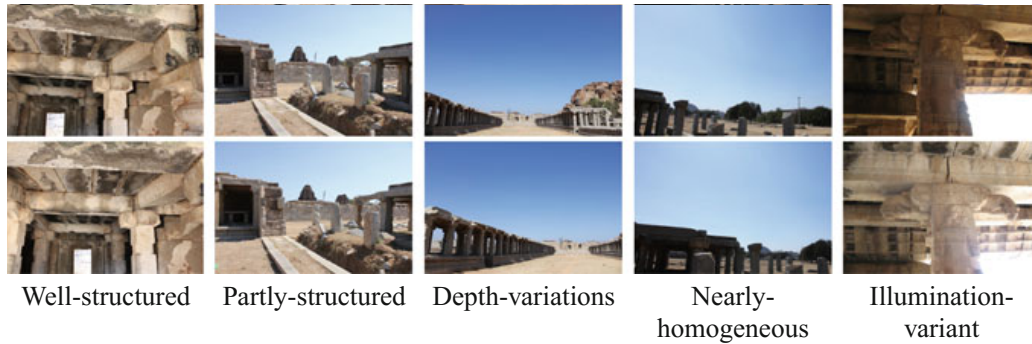
Well-structured   Partly-structured   Depth-variations   Nearly-   Illumination-
                                                            homogeneous   variant

Fig. 1 Challenges that usually beset a feature matcher

1 2 Architectural Scenes and Their Challenges

Figure 1 shows some images from a typical dataset of archaeological sites and historical monuments. Such monuments are usually structured[1] with repeated occurrences of textured regions (col. 1 of Fig. 1) at varying levels of depths (cols. 2 and 3 of Fig. 1). These images may also include large homogeneous regions, especially when shot for a panoramic mosaic or 3-D reconstruction, with a vacant landscape in the front or sky in the back (col. 4 of Fig. 1). Homogeneous regions are poor conveyors of distinctive visual information. So, when large areas of images are covered by homogeneous regions, it becomes important to match the available keypoints from the non-homogeneous regions in a reliable and correct manner, and discard as many pseudo-matches as possible. The descriptors have to be highly distinctive to suit this requirement. Further, the lighting conditions and the time of the day when the images are shot govern the intensities of the pixels and can make them vary in a nonlinear way (col. 5 of Fig. 1), especially in under-exposed or over-exposed regions (for instance, interior structures that are poorly lit). The descriptors need to be resilient to these changes in intensities by adopting a generic normalization technique.

In the next section, we present a brief overview of the four descriptors—*SIFT HRI HRI-CSLTP LIOP*—that are tested in these challenges. While *SIFT* [1] is well known, *HRI-CSLTP* [13] and *LIOP* [14] are recent order-based descriptors that have performed well on the standard datasets [13, 14, 22].

---

[1]A region of image is well structured when it is characterized by regular occurrences of homogeneous or textured patches that are flanked by well-defined object gradients. A typical example is that of a building, as opposed to an image of a scenery.

1 3   Overview of the Descriptors

As the *SIFT* descriptor has been extensively studied, its design, construction and properties are well known. So, we skip its description and proceed to the other feature descriptors.

Histogram of Relative Intensities (*HRI*) descriptors [13] capture the relative orders of the pixels of the patch based on their intensities. Orders have the natural ability to be invariant to monotonic changes in illumination. In a *HRI* descriptor, pixels bin their intensities into intervals that are designed based on the intensity distribution of the overall patch. Linear normalization of intensities yields illumination invariance, wherein the min and the max points of the normalization are adaptively chosen for the saturated and the non-saturated regions.[2] Gaussian pixel noise is handled by a uniform distribution of the intensities into the intervals, and trilinear interpolation and spatial division of the patch into grids handle small pixel movements. It is to be noted that gradient information is not used, in contrast to SIFT [1].

Centre-symmetric local ternary patterns (*CSLTP*) descriptors [13] look at the intensity differences of the diagonal neighbours of each pixel and encode them using 3 categories based on a threshold parameter, T; two of the categories identify differences of opposing contrast, $i_1 - i_2 > T$, while the third identifies pixels of nearly equal intensities, $i_1 - i_2$ T. T helps to choose a certain amount of separation between the diagonal pairs. With 2 diagonal pairs, each being encoded with 3 patterns, there are totally 9 different neighbourhood patterns which can be treated as the 9 bins of the *CSLTP* histogram. Based on its pattern, each pixel contributes a weighted vote to one of the 9 bins. The weight is designed to eliminate a pixel if it has nearly homogeneous neighbours and, thereby, prevent its movement. The patch is divided into 4×4 grids to counter small spatial errors and the *CSLTP* histograms of the grid are concatenated to yield the *CSLTP* descriptor of the patch.

Local Intensity Order Pattern (*LIOP*) descriptors [14] are designed to be rotation and monotonic illumination invariant by using the order of the intensities of the pixels. The local intensity order pattern of a pixel is a weighted vector that encodes the ranking of its 4 neighbours. The neighbours are sampled from a circular neighbourhood in a rotation-invariant manner to avoid the errors in estimation of keypoint orientation [15]. Gaussian noise is handled by giving more weights to the patterns that result from neighbours differing in their intensities by a certain threshold. In addition to the local patterns, the patch is intensity thresholded using multiple values to yield regions of similar intensities, called as ordinal bins. The *LIOP* pattern of an ordinal bin is the weighted summation of those of its pixels; these *LIOP* patterns are concatenated in the order of the ordinal bins resulting in a rotation-invariant *LIOP* descriptor of the patch.

---

[2]A region is saturated if its pixels have intensities either below 10 or above 245.

1 4   Dataset and Evaluation Criterion

We evaluate the descriptors on an architectural dataset which contains images of many archaeological monuments and historical sites. The images,   50 K in all, have been shot in two resolutions (1280×960 and 3648×2736) and are categorized according to varying details of the structures of the sites and thus, made suitable, for different tasks such as panorama stitching and 3D-reconstruction.

For testing the descriptors on image registration for Mosaicking, images shot with the panoramic constraints[3] have been chosen. Following are the challenges based on the nature of the scene that have been used to test the descriptors: (1) well structured with su   cient depth variations (2) partly structured and partly homogeneous (3) nearly homogeneous with a few structured regions and (4) illumination changes. Estimation of homography for a pair of images is done with the manual input of 4 point correspondences.

We use the evaluation criterion proposed by Mikolajczyk and Schmid [3] that identifies the correct and the false descriptor matches using ground truth correspondences at a particular region overlap error (50% in our experiments), as defined by Mikolajczyk et al. [25]. The descriptor matches are obtained using the ratio test proposed by Lowe [1], the threshold for which is varied to obtain the points on the precision–recall response graphs. The correspondences of the regions for a particular overlap error (50%) and the validation of the descriptor matches have been computed using the code available at the A   ne Covariant Features page.[4]

DoG keypoints [1] are detected using the covariant feature detector routine in the *VL-FEAT* library [26]. The minimum absolute value of the cornerness measure is empirically set to 3 for all the experiments. For the *SIFT* and the *LIOP* descriptors, the implementations in the *VL-FEAT* library are used. *HRI* and *HRI-CSLTP* have been implemented by us.

1 5   Performance Evaluation

1.5.1   Images with Illumination Variations

Images taken in an uncontrolled environment such as archaeological sites exhibit wide variety of intensity ranges depending on the ambient light which need not illumine the objects in the scene uniformly, especially the interior parts of structures and can thus, result in under-saturated or over-saturated regions. Such variations in the intensities are usually nonlinear and hence, the descriptors have to deal with an appropriate normalization scheme. Figure 2 shows the performance of the descriptors on images that vary in their illumination patterns. These are usually indoors

---

[3]A set of images is suitable for panoramic stitching if all of them depict a planar scene or are shot with the camera centre being fixed.

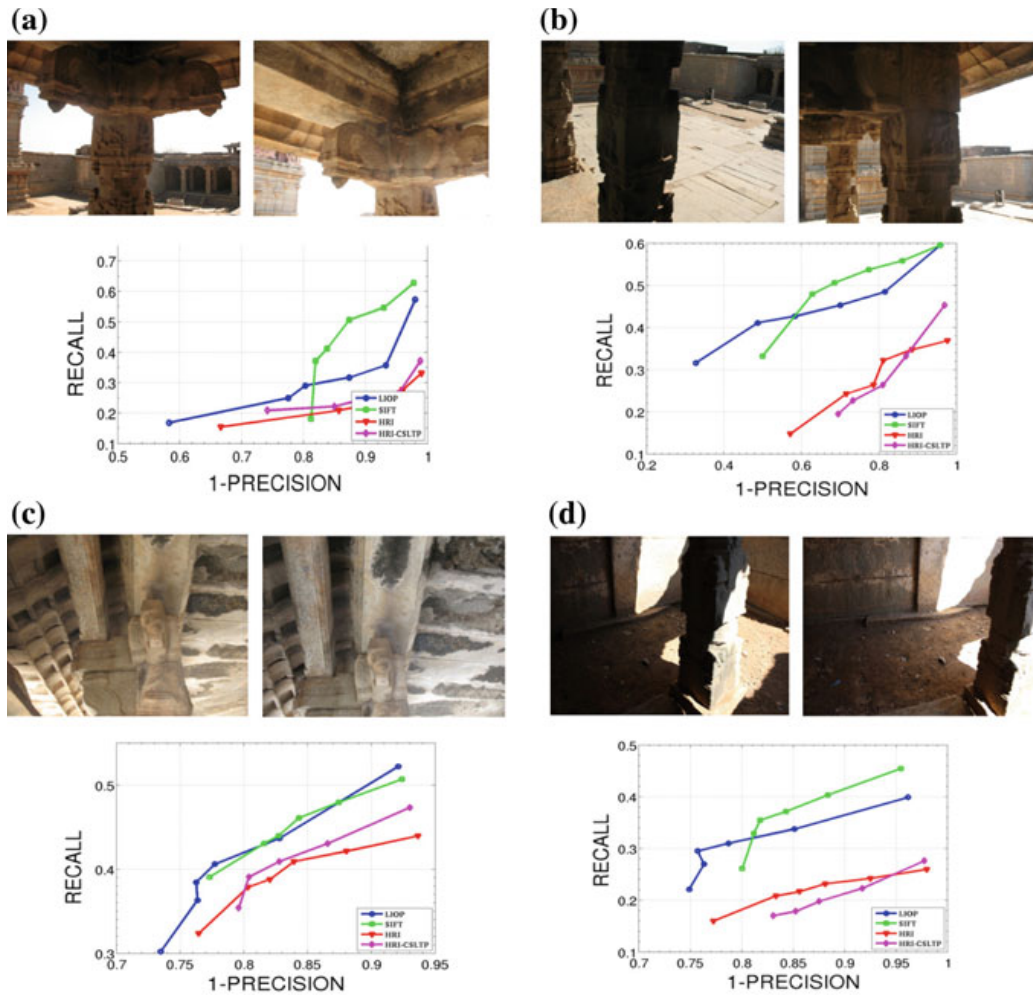[4]https://www.robots.ox.ac.uk/~vgg/research/a   ne/desc_evaluation.html.

Fig. 2 The performance of the descriptors on images with intensity variations. The ranges of the plots have been set different for the sake of clarity

where the natural light does not reach all portions of the scene uniformly. The recall rate is generally low as it is 30% when the precision is 30% for the best performer(s), except in Fig. 2b which might be due to the good matches from the well-lit outdoor structures. *SIFT* seems to be doing consistently well, although *LIOP* is not far behind. Though *HRI-CSLTP* and *HRI* use adaptive binning, the changes in these images might be very nonlinear for these methods to perform well.

1.5.2 Structured Images

Figure 3 shows the performance of the descriptors for images that are well structured with some depth variations and nearly well-lit light conditions. The aim here is to study if the descriptors can match the keypoints output by the detector when they vary in their texture content due to depth and viewpoint changes. *SIFT* and *HRI-CSLTP*
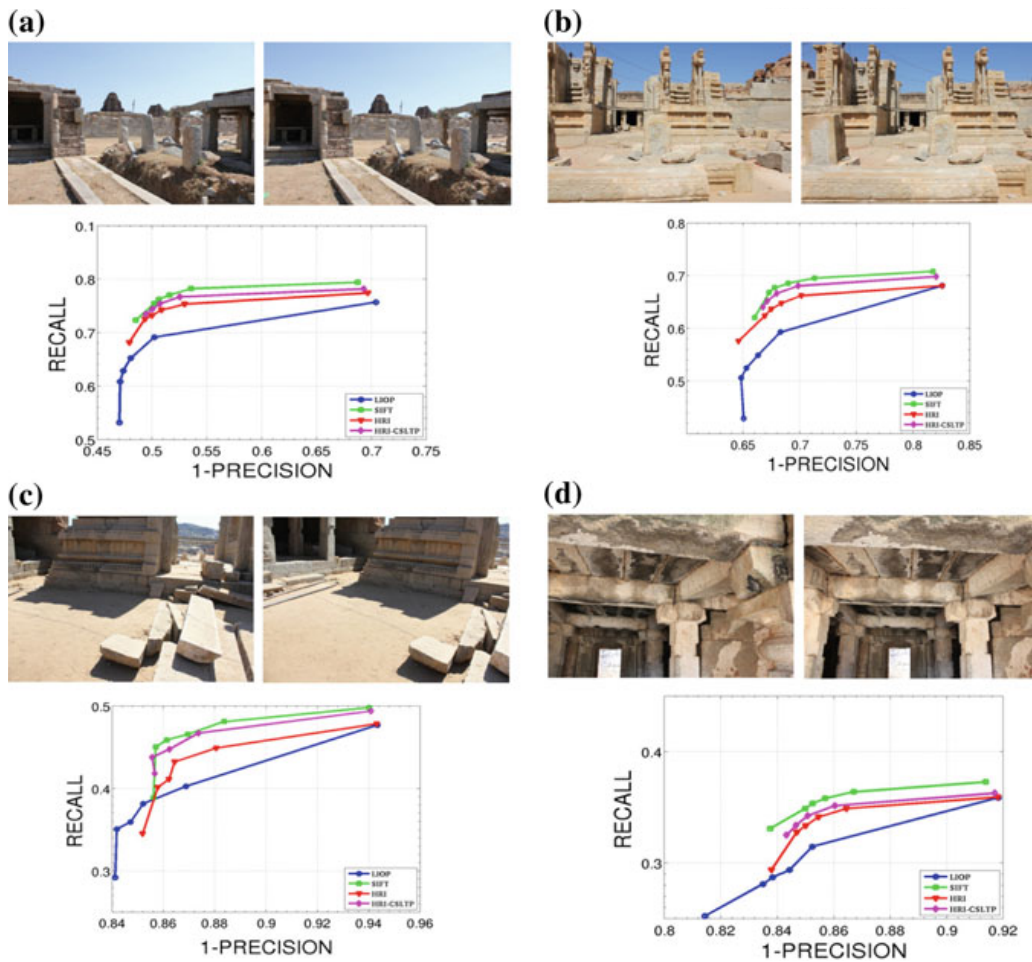
Fig. 3 The performance of the descriptors on well-structured images with some depth variations. The ranges of the plots have been set different for the sake of clarity

perform consistently well in all the four cases. The additional edge direction information in *HRI-CSLTP* definitely helps it score better than *HRI*, although the marginal differences in their performances might suggest that *CSLTP* may have to be combined with other descriptors as it captures directional information only in four orientations.

1.5.3 Partially Homogeneous Images

Figure 4 shows the performance of the descriptors for images that are partially homogeneous containing large depth variations. Such images are usually captured to get a profile of the entire scene when it contains objects that vary significantly in their depths (e.g. a long wall flanked by a bare landscape on its side). For matching, the descriptors have to rely on the keypoints generated from the structured regions of the images. We find that *SIFT* and *HRI-CSLTP* perform well with the differences
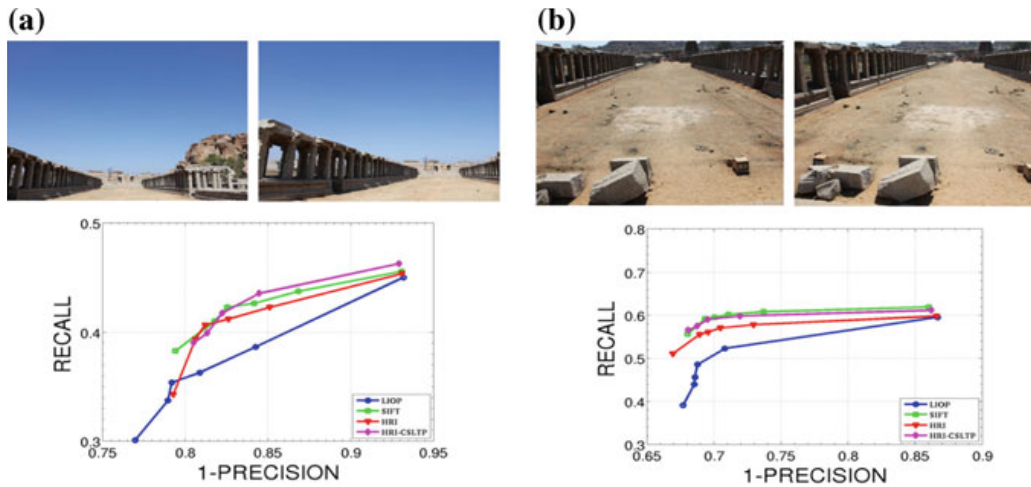
Fig. 4 The performance of the descriptors on partially homogeneous images with significant depth variations. The ranges of the plots have been set different for the sake of clarity

being very marginal in both the test cases. The orders of the pixels considered in *LIOP* can become noisy in homogeneous regions and that may explain the nature of its performance in these cases.

1.5.4 Nearly Homogeneous Images

Figure 5 shows the performance of the descriptors for images that are nearly homogeneous with very little amount of structures in them. Such images are usually captured in a panoramic shot of an architectural monument that has a nearly empty landscape in the front. The low ranges of precision in Fig. 5 can be explained by the fact that nearly homogeneous regions tend to result in large number of false matches. The trend exhibited by the descriptors is the same as in the previous 2 challenges. Though the order patterns used in *LIOP* are weighted, the results suggest that the weighting might not be su cient when there are large areas of homogeneous regions.

1 6 Conclusion

We presented a performance evaluation of four feature descriptors for the task of feature matching in image stitching when the images are of archaeological scenes and architectural sites. As these images are characterized by structures that vary in their textural content and depth and homogeneous regions, we categorized the dataset into four classes and tested the descriptors on them. *SIFT* and *HRI-CSLTP* perform better than the others in many of the test cases highlighting their distinctiveness in representing the keypoint regions. *LIOP* performs well when the intensity variations are
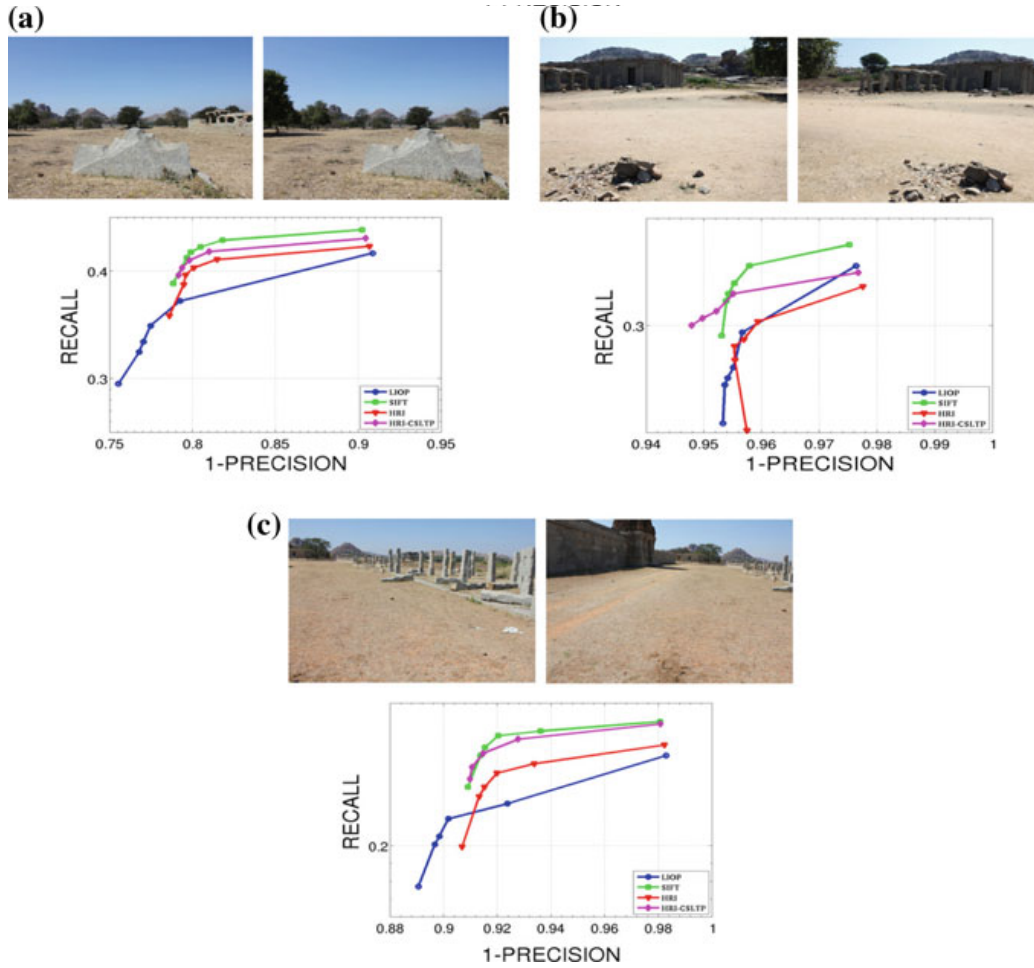
Fig. 5 The performance of the descriptors on nearly homogeneous images with very little structures. The ranges of the plots have been set different for the sake of clarity

complex. Also, the results of *LIOP* show that the order computations have to be done in a noise-resilient manner, especially when homogeneous regions are involved. This performance evaluation can be extended to other applications like 3-D reconstruction to understand the scope of applicability of these descriptors.

2 A New Cross-Bin Distance Measure with Inverse Frequency Component for Comparing Histograms

In the second part of this chapter, we study some distance measures used in computing the similarities between histograms. Distance measures compute the similarity scores by comparing the values of the bins of the histograms. Based on the manner in which the values of the bins are compared, distance measures can be classified as being either *bin-by-bin* or *cross-bin* methods [27]. *Bin-by-bin* methods only compare

the values of the corresponding bins of the histograms while *cross-bin* methods perform a full-blown, exhaustive comparison of all the bins of the histograms. While the former are simple to understand and implement, the latter are more beneficial when the histograms are not well aligned and the bins are not necessarily very different from each other. In the subsequent sections, following a brief revisit of the current distance measures, we propose a new distance measure that combines the advantages of cross-bin methods and methods that adopt inverse frequency normalization. Experiments are performed for the problem of keypoint matching using the SIFT descriptors on natural and architectural images to prove the e cacy of the proposed method.

2 1 Review of Distance Measures

L_1 and L_2 (or 1- and 2-norms) are distance measures which calculate distances (dissimilarities) between two N-bin histograms via a bin-by-bin comparison [28]. L_2 is commonly used to obtain the matches of keypoints, refer to [1, 3]. Mukherjee et al. [29] have shown that L_1 also performs well in certain cases of retrieval.

^{2} distance measure is a variation of the L_2 distance measure which calculates the distance(dissimilarity) between two N-bin histograms via a bin-by-bin comparison and by weighing the bins appropriately [30]. The computation is done as follows. Say, H_1 and H_2 are the two N-bin histograms representing the two features. Then,

$$\text{dist}(H_1, H_2) = \sum_{b=1}^{N} \frac{[H_1(b) - H_2(b)]^2}{H_1(b) + H_2(b)}$$

where $^2(H_1, H_2)$ is the measure of dissimilarity between two histograms H_1 and H_2, $H_i(b)$ is the mass of the bth bin of histogram H_i. The idea is to weigh the bins in an appropriate manner so that their contributions are normalized. Thus, bins with large values are weighed less so that such values do not dominate the computation of distance between histograms while bins with small values are weighed more so that their contribution to the distance is boosted up.

The *Earth Mover's Distance* (*EMD*) is a cross-bin distance (dissimilarity) measure between two histograms [27]. It is defined as the minimal cost that must be paid to transform one histogram into the other. To compute this distance between two given histograms, the cost of moving a unit mass from a bin of the first histogram into a bin of the second histogram needs to be defined. This is known as the *Ground Distance* between a pair of bins. Mathematically, *EMD* is defined as follows. Let $H_1 = [H_1(i)]_1^m$ and $H_2 = [H_2(j)]_1^n$ denote the two histograms of lengths m and n. Let c_{ij} denote the ground distance and f_{ij} denote the amount of mass transferred between $H_1(i)$ and $H_2(j)$. Then, the *EMD*(H_1, H_2) is

$$EMD(H_1, H_2) = \min_{f_{ij}} \{\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} f_{ij}\}, \quad \text{subject to} \tag{1}$$

$$f_{ij} \geq 0, \quad 1 \leq i \leq m, \ 1 \leq j \leq n \tag{2}$$

$$\sum_{i=1}^{m} f_{ij} \leq H_2(j) \tag{3}$$

$$\sum_{j=1}^{n} f_{ij} \leq H_1(i) \tag{4}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} f_{ij} = \min\left(\sum_{i=1}^{m} H_1(i), \sum_{j=1}^{n} H_2(j)\right) \tag{5}$$

Equation 2 refers to the fact that all the mass transfers are unidirectional from H_1 to H_2. Equation 3 requires that the total transfer into a bin of H_2 is bounded by the mass contained in it which can be considered as its 'demand'. Equation 4 requires that the total transfer from a bin of H_1 is bounded by the mass contained in it which can be considered as its 'supply'. Equation 5 imposes the maximum amount of transfer from H_1 to H_2.

The *Quadratic-Chi* Histogram distance [31] aims to combine the advantages of the cross-bin measures and χ^2 like normalization. It is a quadratic form distance that helps in comparing bins that are not corresponding while the χ^2 like normalization helps in normalizing the bins and thereby boost the significance of small bins. Given H_1 and H_2, it is computed as follows:

$$QC_m^A(H_1, H_2) = \sqrt{\sum_{ij}\left(\frac{(H_1(i) - H_2(i))}{(\sum_c (H_1(c) + H_2(c))A_{ci})^m}\right)\left(\frac{(H_1(j) - H_2(j))}{(\sum_c (H_1(c) + H_2(c))A_{cj})^m}\right)A_{ij}},$$

$H_i(c)$ denotes the mass of the cth bin of histogram H_i, A is a positive, semi-definite, bin similarity matrix such that $A_{ij} \leq A_{ii}$.

2 2 *Proposed Approach*

The Earth mover's distance is a measure that compares the value of a bin of the first histogram with all the other bins of the second histogram. Nevertheless, the distance so computed can be influenced by the absolute values of the bins. As a result, 'heavy' bins which represent common but insignificant features can dominate the computation of the distance while 'light' bins that represent rare, yet significant features can be suppressed. The metric of inverse document frequency is an attempt to address

this issue for the problem of text retrieval. The χ^2 distance addresses this by normalizing each bin; however, it fails to compare a bin with all the others using a cross-bin strategy. Our aim is to perform a cross-bin comparison of the given histograms while simultaneously boosting the significance of 'light' bins and suppressing the significance of 'heavy' bins. To achieve the said aim, we propose the following distance measure, $EMD(H_1, H_2)$ which, given two histograms H_1, H_2, is defined as follows:

$$EMD(H_1, H_2) = EMD(\widehat{H}_1, \widehat{H}_2) \text{ where} \tag{6}$$

$$\widehat{H}_1(i) = \frac{H_1(i)}{\left\{ \sum_k a_{ki} \left(H_1(k) + H_2(k) \right) \right\}^m}$$

$$\widehat{H}_2(i) = \frac{H_2(i)}{\left\{ \sum_k a_{ki} \left(H_1(k) + H_2(k) \right) \right\}^m}, \text{ where}$$

$$a_{ki} = 1 - \frac{c_{ki}}{max_{k,i}(c_{ki})},$$

$$c_{ki} = \text{ ground distance between bins } k \text{ and } i$$

$$m >= 0, \text{ a parameter that affects the amount of normalization}$$

To introduce the inverse frequency component, each bin's value is normalized by the weighted sums of the histograms with weights depending on the underlying bin similarities. The notion of using weighted sums of adjacent bins rather a simple sum of the corresponding bins (as used in χ^2) is to enable partially similar bins to play a role in the process of normalization. Intuitively, in a histogram, there is a sense of 'continuity' in the way the distribution of mass across bins is done. Thus, a heavily populated bin also tends to have populated neighbouring (or similar) bins. So, by allowing such neighbouring bins to contribute to the sum (in the normalization term of the denominator), we hope to achieve a more robust and meaningful normalization of the bins to the effect that 'heavy' and 'light' bins contribute equally to the computation of distances.

Although the said normalization transforms and brings the heavily and sparsely populated bins on an equal footing, it will still be to our advantage to consider non-corresponding bins (or cross bins) in the distance computation. A cross-bin measure such as EMD that utilizes the underlying similarities of the normalized bins by computing the appropriate mass flows amongst them has the potential to combine the merits of χ^2 like normalization with those of EMD's. The experimental Sect. 2.3 shows some results to this end.

2 3   Experimentation and Results

The proposed distance measure is tested on the problem of matching of keypoints described using the SIFT feature (Difference of Gaussian keypoints described using SIFT) [1]. Given p keypoints from an *image*1, the task is to find a corresponding keypoint for each of them from the q keypoints of *image*2. The images can differ by simple geometric transformations, complex viewpoint changes, compression factors, blur and light changes. We show the results on the Oxford Dataset [3] and the images from the Hampi Dataset. We follow the same procedure as described in Sect. 1.4 to evaluate the matches obtained by the distance measures.

In Table 1, the areas under the curve (*AUC*)% of the P-R graphs for the various distance measures are shown. In each case, the keypoints of the first image, *Image*1, have been matched with those of the corresponding second image which differs from the former by the said variation. For *EMD* and   *EMD*, we used the 0-1 ground distance as suggested in [32, 33]. This is defined as $c_{ij} = 0$, for $i = j$ and $c_{ij} = 1$, for $i \neq j$. For *QC*, we defined a_{ij} as given in Eq. 6 for the 0-1 ground distance. Further, it has been found for the SIFT descriptor that computing the *EMDs* in a grid-wise fashion (a SIFT patch typically has 4x4 grids) and finally summing them perform better than computing on the whole of the patch, refer to [33].

We observe that   *EMD* does well especially when there is a geometric change (e.g. viewpoint, rotation, zoom) in between the pair of images. This, we believe, is because of the effect that cross-bin normalization has on the distance computation. We also observe that a single value of m does not come out as a winner in all the challenges. The choice of m and the ground distance has been empirical and it will be illustrative to cast their estimation in a learning framework.

In Table 2, the performances of the distance measures on some of the Hampi images have been shown. The images have been chosen such that there exists an inherent homography between every pair of them. The images are shown in Fig. 6. We see that the cross-bin measures, especially the *EMD* variants, are generally doing well on these image pairs, with the proposed measure outperforming the others on a few cases. Figure 7 shows a stitched image that is obtained from the first pair of images of Fig. 6. For this, the matches between the keypoints were obtained from   *EMD* and the homography was estimated using RANSAC following which registration and stitching were performed.

2 4   Conclusion

We proposed a new cross-bin measure with an inverse frequency component for comparing histograms. Cross-bin measures were shown to be better than the naive bin-by-bin measures, while boosting the contribution of rare bins and suppressing that of frequent bins improved their performances more. The proposed measure was compared with others for the problem of keypoint matching on natural and

Table 1 (*AUC%*) values for matching the keypoints of image 1 with those of the corresponding second image for different image pairs from the Gra ti, Boat, Bark and UBC datasets, refer [3]

Gra ti (view-point)	img1–img2	img1–img3	img1–img4	img1–img5	img1–img6
l_1	89.077	70.907	52.406	15.477	2.4556
l_2	87.554	69.178	49.706	11.827	1.601
chi^2	89.253	72.202	54.427	**19.901**	4.0098
EMD	88.61	70.916	53.037	16.229	3.3225
Emd($m = 0.4$)	**89.339**	**72.281**	**55.434**	19.486	**4.8785**
QC	86.437	56.449	36.307	15.16	3.1237
Boat(zoom, rotation)					
l_1	90.755	92.241	77.481	79.312	37.257
l_2	89.97	91.532	75.696	78.312	33.875
chi^2	91.218	92.888	79.081	**81.266**	38.851
EMD	91.103	92.954	77.729	80.055	39.64
Emd($m = 0.4$)	**91.386**	**92.987**	**79.242**	80.757	**42.308**
QC	89.402	90.791	76.486	76.457	30.769
Boat(zoom, rotation)					
l_1	86.765	**55.081**	71.815	96.573	72.418
l_2	86.032	54.551	71.769	95.563	72.231
chi^2	86.43	54.234	64.39	92.255	37.391
EMD	87.317	54.592	**71.865**	96.801	73.147
Emd($m = 0.1$)	**87.406**	54.652	71.847	**97.083**	**73.507**
QC	82.716	52.158	70.457	94.041	69.631
Ubc(compression)					
l_1	97.792	97.08	94.583	90.469	72.204
l_2	98.002	96.917	**95.248**	89.622	71.534
chi^2	68.229	85.529	93.855	**91.816**	**75.719**
EMD	97.541	96.96	94.003	88.819	70.816
Emd($m = 0.6$)	97.972	**97.081**	93.005	85.819	65.809
QC	**98.465**	96.902	94.012	88.943	75.204
Bikes(blur)					
l_1	89.726	**86.273**	85.373	83.387	82.905
l_2	88.631	84.546	83.587	81.832	81.129
chi^2	87.219	84.461	**85.409**	**83.647**	**83.176**
EMD	90.22	86.091	84.496	83.107	81.308
Emd($m = 0.1$)	**90.226**	86.254	84.771	83.28	81.504
QC	84.585	79.62	80.435	77.975	75.595

Table 2 (*AUC%*) values for matching the keypoints of image 1 with those of the corresponding second image of different image pairs from Fig. 6

	img1–img2	img1–img3	img1–img4	img1–img5	img1–img6	img1–img7
l_1	32.773	32.12	**25.732**	43.011	38.577	**32.186**
l_2	32.71	31.635	25.343	43.348	38.55	31.915
chi^2	9.4066	25.125	18.947	40.778	35.753	30.903
EMD	33.636	31.848	25.713	43.572	**38.943**	32.16
$Emd(m = 0.3)$	**33.646**	**32.451**	25.177	**45.925**	38.381	31.507
QC	32.476	31.308	22.769	37.241	35.917	27.262

a) Image 1 b) Image 2 c) Image 3 d) Image 4 e) Image 5 f) Image 6 g) Image 6

Fig. 6 Figure best viewed in colour. A few images taken from the Hampi dataset on which the experiments were performed. A pair of these images is related by a homography which is estimated by a manual selection of 4 point correspondences

Fig. 7 Figure best viewed in colour. A stitched image from the first pair of Fig. 6 using the keypoint matches generated by *EMD*

EM used in stitching.

architectural scenes. Performance gains were reported. The proposed method can be applied to other tasks such as texture recognition, shape matching and image retrieval. Also, in order to make the method completely online, it will be useful to learn to estimate the values of the parameters and the ground distance from the images and their keypoints.

References

1. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110
2. Bay H, Tuytelaars T, Van Gool L (2006) Surf: Speeded up robust features. In: Leonardis A, Bischof H, Pinz A (eds) The proceedings of the 9th European conference on computer vision, vol 3951 of Lecture notes in computer science. Springer, Heidelberg, pp 404–417
3. Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans Pattern Anal Mach Intell 27:1615–1630
4. Ke Y, Sukthankar R (2004) Pca-sift: a more distinctive representation for local image descriptors. In: The proceedings of the IEEE conference on computer vision and pattern recognition, vol 2, pp II–506–II–513
5. Mori G, Belongie S, Malik J (2005) E cient shape matching using shape contexts. IEEE Trans Pattern Anal Mach Intell 27:1832–1837
6. Lazebnik S, Schmid C, Ponce J (2005) A sparse texture representation using local a ne regions. IEEE Trans Pattern Anal Mach Intell 27:1265–1278
7. Mikolajczyk K, Matas J (2007) Improving descriptors for fast tree matching by optimal linear projection. In: The proceedings of the eleventh IEEE international conference on computer vision, pp 1–8
8. Freeman W, Adelson E (1991) The design and use of steerable filters. IEEE Trans Pattern Anal Mach Intell 13:891–906
9. Zabih R, Woodfill J (1994) Non-parametric local transforms for computing visual correspondence. In: Eklundh JO (ed) The proceedings of the 3rd European conference on computer vision. Volume 801 of Lecture notes in computer science. Springer, Heidelberg, pp 151–158
10. Bhat DN, Nayar SK (1998) Ordinal measures for image correspondence. IEEE Trans Pattern Anal Mach Intell 20:415–423
11. Mittal A, Ramesh V (2006) An intensity-augmented ordinal measure for visual correspondence. Proc IEEE Conf Comput Vis Pattern Recogn 1:849–856
12. Tang F, Lim SH, Chang N, Tao H (2009) A novel feature descriptor invariant to complex brightness changes. In: The proceedings of the IEEE conference on computer vision and pattern recognition, pp 2631–2638
13. Gupta R, Patil H, Mittal A (2010) Robust order-based methods for feature description. In: The proceedings of the IEEE conference on computer vision and pattern recognition, pp 334–351
14. Wang Z, Fan B, Wu F (2011) Local intensity order pattern for feature description. In: The Proceedings of the thirteenth IEEE international conference on computer vision, pp 603–610
15. Fan B, Wu F, Hu Z (2012) Rotationally invariant descriptors using intensity order pooling. IEEE Trans Pattern Anal Mach Intell 34:2031–2045
16. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24: 971–987
17. Heikkilä M, Pietikäinen M, Schmid C (2009) Description of interest regions with local binary patterns. Pattern Recogn 42:425–436
18. Calonder M, Lepetit V, Ozuysal M, Trzcinski T, Strecha C, Fua P (2012) BRIEF: computing a local binary descriptor very fast. IEEE Trans Pattern Anal Mach Intell 34:1281–1298
19. Rublee E, Rabaud V, Konolige K, Bradski G (2011) ORB: an e cient alternative to Sift or Surf. In: The proceedings of the thirteenth IEEE international conference on computer vision. IEEE, pp 2564–2571
20. Leutenegger S, Chli M, Siegwart R (2011) BRISK: binary robust invariant scalable keypoints. In: The proceedings of the thirteenth IEEE international conference on computer vision
21. Alahi A, Ortiz R, Vandergheynst P (2012) FREAK: fast retina keypoint. In: The proceedings of the IEEE conference on computer vision and pattern recognition, pp 510–517
22. Miksik O, Mikolajczyk K (2012) Evaluation of local detectors and descriptors for fast feature matching. In: The 21st international conference on pattern recognition, pp 2681–2684

23. Heinly J, Dunn E, Frahm JM (2012) Comparative evaluation of binary features. In: Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C (eds) The proceedings of the 12th European conference on computer vision. Volume 7573 of Lecture notes in computer science. Springer, Heidelberg, pp 759–773
24. Balasubramanian P, Verma VK, Mittal A (2015) A performance evaluation of feature descriptors for image stitching in architectural images. In: Jawahar CV, Shan S (eds) Computer vision—ACCV 2014 workshops, Singapore, Singapore, November 1–2, 2014, Revised selected papers, Part II, pp 517-528. Springer. https://doi.org/10.1007/978-3-319-16631-5_38
25. Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, Schaffalitzky F, Kadir T, Gool L (2005) A comparison of a ne region detectors. Int J Comput Vis 65:43–72
26. Vedaldi A, Fulkerson B (2008) VLFeat: an open and portable library of computer vision algorithms. http://www.vlfeat.org/
27. Rubner Y, Tomasi C, Guibas LJ (2000) The earth mover's distance as a metric for image retrieval. Int J Comput Vis 40:99–121
28. Hafner J, Sawhney HS, Equitz W, Flickner M, Niblack W (1995) E cient color histogram indexing for quadratic form distance functions. IEEE Trans Pattern Anal Mach Intell 17: 729–736
29. Mukherjee J, Mukhopadhyay J, Mitra P (2014) A survey on image retrieval performance of different bag of visual words indexing techniques. In: The proceedings of the IEEE students' technology symposium (TechSym)
30. Zelnik-Manor L, Irani M (2001) Event-based analysis of video. In: The proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp II–123
31. Pele O, Werman M (2010) The quadratic-chi histogram distance family. In: Daniilidis K, Maragos P, Paragios N (eds) The proceedings of the 11th European conference on computer vision. Volume 6312 of Lecture notes in computer science. Springer, Heidelberg, pp 749–762
32. Pele O, Werman M (2009) Fast and robust earth mover's distances. In: The proceedings of the twelfth IEEE international conference on computer vision, pp 460–467
33. Pele O, Werman M (2008) A linear time histogram metric for improved SIFT matching. In: The proceedings of the tenth European conference on computer vision, pp 495–508

Recovering the 3D Geometry of Heritage Monuments from Image Collections

Rajvi Shah Aditya Deshpande Anoop M. Namboodiri and P. J. Narayanan

1 Introduction

The field of large-scale structure from motion (SFM) and 3D reconstruction has seen a steady progress in the past decade. [44] presented Photo Tourism, a system for navigation, visualization, and annotation of unordered Internet photo collections based on a robust method for incremental structure from motion [9]. Snavely's incremental SFM software, Bundler, is widely used since then. Bundler is a robust and e ective system but one with quadratic and cubic costs associated with exhaustive pairwise feature matching and bundle adjustment. The e ectiveness of this system, however, inspired attempts to yield city-scale 3D reconstructions under a day by identifying many independent sub-tasks involved, and leveraging multi-core clusters and GPUs to parallelize these tasks [3, 18]. Since then, researchers have continued to improve the large-scale reconstruction pipeline in many ways.

The large-scale SFM pipeline can broadly be divided into five steps (see Sect. 2). In steps 1 and 2, a large reconstruction problem is broken down into multiple components based on the image connections. Steps 3 and 4 involve pairwise feature matching and 3D reconstruction of a single component; these were at the core of photo tourism. While the original incremental SFM method is still widely used, hierarchical and global methods that di er significantly have been proposed subsequently. In this paper, we present a *multistage* method for steps 3 and 4 that provide greater

R. Shah A. Deshpande A. M. Namboodiri ( ) P. J. Narayanan
CVIT, IIIT Hyderabad, Hyderabad, India
e-mail: anoop@iiit.ac.in

R. Shah
e-mail: rajvi.shah@research.iiit.ac.in

A. Deshpande
e-mail: adity.deshapandeug08@students.iiit.ac.in

P. J. Narayanan
e-mail: pjn@iiit.ac.in

efficiency and completeness to SFM. Our method builds on several prior e orts for reconstructing a single component typically of about 1000 pictures, representing a facade, a building, a street, or parts of it.

The motivation behind our multistage approach is akin to the coarse-to-fine strategies of several vision algorithms. We wish to quickly recover a coarse yet global model of the scene using fewer features and leverage the constraints provided by the coarse model for faster and better recovery of the finer model using all features in the subsequent stages. Feature selection for recovering the coarse model can be based on several aspects. We use the scales of SIFT features for this. The coarse model provides significant geometric information about the scene structure and the cameras, in the form of point-camera visibility relations, epipolar constraints, angles between cameras, etc. By postponing the bulk of the processing until after the coarse model reconstruction, our approach can leverage rich geometric constraints in the later stages for e ective, efficient, and highly parallel operations. We demonstrate that the proposed staging results in richer and faster reconstructions by using more compute power on several datasets.

2   Background and Related Work

Recovering structure and motion from multiple images is a long-studied problem in computer vision. Early e orts to solve this problem were mostly algebraic in nature, with closed-form; linear solutions for two, three, and four views [20] provide a comprehensive account of these now standard techniques. For multi-image sequences with small motions, factorization-based solutions were proposed by [48, 52]. Algebraic methods are fast and elegant but sensitive to noisy feature measurements, correspondences, and missing features. Another class of algorithms took a statistical approach and iteratively solved the reconstruction problem by minimizing the distance between the projections of the 3D points in images and feature measurements ("reprojection error") using nonlinear least-squares technique [50, 51]. These methods are robust to noise and missing correspondences but computationally more expensive than linear methods. The joint optimization of all camera parameters and 3D points by minimization of the reprojection error is now commonly referred to as Bundle Adjustment [53], which has been a long-studied topic in the field of photogrammetry. Advances in robust feature detection and matching [29] and sparse bundle adjustment made the structure from motion techniques applicable to unordered photo collections [9]. Reference [44] presented the first system for large-scale 3D reconstruction using the incremental SFM algorithm on Internet photo collections. Since then, many e orts have been made to push the state of the art.

There are two main tasks involved in a typical reconstruction pipeline, (i) match graph construction—that computes pairwise geometric constraints between the image pairs, and (ii) structure from motion reconstruction—that recovers a globally consistent structure from the match graph. However, in the context of large-scale reconstruction, often these tasks are further divided into sub-tasks.

Match graph construction begins with a filtering step that identifies image pairs that can potentially have a visual overlap (step 1). For city-scale reconstructions, multiple connected components that can be reconstructed independently are identified from the potential image connections (step 2). For each of the connected components, a match graph (or a view graph) is constructed by performing pairwise feature matching for all directly connected nodes and by verifying the matches based on two-view geometric constraints (step 3). Step 2 and step 3 are sometimes performed in a reverse order, i.e., the connected components are identified after feature matching. Each connected component is reconstructed from the pairwise correspondences using structure from motion (step 4) and finally merged into a single reconstruction (step 5). We now explain each of these steps and discuss the related literature in the remainder of this section.

2.1 Selecting Image Pairs to Match

Large-scale image collections often contain images that do not capture the structure of interest. Also, a large number of good images do not match with the majority of the other images, as they capture different parts of the structure and have no visual overlap. With tens of thousands of features per image, the cost of pairwise feature matching is nontrivial. Hence, exhaustively matching features between all pairs of images ($O(n^2)$) would result in a wasted effort in performing expensive feature matching between a large number of unmatchable images. Due to this, most large-scale pipelines first attempt to identify the image pairs that can potentially match using computationally cheaper methods.

Many methods use the global similarity between two images as a measure of matchability and employ retrieval techniques to identify image pairs with potential overlap [3, 8, 14, 18, 28]. Another class of methods use learning techniques to evaluate whether an image pair would match [11, 39]. A preemptive matching (PM) scheme to quickly eliminate non-promising image pairs attempts is also efficient in discarding unlikely pairs [56]. If present, geographical information such as GPS coordinates or geotags can also be utilized to restrict the pairs to match.

2.2 Finding Connected Components

The image connections found in step 1 define an approximate match graph with edges between nodes corresponding to matchable pairs. A connected component is found by performing a depth-first search on this approximate match graph and later pruned by feature matching and geometric verification [8, 17, 18]. Some other pipelines

perform pairwise feature matching (step 3) first to compute a geometrically verified match graph (or a view graph) and then find the connected components [3, 45].

Some methods also propose to make the connected components sparser to improve the efficiency of the SFM step. Especially for incremental SFM with $O(n^4)$ cost, the improvement is significant [24, 46]. These methods solve for global structure recovery of the sparse connected component first and augment the remaining image pairs to the global structure using pairwise geometry.

2.3 Pairwise Feature Matching

Features of two images are matched by computing L_2-norm between the corresponding descriptors and finding the closest feature as a candidate match. Candidate match is confirmed using a ratio test that checks if best match is significantly closer to the query than the next best match. Without leveraging massively multithreaded hardwares like GPUs, exhaustively comparing features between two images is computationally prohibitive ($O(m^2)$ for m features per image) even after reducing the image pairs. Hence, it is common to use approximate nearest neighbor search ($O(m \log m)$) using accelerated search structures such as Kd trees, cascade hash [26, 31].

Alternatively, the efficiency of feature matching can be improved by reducing the search space. [21] train random forest classifiers to learn indistinctive features and eliminate these from the matching pool. [22] suggest that if features are quantized into a very large vocabulary, the quantization would be sufficiently fine to assume that features from multiple images belonging to the same visual word are matches. In [42], we presented a two-stage geometry-aware scheme that leverages coarse epipolar geometry to reduce the number of candidate features to match and also produces denser correspondences by retaining good correspondences on repetitive structures.

2.4 Reconstruction of a Connected Component

Given the match graph/view graph for a connected component, reconstruction can be performed by various structures from motion pipelines. Most SFM reconstruction techniques can be categorized into (i) Incremental SFM, (ii) Global SFM, and (iii) Hierarchical SFM.

Incremental SFM [9, 44] reconstructs the cameras and points starting with a good seed image pair. The reconstruction grows incrementally by adding a few well-connected images, estimating the camera parameters, and triangulating feature matches. To avoid drift accumulation, this is followed by a global bundle adjustment (BA) which refines camera poses and 3D point positions. The complexity of the incremental SFM is $O(n^4)$ due to repeated BA. To improve the efficiency of this step, many methods propose fast approximations of the sparse bundle adjustment and/or exploit many-core architectures to parallelize it [2, 3, 10, 56, 57].

Another class of methods can be classified as global SFM methods as they aim to reconstruct all images at once as opposed to a sequential solution. At its core, these methods attempt to identify global rotations directly from relative rotations without solving for structure [12, 17, 30, 32, 43]. Once the rotations are known, SFM boils down to solving a linear problem of estimating camera translations and 3D structure. [54] estimate global translations by solving for 1D ordering in a graph problem.

References [19, 23] proposed hierarchical methods for SFM that attempts to avoid fully sequential reconstruction typical to incremental SFM methods without using global estimations. In [40], we proposed a multistage approach for SFM that first reconstructs a coarse global model using a match graph of a few high-scale features and enriches it later by simultaneously localizing additional images and triangulating additional points. Leveraging the known geometry of the coarse model allows the later stages of the pipeline to be highly parallel and efficient for component-level reconstruction.

2.5 Merging Reconstructed Components

The connected components of the match graph are independently reconstructed using methods discussed before and later merged into a single reconstruction. Many pipelines merge multiple sub-models by finding the common 3D points across the models and by robustly estimating a similarity transformation using ARRSAC/RANSAC/MSAC [18, 19, 23, 36]. These merging methods mainly di er in their ways of identifying common 3D points. [8], while dividing the match graph into multiple components ensure that common images exist between two components and estimate the similarity transform between two models by leveraging the pairwise epipolar geometry of the link images. Recently, [15] presented a combinatorial approach for merging visually disconnected models of urban scenes.

3 Our Multistage SFM Algorithm

The flow of our algorithm is depicted in Fig. 1. We begin with a set of roughly connected images that represent a single a monument or geographic site. Appearance techniques and geotags can be used to obtain such image components from larger datasets as explained in Sect. 2. Alternatively, images of a site may be captured or collected specifically for image-based modeling, e.g., for digital heritage applications. We first extract SIFT features from these images and sort them based on their scales. Our algorithm then operates in the following main stages.
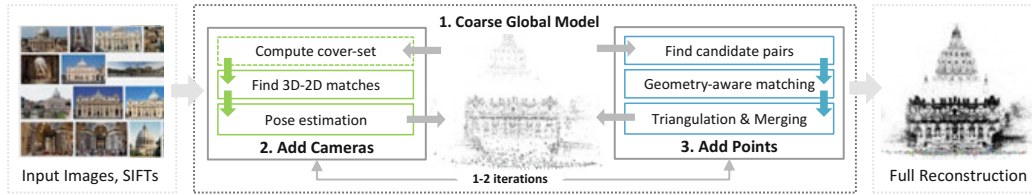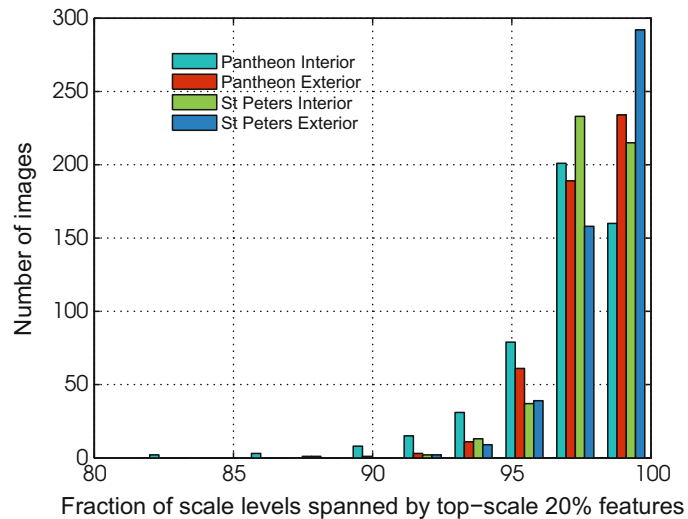
Fig. 1 Flow of our multistage algorithm. Given images of a component, in stage 1, we match a small fraction of the image SIFTs and recover a coarse but global model of the scene using any robust SFM method. In stage 2, camera poses for the images unregistered in the first stage are estimated using fast 3D–2D correspondence search-based localization. In stage 3, the unmatched features of the localized images are matched with a set of candidate images using geometry-aware matching and triangulated to produce the final model. Stages 2 and 3 are highly parallel and do not require bundle adjustment. These stages can be repeated for more iterations if needed

3.1 Coarse Model Estimation

In this stage, a coarse global model of the scene is estimated by SFM reconstruction of high-scale feature match graph. Any robust SFM method can be used for this reconstruction. We experimented with both incremental and global methods for coarse reconstruction as discussed in Sect. 4.

For match graph construction in this stage, we select only % features from all images in descending order of their scales. One should note that this is very different from extracting features from down-sampled images or picking random features. There are two main reasons why we favor higher scale features for reconstruction: (i) Due to successive Gaussian blurring applied to create a scale-space signal, fewer and fewer features are detected at higher scales. Hence, the selected coarse features span across many scales of the scale-space structure. Figure 2 shows the histograms of fractions of scale levels spanned by the top-scale 20% features of randomly

Fig. 2 Histograms showing fractions of scale levels spanned by the top-scale 20% features for sets of 500 images each randomly sampled from four datasets. High-scale features cover many scale levels

sampled 500 images for four datasets. It can be seen that for most of the images across all datasets, more than 90% of the scale levels are spanned by the selected coarse features, indicating broad coverage; (ii) Features detected from the top-scale signals correspond to more stable structures in the scene as compared to the features detected at high-resolution bottom scales which are more susceptible to change with minor variations in the imaging conditions. Due to these two reasons, we consider high-scale features both reliable and sufficient for coarse image matching and geometry estimation.

The latter observation is empirically verified by analyzing the distribution of features by their scales in different models reconstructed using a standard structure from motion pipeline—Bundler. Figure 3a shows the distribution of reconstructed features vs. their percentile rank by scale for four models. Higher scale points clearly are part of more 3D points. The area under the curve is high for value of 10–30. Choosing these features for coarse model reconstruction would enable us to recover many 3D points. Figure 3b shows the number of 3D point tracks that would survive if the top 20% and bottom 20% features by scale are removed from the tracks. The high-scale features are clearly more important than the low-scale ones, as more points are dropped when they are removed. It also indicates that high-scale features not only match well but they also match more frequently to other features of higher scales. We also performed experiments with matchability prediction [21] for feature selection but found the scale-based selection strategy to be more effective for coarse reconstruction.

We performed various experiments to see the effect of on completeness of reconstruction and runtime. We conclude that selecting 20% high scale for initial matching offers an optimum trade-off between connectivity and matching efficiency
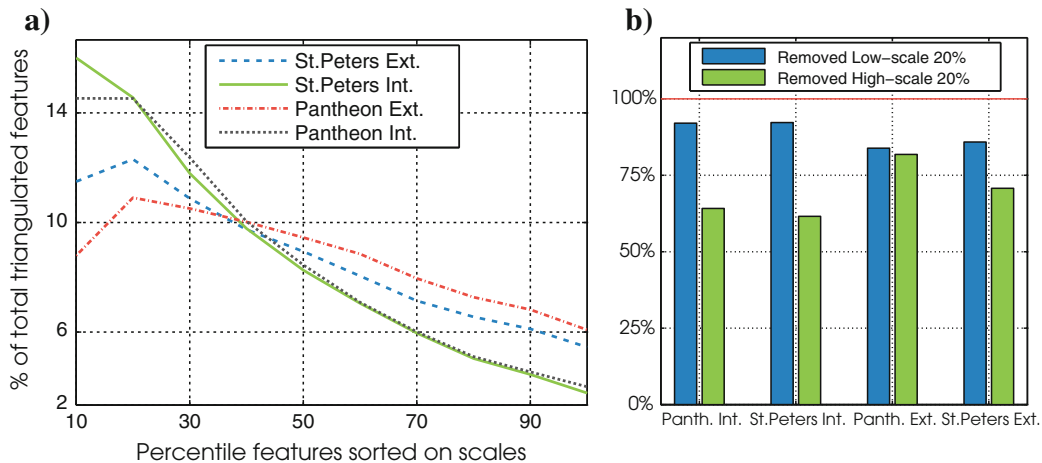

Fig. 3 Analysis of triangulated features by scales in reconstructed models: **a** illustrates the distribution of triangulated features versus their percentile scale rank, **b** illustrates the effect of removing high versus low-scale features on total number of triangulated points. These plots indicate that high-scale features participate more commonly in triangulated feature tracks and are clearly more important than low-scale features for reconstruction

for images with 10–30K features. The complexity for Kd tree based pairwise feature matching is $O(n^2 m \log m)$, for n images and m features per image. Most literature on SFM ignores m, assuming it to be a small constant. However, typical images have tens of thousands of features and m does have an impact on runtime in practice. Since we use only % of features, feature matching is nearly 100 times faster for components of ∼1000 images. Fewer features also have a notable advantage on practical runtime of bundle adjustment during reconstruction.

To further improve the efficiency, we adopt a hybrid matching scheme inspired by preemptive matching. Here, the first 10% of high-scale query features are matched with 20% features of the reference image. Then, the next 10% query features are matched only if sufficient matches are found (>4) in the first batch. Matching can be terminated early if sufficient matches for geometry estimation are found. Please note that for images with very few features (<1000), we simply use all features for matching. We use the Approximate Nearest Neighbor (ANN) library [6] for Kd tree based feature matching on CPU [6] and SIFTGPU library [55] for global feature matching on GPU.

We denote the recovered model as $\mathcal{M}_0 = \langle \mathbb{C}_0, \mathbb{P}_0 \rangle$, where $\mathbb{C}_0$ is the set of pose-estimated (localized) images and $\mathbb{P}_0$ is the set of 3D points mapped to triangulated feature tracks. The model $\mathcal{M}_0$ is *coarse* but *global*. That is, as compared to the full models reconstructed using all features, the 3D points in $\mathbb{P}_0$ are sparser but $\mathbb{C}_0$ covers most of the cameras. In our experiments, $\mathcal{M}_0$ contained 80–100% of the cameras of the full construction and roughly % of the 3D points. The coarse model, however, contains enough information to successfully add remaining cameras and points in subsequent stages.

3.2 Adding Cameras

After the reconstruction of the coarse model $\mathcal{M}_0$ in the first stage, some of the images would remain unregistered and a large number of features in all images would remain yet to be matched and triangulated. In this stage, we enrich the model $\mathcal{M}_0$ by registering the remaining images to the model. This stage is later repeated after the point addition stage if needed.

Registering a new image to an existing SFM model is known as the localization problem. For localization, it is necessary to establish correspondences between 3D points in the model and 2D features of the image to be localized. Once reliable and sufficient 3D–2D matches are established, the camera pose can be estimated using PnP solvers. Since a global model of the scene is known, localization of each image can be done independently of others. Unlike the traditional incremental SFM process where images can only be added to a growing reconstruction in a sequential manner, the coarse model allows us to localize all unregistered images simultaneously, and in parallel. Many methods have been proposed for efficient image localization [13, 25, 27, 37, 38]. These methods mainly di er in their strategies for 3D–2D

correspondence search. We have experimented with three different strategies for correspondence search in our pipeline.

Direct 3D–2D matching

In this method, each 3D point P_i in the model is represented by the mean SIFT descriptor of all features in its track, $Track(P_i)$ and queried into the Kd tree of all feature descriptors of the image to be localized. A match is declared by the ratio test. With this method for correspondence search, localization takes around 1–5 s to localize a single image, for models of about 100–200K 3D points.

Active correspondence search

In this technique proposed by [38], the pitfalls of both 3D-to-2D search and 2D-to-3D search are avoided by a combined approach. Moreover, the technique is made efficient by incorporating prioritized search based on visual words. This localization technique is significantly faster and superior in quality as compared to our previous technique. Localizing an image using active correspondence search takes between 0.5 and 1 s. The preprocessing steps for this search involve computing mean SIFT descriptors and visual word quantization which are easy to parallelize.

Ranked 2D–2D matching

Many image pairs with coarse feature matches do not participate in the SFM step for coarse model estimation due to either insufficient matches or inliers. Nevertheless, matches between the coarse features of an image pair offer an important cue that the images could visually overlap. We leverage this cue and propose a ranked 2D–2D matching scheme for 3D–2D correspondence search when active search fails. Given an unlocalized image, we find the top-K localized images sorted on the number of coarse feature matches they share with the unlocalized image. For each of the K localized images, we find the subset of 2D features that participate in tracks of 3D points in the current model and use these feature descriptors as proxies for 3D–2D matching. We create a Kd tree of all features in the unlocalized image, query the subset of 2D features of the nearby localized images into this Kd tree, and verify matches by ratio test, thereby establishing correspondences (>16) between the parent 3D points and the 2D features in the unlocalized image.

While the coarse model is typically small and localization is fast in the first iteration, the model after the first point addition stage gets heavy in 3D points for efficient localization in the later iterations. To avoid this, we use set cover representation of a 3D model, if it contains > 100K 3D points. The set cover of a model is a reduced set of points that cover each camera at least k (300–500) times [27]. Upon obtaining sufficient number of 3D–2D matches, RANSAC-based pose estimation and nonlinear pose refinement are performed, and finally the model is updated with all localized images.

By addition of newly localized cameras, the model $\mathcal{M}_i = \langle \mathbb{C}_i, \mathbb{P}_i \rangle$ upgrades to an intermediate model $\mathcal{M}_i^l = \langle \mathbb{C}_{i+1}, \mathbb{P}_i \rangle$. For each localized camera C_q, we have the inlier 3D–2D correspondences (P_j, f_k). We update all $Track(P_j)$s to contain (C_q, f_k) after adding each camera C_q. The new cameras each have a few points at this stage. More points are added for all pose-estimated cameras in the subsequent point addition stage.

3.3 Adding Points

The point addition stage updates the model $\mathcal{M}_i^l = \langle \mathbb{C}_{i+1}, \mathbb{P}_i \rangle$ to $\mathcal{M}_{i+1} = \langle \mathbb{C}_{i+1}, \mathbb{P}_{i+1} \rangle$ by triangulating several unmatched features of images in $\mathbb{C}_{i+1}$. The model after first camera addition stage is nearly complete in cameras but consists of points corresponding to only % high-scale features of localized cameras. After the first point addition step, the model is dense in points. This step is repeated after every round of camera addition to triangulate and merge features of the newly added cameras. This is done to ensure that unlocalized cameras can form 3D–2D connections with newly localized cameras too in the upcoming camera addition stage. To accelerate this stage, we leverage the known geometry of the existing model in the following two ways: (i) we use the visibility relations between localized cameras and triangulated coarse features to restrict feature matching to only pairs with sufficiently many co-visible points; and (ii) we use the epipolar geometry between the localized cameras to accelerate feature correspondence search. In the following sections, we explain these individual steps in detail.

3.3.1   Finding Candidate Images to Match

Given a set of images of a monument or a site, each image would find sufficient feature matches with only a small fraction of total images, those looking at common scene elements. Ideally, we would like to limit our search to only these *candidate* images. We use the point-camera visibility relations of the model $\mathcal{M}_1^l = \langle \mathbb{C}_1, \mathbb{P}_0 \rangle$ to determine whether or not two images are looking at common scene elements.

   Let I_q denote the query image and $F_q = \{f_1, f_2, ..., f_m\}$ denote the features that we wish to match and triangulate. Traditionally, we would attempt to match the features in image I_q with the features in set of all localized images I_L, where $I_L = \{I_i \mid C_i \in \mathbb{C}_1,\ C_i \neq C_q\}$. However, we wish to match the features in query image I_q with features in only a few *candidate* images that have co-visible points with image I_q. We define the set of all co-visible points between two images I_i and I_j as $P_{cv}(I_i, I_j) = Points(C_i) \cap Points(C_q)$. Using this visibility relation, we define the set of candidate images for image I_q as $S_q = \{I_i \mid |P_{cv}(I_q, I_i)| > T\}$ ($T = 8$ for our experiments). We select only top-k candidate images ranked based on the number of co-visible points. Our experiments show it is possible to converge to a full match graph of exhaustive pairwise matching even when the number of candidate images k is limited to only 10% of the total images. We find unique image pairs from candidate image sets for all query images and match these pairs in parallel using fast geometry-aware feature matching.

3.3.2   Fast Geometry-Aware Feature Matching

Given a query image I_q and its candidate set S_q, we use the guided matching strategy to match the feature sets $(F_q, F_c \mid I_c \in S_q)$. In traditional feature matching, each query feature in F_q is compared against features in a candidate image using a Kd tree of features in F_c.

Since query image I_q and candidate image I_c both are localized, their camera poses (epipolar geometry) are known. For a query feature point $p_q = (x_q\ y_q\ 1)$ in feature set F_q of image I_q, the corresponding epipolar line $l_q = (a_q, b_q, c_q)$ in image I_c is given by $l_q = F_{qc}\ p_q$. If $p'_q = (x'_q\ y'_q\ 1)$ denotes the corresponding feature point in image I_c, then as per the epipolar constraint $p'_q\ ^\top F_{qc} p_q = 0$, point p'_q must lie on the epipolar line, i.e., $p'_q\ ^\top l_q = 0$. Due to inaccuracies in estimation, it is practical to relax the constraint to $p'_q\ ^\top l_q < \epsilon$. To find the corresponding point p'_q, instead of considering all features in set F_c, we limit our search to only those features which are close to the epipolar line l_q. We define the set of candidate feature matches $\mathcal{C}$ as

$$\mathcal{C} = \{p'\ |\ \mathrm{dist}(p', l_q)\ \leq\ d\} \tag{1}$$

$$\mathrm{dist}(p', l_q) = \frac{a_q x' + b_q y' + c_q}{\sqrt{a_q{}^2 + b_q{}^2}} \tag{2}$$

We propose a fast algorithm for finding the set of candidate feature and propose an optimal strategy for correspondence search based on the dual nature of epipolar lines.

Grid-based search:

We optimize the candidate search by using a grid-based approach. We first divide the target image I_c into four overlapping grids of cell size $2d \times 2d$ and overlap of d along x, y, and x-y directions, as shown by dotted lines in Fig. 4c. We then bin all feature points of the set F_c into cells of the overlapping grids based on their image coordinates. Each feature point (x, y) would fall into four cells.

Given a query point p_q, we find its epipolar line l_q and the equidistant points (x_k, y_k). For each of the equidistant points on the epipolar line, we find the center most cell and accumulate all feature points binned into these cells to obtain an approximate set of candidate matches $\mathcal{C}'$. Red squares in Fig. 4c indicate the coverage of true candidate matches in set $\mathcal{C}$ by grid-based approximate search. In practice, we use a slightly larger grid size to account for misses due to grid approximation. Since feature points are binned only once per image, the time complexity for searching candidate matches is $O(1)$ in grid-based approach.

Further optimization

To finalize a match from candidate set $\mathcal{C}'$, a Kd tree of feature descriptors in $\mathcal{C}'$ is constructed, closest two features from the query are retrieved, and ratio test is performed. The number of candidate feature matches $|\mathcal{C}'|$ is a small fraction of total points $|F_c|$ (typically 200:1 in our experiments). Since the top two neighbors are searched in a much smaller Kd tree of size $|\mathcal{C}'|$, geometry-aware search reduces the operations required for two-image matching from $(|F_q| \log |F_c|)$ to $(|F_q| \log |\mathcal{C}'|)$, with an overhead of constructing a small Kd tree of size $|\mathcal{C}'|$ for each query feature.

To reduce the overhead of redundant Kd tree construction, we exploit the dual nature of epipolar lines, i.e., for all points that lie on line l in image I_c, their corresponding points must lie on the dual line l' in image I_q. We use this property, to group the query points in I_q whose epipolar lines intersect the boundaries of I_c in nearby

(a) Linear Search $O(|F_c|)$

(b) Radial Search $O(\log|F_c| \cdot |\cdot|)$

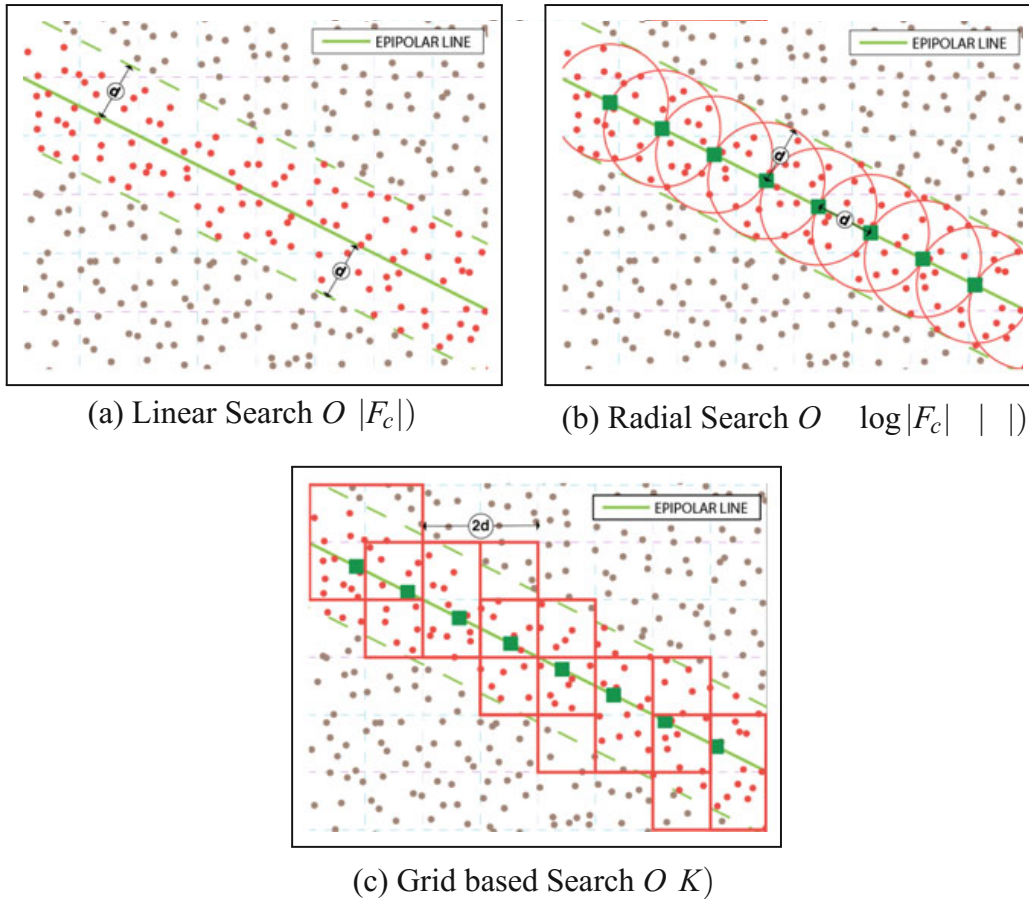(c) Grid based Search $O(K)$

Fig. 4 Illustration of geometry-aware search strategy. Search for points within distance d from epipolar line (shown by red dots) can be approximated by radial search and more efficient grid-based search. Red squares in (**c**) show the centermost cell of the overlapping grids selected for each equidistant point along the epipolar line

points (within two pixels) and search for matches group by group. Since all feature points in a group have the same epipolar line and hence the same set of candidate matches, we avoid redundant construction of the small Kd tree of size $|\cdot'|$ for points in a group.

Apart from being faster than geometry-blind global correspondence search, our grid-based geometry-aware search produces denser correspondences and is easy to parallelize.

3.4 Triangulation and Merging

After pairwise image matching is performed, we form tracks for features in a query image by augmenting matches found in all candidate images and triangulating these feature tracks using a standard least mean squared error method. We perform this

operation independently for all images. This would typically results in duplication of many 3D points because a triangulated feature pair (C_i, f_k) (C_j, f_l) for image C_i would also match and triangulate in reverse order for image C_j. Also, since we limited our matching to only candidate images, the longest track would only be as long as the size of the candidate set. We solve both of these problems in a track merging step. Our track merging step is similar to [44] and uses the standard sequential depth-first-search (DFS) algorithm to find connected components. It is possible to substitute our sequential implementation with a faster multi-core CPU or GPU implementation.

The models reconstructed using our multistage approach converge to full models reconstructed using all features and traditional pipelines in 1–2 iterations of the camera and point addition stages. Since we begin with a global coarse model, our method does not su er from accumulated drifts (for datasets observed so far), making incremental bundle adjustment optional in later stages of our pipeline. Please see Sect. 4 for a detailed discussion of these results.

4 Results and Discussion

We evaluate our method on several datasets and show qualitative results. For a detailed discussion on quantitative results and runtime performance, we encourage the interested reader to refer to [41].

Table 1 Datasets used in various experiments

Dataset	Label	#images	#feat (avg.)
Notre Dame Paris (subset)[1]	NDP	99	21K
Tsinghua School Building[2]	TSB	193	26K
Barcelona National Museum[3]	BNM	191	18K
Pantheon Interior[4]	PTI	587	9K
Pantheon Exterior[4]	PTE	782	13K
St. Peters Interior[4]	SPI	953	15K
St. Peters Exterior[4]	SPE	1155	17K
Hampi Vitthala Temple[5]	HVT	3017	39K
Cornell Arts Quad[6]	AQD	6014	18K

[1] [44] http://phototour.cs.washington.edu/datasets/
[2] [26] http://vision.ia.ac.cn/data/index.html
[3] [16] https://www.inf.ethz.ch/personal/acohen/papers/symmetryBA.php
[4] [27] http://www.cs.cornell.edu/projects/p2f/
[5] CVIT-IDH
[6] [17] http://vision.soic.indiana.edu/projects/disco/

Table 2 Comparison of runtime for match graph construction and number of 3D points in final models for three datasets. For NDP and TSB, all images (99 and 193, respectively) are registered for all methods. For BNM, Kd tree and cascadeHash matching based reconstructions register only 119 and 136 images, respectively, while geometry-aware matching registers 181 of 191 images. Also, point clouds for SFM with our method are denser

Dataset	Number of 3D points in reconstructions						Runtime for match graph construction				
	Kd tree		CascadeHash		Our		Kd tree	CascadeHash	Our (CPU)	Our (GPU)	SIFTGPU
	#pts	#pts3+	#pts	#pts3+	#pts	#pts3+	sec.	sec.	sec.	sec.	sec.
NDP	85K	46K	82K	43K	**109K**	**65K**	6504	1408	3702	171	999
TSB	178K	112K	180K	111K	**204K**	**132K**	27511	8660	8965	857	7019
BNM	39K	12K	40K	11K	**179K**	**77K**	18282	3662	5120	553	4799

Table 3 Statistics for baseline models reconstructed using Bundler with Kd tree based pairwise matching of all features for all image pairs. "#pairs" indicate the image pairs connected by co-visible 3D points. The columns under "cam. dists" indicate the average and median distances between the locations of the reconstructed cameras

Dataset	#cams	#pts (K)	#pts3+ (K)	#pairs	Repro. error		Cam. dists	
					Mean	Med.	Mean	Med.
PTI	574	126	57	66982	0.86	0.51	2.51	2.36
PTE	782	259	124	303389	0.76	0.49	0.811	0.78
SPI	953	301	140	227330	0.96	0.63	29.24	28.23
SPE	1155	380	180	575134	0.70	0.47	3.09	1.99
AQD	5147	–	1402	538131	0.41	0.30	179.01	172.77

4.1 Evaluation of Geometry Aware Matching

To evaluate the e ectiveness of geometry-aware matching, we perform 3D reconstruction using Bundler from match graphs constructed using Kd tree based matching, cascade hashing based matching (CascadeHash), and two-stage geometry-aware matching methods for three small datasets (NDP, TSB, and BNM in Table 1). Since geometry-aware method depends upon the coarse epipolar geometry between the image pairs, we first match the high-scale 20% features using Kd tree based matching and estimate pairwise fundamental matrices from the initial matches using DLT and RANSAC. The estimated fundamental matrices are then directly used for geometry-aware matching of the unmatched features as explained in Sect. 3.3.2. For this set of experiments, SFM on coarse match graph is not performed.

Table 2 compares the match graph construction time and the number of 3D points in the final reconstruction for the three methods. Geometry-aware matching clearly outperforms other methods. Figure 5 shows the reconstruction of BNM dataset with unguided matching and geometry-aware matching. Table 3 provides a quantitative baseline of Kd tree based matching of all image pairs on various datasets. The reconstruction with geometry-aware matching is more complete compared to other methods.

4.2 Evaluation of Multistage SFM Pipeline

To evaluate our multistage SFM pipeline, we reconstruct components of ∼500–1000 images from Rome16K dataset (PTI, PTE, SPI, and SPE in Table 1). We also use our pipeline to reconstruct two large datasets of multiple structures, HVT and AQD. In practice, such large datasets should be divided into multiple components similar to [8, 18] and our pipeline should only be used reconstruct individual components to be merged later.
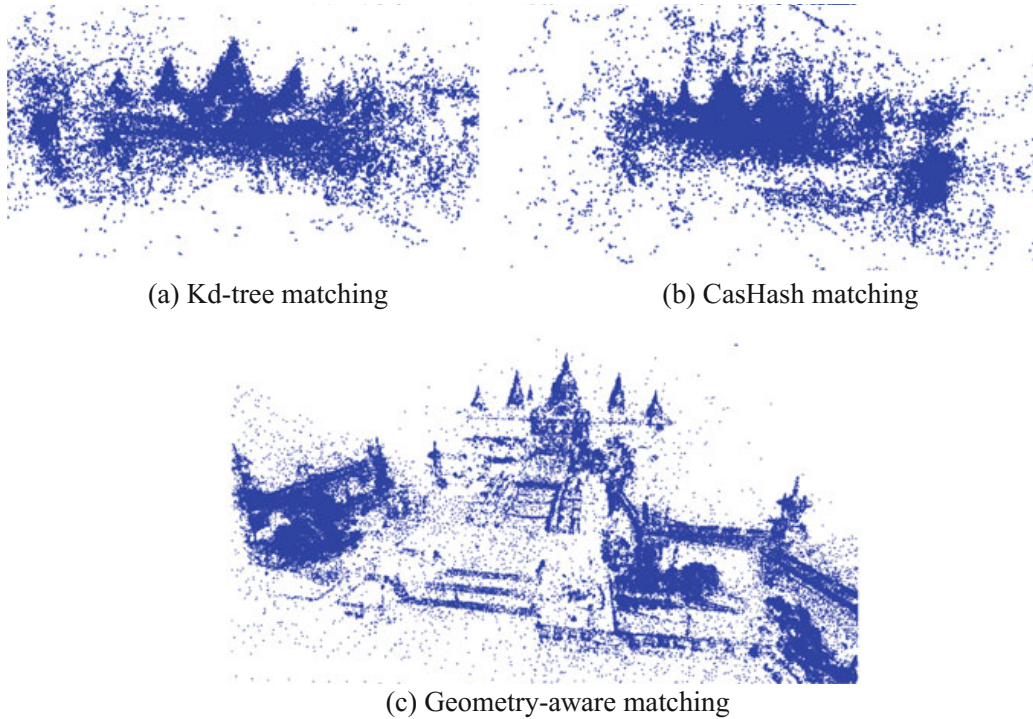
(a) Kd-tree matching   (b) CasHash matching

(c) Geometry-aware matching

Fig. 5 Reconstructions of BNM dataset using undler with three diﬀerent match graphs produced by **a** Kd tree matching, **b** CascadeHash matching, and **c** geometry-aware matching. Models with unguided matching (**a**,**b**) remain incomplete
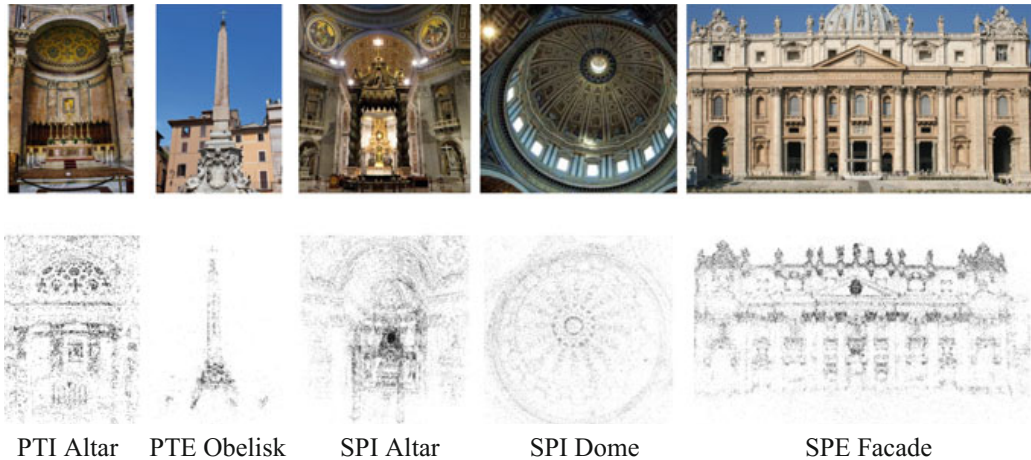

PTI Altar   PTE Obelisk   SPI Altar   SPI Dome   SPE Facade

Fig. 6 Selected renders of reconstructed components of Rome16K dataset

For coarse model reconstruction of datasets, we use three popular SFM implementations: (i) Bundler BDLR) [44], (ii) VisualSFM VSFM) [57], and (iii) Theia [49] with incremental SFM (i & ii) and global SFM (iii) as underlying pipelines. To demonstrate that the traditional preprocessing steps of large-scale SFM pipeline can be used in connection with our method, we also create match graphs by matching coarse features for only the image pairs selected using preemptive matching [56] and later perform reconstruction using VisualSFM.

HVT oor plan

top-view of the full reconstruction

coarse reconstruction

full reconstruction (multistage pipeline)

Utsava Mandapa : coarse (left), full (right)

North Gopura : coarse (left), full (right)

Fig. 7 Point clouds of Hampi Vitthala Temple Complex reconstruction

We observed that all coarse reconstruction methods are able to register between 70 and 100% cameras w.r.t. the baseline models. The coarse models are also fairly accurate as compared to the baseline models. The mean and median rotation errors w.r.t the baseline models are less than 0.05 degree for most models. The mean and median relative translation errors are also below 2% of the mean/med. camera distances of the baselines models. We enrich and complete the coarse models using two iterations of camera and point addition stages of our multistage pipeline as explained in Sects. 3.2 and 3.3. Despite being initialized with di erent coarse models, all final models are nearly complete in number of cameras and have higher or comparable number of 3D points w.r.t. the baseline models, except for the AQD model.

The coarse and final reconstructions for these datasets are shown in Figs. 6 and 7. For HVT coarse reconstruction, we only use V+PM method due to practical constraints imposed by its large scale.

As shown before, feature matching using multistage strategy has significant runtime advantage. Moreover, despite using all images, SFM with coarse match graph is 2–5 times faster than SFM with full match graph. Since our method employs the SFM step for only coarse features based match graph, it is clearly more advantageous. Furthermore, we do not perform incremental BA during later stages of the pipeline; the runtimes for feature matching and coarse global model estimation dominate the total runtime of our pipeline. In comparison, the runtime for the remaining steps of our pipeline is mostly trivial.

5   Conclusion

In this chapter, we presented a multistage approach as an alternative to match graph construction and structure from motion steps of the traditional large-scale 3D reconstruction pipeline. The proposed approach provides an opportunity to leverage the constraints captured by the coarse geometry for making the rest of the processing more efficient and parallel. We evaluated our method comprehensively and showed that it can produce similar or better quality reconstructions as compared to the traditional methods while being notably fast. Study of real-time reconstruction applications with multistaging and the possibility of extending the framework for performing fast multi-view stereo are interesting leads that could be explored.

References

1. Agarwal S, Mierle K (2010) Others: Ceres solver. http://ceres-solver.org
2. Agarwal S, Snavely N, Seitz SM, Szeliski R (2010) Bundle adjustment in the large. In: Proceedings ECCV
3. Agarwal S, Snavely N, Simon I, Seitz SM, Szeliski R (2009) Building rome in a day. In: Proceedings ICCV
4. Agrawal A, Raskar R, Chellappa, R (2006) What is the range of surface reconstructions from a gradient field?. In: Proceedings of the European Conference on Computer Vision
5. Aharon M, Elad M, Bruckstein A (2006) K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Signal Process
6. Arya S, Mount DM, Netanyahu NS, Silverman R, Wu AY (1998) An optimal algorithm for approximate nearest neighbor searching fixed dimensions. J ACM 45(6)
7. Barron J, Malik J (2012) Color constancy, intrinsic images, and shape estimation. In: Proceedings of the European Conference on Computer Vision
8. Bhowmick B, Patra S, Chatterjee A, Govindu V, Banerjee S (2014) Divide and conquer: Efficient large-scale structure from motion using graph partitioning. In: Proceedings ACCV, pp. 273–287
9. Brown M, Lowe D (2005) Unsupervised 3d object recognition and reconstruction in unordered datasets. In: 3-D Digital Imaging and Modeling
10. Byröd M, Åström K (2010) Conjugate Gradient Bundle Adjustment
11. Cao S, Snavely N(2012) Learning to match images in large-scale collections. In: Proceedings ECCV Workshop
12. Chatterjee A, Govindu VM (2013) Efficient and robust large-scale rotation averaging. In: 2013 IEEE ICCV
13. Choudhary S, Narayanan P (2012) Visibility probability structure from SfM datasets and applications. In: Proceedings ECCV
14. Chum O, Matas J (2010) Large-scale discovery of spatially related images. IEEE Trans Pattern Anal Mach Intell 32(2):371–377
15. Cohen A, Sattler T, Pollefeys M (2015) Merging the unmatchable: Stitching visually disconnected SfM models. In: Proceedings IEEE ICCV
16. Cohen A, Zach C, Sinha S, Pollefeys M (2012) Discovering and exploiting 3d symmetries in structure from motion. In: Proceedings CVPR
17. Crandall D, Owens A, Snavely N, Huttenlocher D (2011) Discrete-continuous optimization for large-scale structure from motion. In: Proceedings IEEE CVPR
18. Frahm JM, Fite-Georgel P, Gallup D, Johnson T, Raguram R, Wu C, Jen YH, Dunn E, Clipp B, Lazebnik S, Pollefeys M (2010) Building rome on a cloudless day. In: Proceedings ECCV
19. Gherardi R, Farenzena M, Fusiello A (2010) Improving the efficiency of hierarchical structure-and-motion. In: Proceedings IEEE CVPR
20. Hartley R, Zisserman A (2003) Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge
21. Hartmann W, Havlena M, Schindler K (2014) Predicting matchability. Proceedings IEEE CVPR. CVPR '14. IEEE Comput Society, Washington, DC, USA, pp 9–16
22. Havlena M, Schindler K (2014) Vocmatch: Efficient multiview correspondence for structure from motion. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds.) Proceedings ECCV 2014
23. Havlena M, Torii A, Knopp J, Pajdla T (2009) Randomized structure from motion based on atomic 3d models from camera triplets. In: Proceedings IEEE CVPR
24. Havlena M, Torii A, Pajdla T (2010) Efficient structure from motion by graph optimization. In: Proceedings ECCV 2010
25. Irschara A, Zach C, Frahm JM, Bischof H (2009) From structure-from-motion point clouds to fast location recognition. In: Proceedings IEEE CVPR
26. Jian C, Cong L, Jiaxiang W, Hainan C, Hanqing L (2014) Fast and accurate image matching with cascade hashing for 3d reconstruction. In: Proceedings IEEE CVPR

27. Li Y, Snavely N, Huttenlocher DP (2010) Location recognition using prioritized feature matching. In: Proceedings ECCV
28. Lou Y, Snavely N, Gehrke J (2012) Matchminer: Efficient spanning structure mining in large image collections. In: Proceedings ECCV
29. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vision 60(2)
30. Moulon P, Monasse P, Marlet R (2013) Global fusion of relative motions for robust, accurate and scalable structure from motion. In: IEEE ICCV
31. Muja M, Lowe DG (2014) Scalable nearest neighbor algorithms for high dimensional data. IEEE Trans Pattern Anal Mach Intel 36
32. Olsson C, Enqvist O (2011) Stable structure from motion for unordered image collections. In: Proceedings of the 17th Scandinavian conference on Image analysis, ser. SCIA11, pp 524–535
33. Panagopoulos A, Hadap S, Samaras D (2012) Reconstructing shape from dictionaries of shading primitives. In: Proceedings of the Asian Conference on Computer Vision
34. Petschnigg G, Szeliski R, Agrawala M, Cohen M, Hoppe H, Toyama K (2004) Digital photography with flash and no-flash image pairs. In: Proceedings of the ACM SIGGRAPH
35. Ping-Sing T, Shah M (1994) Shape from shading using linear approximation. Image Vision Comput 12(8):487–498
36. Raguram R, Wu C, Frahm JM, Lazebnik S (2011) Modeling and recognition of landmark image collections using iconic scene graphs. Intern J Comput Vision 95(3):213–239
37. Sattler T, Leibe B, Kobbelt L (2011) Fast image-based localization using direct 2d-to-3d matching. In: Proceedings IEEE ICCV
38. Sattler T, Leibe B, Kobbelt L (2012) Improving image-based localization by active correspondence search. In: Proceedings ECCV
39. Schönberger JL, Berg AC, Frahm JM (2015) Paige: Pairwise image geometry encoding for improved efficiency in structure-from-motion. In: IEEE CVPR
40. Shah R, Deshpande A, Narayanan PJ (2014) Multistage sfm: Revisiting incremental structure from motion. In: International Conference on 3D Vision (3DV), vol. 1, pp. 417–424
41. Shah R, Deshpande A, Narayanan PJ (2015) Multistage SFM: A Coarse-to-Fine Approach for 3D Reconstruction. In:CoRR (2015)
42. Shah R, Srivastava V, Narayanan PJ (2015) Geometry-aware feature matching for structure from motion applications. In: IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 278–285
43. Sinha S, Steedly D, Szeliski R (2010) A multi-stage linear approach to structure from motion. In: Proceedings ECCV RMLE Workshop
44. Snavely N, Seitz SM, Szeliski R (2006) Photo tourism: Exploring photo collections in 3d. ACM Trans Graph 25(3)
45. Snavely N, Seitz SM, Szeliski R (2008) Modeling the world from internet photo collections. Int J Comput Vision 80(2)
46. Snavely N, Seitz SM, Szeliski R (2008) Skeletal graphs for efficient structure from motion. In: Proceedings IEEE CVPR
47. Soman J, Kothapalli K, Narayanan PJ (2010) Some GPU algorithms for graph connected components and spanning tree. Parallel Process Lett 20(04)
48. Sturm PF, Triggs B (1996) A factorization based algorithm for multi-image projective structure and motion. In: Proceedings of the 4th European Conference on Computer Vision, ECCV '96, pp 709–720
49. Sweeney C (2015) Theia Multiview Geometry Library: Tutorial & Reference. University of California, Santa Barbara
50. Szeliski R, Kang SB (1993) Recovering 3d shape and motion from image streams using non-linear least squares. In: Proceedings CVPR '93, 1993 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 752–753 (1993)
51. Taylor C, Kriegman D, Anandan P (1991) Structure and motion in two dimensions from multiple images: a least squares approach. In Proceedings of the IEEE Workshop on Visual Motion, pp 242–248

52. Tomasi C, Kanade T (1992) Shape and motion from image streams under orthography: a factorization method. Intern J Comput Vision 9(2):137–154
53. Triggs B, McLauchlan P, Hartley R, Fitzgibbon A (2000) Bundle adjustment a modern synthesis. In: Triggs B, Zisserman A, Szeliski R (eds.) Vision Algorithms: Theory and Practice, vol. 1883, pp 298–372
54. Wilson K, Snavely N (2014) Robust global translations with 1DSfM. In: Proceedings ECCV
55. Wu C (2007) SiftGPU: A GPU implementation of scale invariant feature transform (SIFT). http://cs.unc.edu/~ccwu/siftgpu
56. Wu C (2013) Towards linear-time incremental structure from motion. In: 3DV Conference
57. Wu C, Agarwal S, Curless B, Seitz SM (2011) Multicore bundle adjustment. In: Proceedings IEEE CVPR
58. Zhang R, Tsai P, Cryer J, Shah M (1999) Shape-from-shading: A survey. IEEE Transac Pattern Anal Mach Intel

Realistic Walkthrough of Cultural Heritage Sites

Uma Mudenagudi Syed Altaf Ganihar and Shankar Setty

1 Introduction

In this chapter, we describe the framework for the generation of realistic digital walkthrough of cultural heritage sites. With the availability of various 3D data acquisition techniques and with the increase in computational power have made the digital reconstruction and realistic rendering of 3D models a topic of prime interest. With the advent of digital technology, there is a great surge of interest among the computer graphics and vision community in digital restoration and preservation of cultural heritage sites [14, 19]. A large number of cultural heritage sites are deteriorating or being destroyed over a period of time due to natural weathering, natural disasters, and wars. The heritage sites in Hampi, India are largely composed of rock structures which are in a grievous situation as can be seen in Fig. 1a and this necessitates the need for reconstruction of 3D models. The digital preservation of cultural heritage sites is an area of application of 3D reconstruction. However, many challenges still exist during building of the 3D reconstruction pipeline, as presented by [3, 14, 28].

Digital restoration of cultural heritage sites has been in the purview of computer graphics and vision research since a long time. Some of the prominent works in literature sought to overcome the problems pertaining to digital reconstruction of complex 3D models at heritage sites are as follows: The Stanford University's Michelangelo project [19] describes a hardware and software system for digitizing (acquiring, aligning, merging, and viewing scanned large data) the shape and color of large fragile objects under non-laboratory conditions. The Minerva project [6] which

U. Mudenagudi ( )   S. A. Ganihar   S. Setty
B.V.B. College of Engineering and Technology, Hubli, India
e-mail: uma@bvb.edu

S. A. Ganihar
e-mail: altafganihar@gmail.com
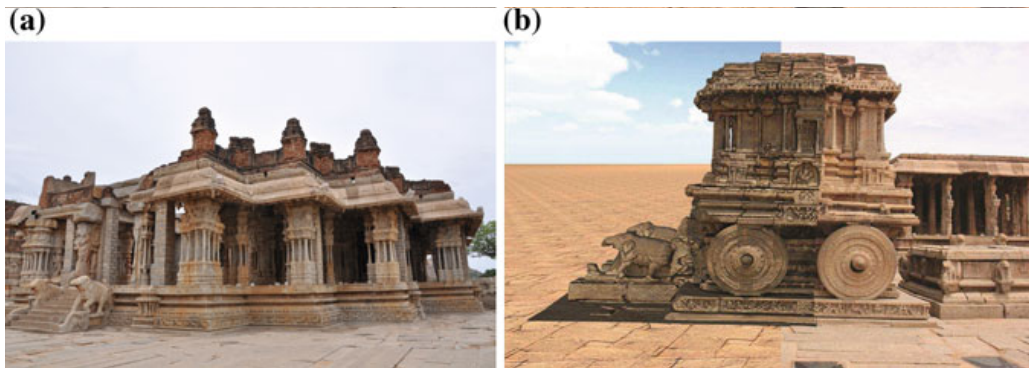
S. Setty
e-mail: shankar@bvb.edu

Fig. 1 **a** The ruins at Vittala Temple, Hampi, India and **b** comparison of the rendered scene and the original image of the Stone Chariot at Vittala Temple, Hampi: left half is the rendered image and the right half is the original image

is related to Minerva of Arezzo describes the usage of high-resolution 3D triangulation laser scanner for keeping track of the variations that occur during the restoration process of the 3D models. The IBM's Pieta project [29] describes the usage of triangulation scanner and a video camera for enhancing texture for digitization of a large marble sculpture created by Michelangelo. The Beauvais Cathedral project [1] presented a new method for automatic registration of range images and texture mapping between point clouds and 2D image. In the Angkorian Temples project [27], acquisition of images was aerial and reconstruction is based on photogrammetric 3D modeling. The Great Buddha of Bamiyan project [13] presents methods in digitization of two tall Buddha statues. The Great Buddha Project [14] describes the full pipeline, consisting of acquiring data, aligning data, aligning multiple range images, and merging of range images and challenges for the digital preservation and restoration of Great Buddha statues existing in the outdoor environment. Columbia University's French cathedral project [26] describes building of a system which can automatically acquire 3D range scans and 2D images to build 3D models of urban environments. The 3D digitization and its applications in cultural heritage reported in [12, 20] o er comprehensive survey on 3D reconstruction methods for digital preservation of cultural heritage sites.

The acquisition of the 3D data is an integral step in the digital preservation of the cultural heritage sites. The classic 3D modeling tools are often derisory to accurately portray the complex shape of sculptures found at cultural heritage sites. The advent of inexpensive 3D scanning devices like Microsoft Kinect and Time of Flight (ToF) cameras has simplified the 3D data acquisition process. The state-of-the-art 3D laser scanning devices generate very accurate 3D data of the objects. However, the scanning of large outdoor objects at the cultural heritage sites invites a lot of tribulations due to the generation of partial 3D models. The image-based methods like SFM (Structure from Motion) [25] and PMVS (Patch based Multi-View Stereo) [7] consolidate the 3D data acquisition process but do not generate high-resolution 3D data to accurately depict the artwork at the heritage sites. The occlusions during the scanning process result in the occurrence of missing regions or holes in the 3D

data and generation of partial models. Also with the rapid growth of large collection of 3D models on the internet, it is essential to categorize 3D models for better search and retrieval. This warrants the need for efficient data processing techniques and realistic interactive rendering for the digital preservation of the cultural heritage sites.

The comparison of the rendered scene of the Stone Chariot at Vittala Temple, Hampi with the original image of the scene is shown in Fig. 1b. Our framework generates a realistic walkthrough of the cultural heritage sites using a coarse to detail 3D reconstruction pipeline. The proposed 3D reconstruction pipeline consists of three stages, viz., 3D data acquisition, data processing, and interactive rendering of 3D models. The 3D models are acquired from different modalities like computer-aided design (CAD), single-view reconstruction, Kinect scanning, and multi-view reconstruction. We put into service several data processing algorithms like noise filtering, 3D super resolution, 3D hole filling, and 3D object categorization for the detail level 3D reconstruction of the models. The detail level 3D reconstructed models at the cultural heritage sites are registered with the coarse level models to generate a coarse to detail 3D reconstructed model. The coarse to detail 3D reconstructed models are subsequently rendered to obtain a digitally realizable walkthrough of the heritage site. Towards this we make the following contributions:

1. We propose a framework for the generation of realistic walkthrough of cultural heritage sites with coarse to detail level 3D reconstruction.
2. We propose 3D super resolution, hole filling, and object categorization algorithms for efficient data processing using concepts of Riemannian geometry with metric tensor and Christoffel symbols as a novel set of features.
3. We propose rendering of reconstructed 3D models using rendering engine in an attempt to restore the original appearance of cultural heritage sites.
4. We demonstrate the proposed framework for artifacts at Vittala Temple, Hampi, India.

The remainder of this chapter is organized as follows. In Sect. 2, we present the 3D reconstruction framework. In Sect. 3, we demonstrate the results of the proposed framework and conclude in Sect. 4.

2 The Realistic Walkthrough: 3D Reconstruction Framework

The proposed 3D reconstruction framework to generate realistic walkthrough [22] at cultural heritage sites is as shown in Fig. 2. The proposed framework consists of the following stages: (i) 3D Data acquisition—includes acquisition of 3D data of different modalities like CAD, single-view reconstruction, multi-view reconstruction, and Kinect-based reconstruction. (ii) 3D Data processing—includes various data processing steps like outlier filtering, extraction of geometric features, super
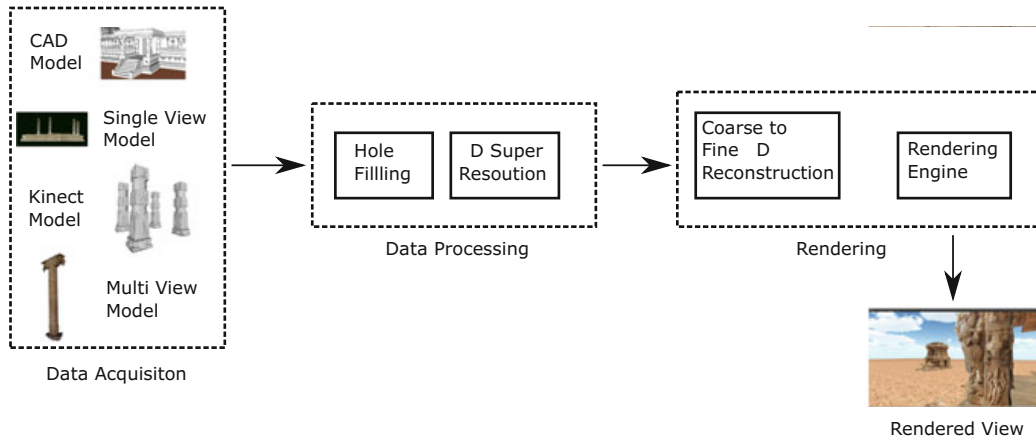
Fig. 2 An overview of 3D reconstruction framework

resolution, hole filling, and object categorization to refine the point cloud data. (iii) Coarse to detail 3D reconstruction and interactive rendering—the refined data is 3D reconstructed using a coarse to detail 3D reconstruction framework and the reconstructed models are rendered using rendering engine for the generation of the realistic walkthrough at cultural heritage sites.

2 1   3D Data Acquisition

The 3D data acquisition of the cultural heritage sites is the process of capturing 3D models from the on-site real-world objects and is an important part of the digital restoration process. The coarse level models are obtained either using CAD modeling tools or using single-view reconstruction. The CAD models obtained do not accurately depict the geometry of the artwork at the cultural heritage sites. The CAD models while modeling are recreated or restored in order to incorporate some of the missing, withered or prophesied part of the cultural heritage site. The CAD models or the single-view reconstructed models do not accurately portray the artworks at the cultural heritage sites. The detail level models are, hence, required to precisely represent the artworks. We acquire the detail level models at the cultural heritage sites in the following ways depending upon the location, size, and feasibility of the method.

1. The Microsoft Kinect 3D sensor consisting of a depth and an RGB camera is employed to scan the 3D models. Under appropriate lighting conditions, scanning is done on a 3D model and we use the Kinect Fusion (KinFu) [15] to generate a dense point cloud of the scanned model.
2. A set of images of a 3D model to be reconstructed is captured under appropriate lighting conditions. The images are then fed to dense reconstruction algorithms like SFM [25] or PMVS [7] to generate point cloud models.

2 2   Data Processing

The data processing algorithms are a vital component in the digital restoration of cultural heritage sites. The acquired data is in the form of a point cloud which is filtered using statistical outliers filter in order to eliminate any noisy data acquired during the scanning process. The data acquired using scanners like laser scanners, Microsoft Kinect, or image-based methods comprises certain missing regions (holes), partial models, or low resolution models. To address these issues, we propose geometry-based data processing algorithms such as 3D data super resolution, hole filling, and object categorization. The pipeline for the generation of detail level models is as shown in Fig. 3. In this section, we discuss the problems addressed during the data processing stage.

2.2.1   Geometric Features and Decomposition Framework

We observe that the captured 3D point cloud at cultural heritage sites exhibits nonuniform distribution of geometrical properties. The portrayal of these nonuniform geometrical properties is essential for the accurate depiction of the models to generate *realistic walkthrough* of cultural heritage sites. Most of the low-cost 3D data acquisition techniques as discussed in Sect. 2.1 fail to accurately capture the geometric properties. In order to address this issue, we propose to use Riemannian metric tensor together with Christo el symbols as the geometric features from our previous work [10].

Fig. 3 Overview of the data acquisition and data refinement pipeline for detailed reconstruction

Our work in [10] is summarized as follows: Given a 3D object V x, y, z) in Euclidean space, we model it as a set of unique discretized piecewise Riemmanian manifolds M, g) [16] in geometric space to account for discontinuities in the geometry. A Riemannian manifold is a real smooth differential manifold M equipped with an inner product g on the tangent space at each point p. Metric tensor g on the smooth manifold alone cannot represent the inherent local geometry of the 3D model as the 3D model exhibits non-smooth behavior at certain positions due to nonuniform geometry. So Christoffel symbols are considered for the deviation in the geometric properties from neighboring manifolds. Thus, 3D model under consideration is uniquely represented by the pair of tensors g,). Finally, the pair of tensors g,) of 3D model are decomposed into basic shapes, viz., sphere, cone, and cylinder in the basis space which is further used for solving data processing problems such as super resolution, hole filling, and object categorization.

2.2.2   Super Resolution

The point cloud data obtained from the low-resolution 3D scanner like the Microsoft Kinect or from sparse reconstruction algorithms usually fail to capture the accurate geometric properties and detailed structure of the 3D model either due to the presence of occlusions during the scanning process, non-feasibility of the sparse reconstruction algorithm, or adverse scanning environment. As a result, these techniques fail to portray all the details in a model's surface resulting in a low-resolution point cloud data. The generation of high-resolution 3D data is important for the realistic rendering of cultural heritage sites. Hence, there is an immense requirement to produce a high-resolution point cloud data from a given low-resolution point cloud data. We propose to solve the problem of 3D super resolution by proposing a selective super-resolution technique for super resolving the 3D model. The overview of the proposed 3D super-resolution technique is as shown in Fig. 4.

Given 3D model is modeled as a set of Riemannian manifolds [16] in continuous and discretized space. A kernel-based support vector machine (SVM) learning framework [4] is employed to decompose a given 3D model into basic shapes, viz., sphere, cone, and cylinder using metric tensor and Christoffel symbols as a set of novel geometric features [8, 9]. The decomposed models are then independently super-resolved using selective interpolation techniques. Consider, for example, the spherical decomposed part of the 3D model is interpolated using spherical surface interpolation method. Similarly, the conical and the cylindrical decomposed parts are interpolated using conical surface interpolation and cylindrical surface interpolation methods, respectively. The independently super-resolved algorithms are merged to obtain the final super-resolved model.

2.2.3   Hole Filling

The 3D data acquired using the proposed techniques consists of missing regions or holes due to occlusions in the surface to be scanned. To address this issue, we propose

a hole filling algorithm using metric tensor and Christo el symbols as features. The overview of the proposed hole filling algorithm is as shown in Fig. 5. The holes are identified by using the boundary detection algorithm used in [21]. The neighborhood of the hole is decomposed into basic shapes, viz., sphere, cone, and cylinder using a kernel-based SVM learning framework with metric tensor and Christo el symbols as features [24]. The decomposed regions in the neighborhood of the hole are interpolated using selective surface interpolation techniques. The centroid of the hole region is computed and the selective surface interpolation is carried out along the directional vector for best surface fit to recover missing regions.

Fig. 4 Proposed 3D super resolution algorithm

Fig. 5 Proposed hole filling algorithm

2.2.4 Object Categorization

Today, with the presence of inexpensive scanning devices, capturing large collections of 3D data is growing at a rapid rate, and this has increased the interest in 3D search and object retrieval. In the context of cultural heritage sites, the artifacts and sculptures at cultural heritage sites show similarities towards other sculptures or artifacts (at the same or di erent locations) and hence categorizing the artifacts based on their geometric properties facilitates the management of the 3D data. The geometric categorization of the sculptures or artifacts provides insights into the history of the artwork as well as the techniques employed for their construction.

The implementation of the proposed categorization framework is as shown in Fig. 6. We propose a learning framework to categorize 3D models into predefined classes using kernel-based SVM by capturing the geometry features metric tensor and Christo el symbols [10, 11]. The geometric features are aggregated into a local patch based approach and a Bag of Words (BoW) approach (see [10]) which are then fed into SVM for categorization. In local patch based approach, each local patch comprising 12 points is created. Based on the local patches, 3D models are categorized into appropriate predefined classes. In BoW approach, a vocabulary set is built by using k-means clustering algorithm for the features on training dataset. The vocabulary is used to compute histogram on the set of features. For testing, the one-against-all method is adopted. Thus, the testing data fed to learning framework based on local patch based and BoW approach classifies the 3D models into predefined set of categories.
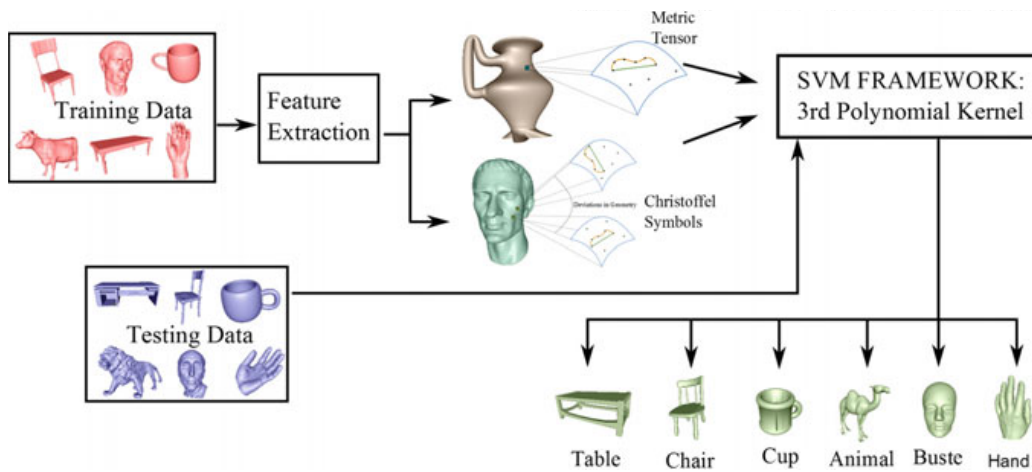

Fig. 6 Proposed 3D object categorization technique

2 3 Coarse to Detail 3D Reconstruction and Interactive Rendering

In this section, we present the coarse to detail 3D reconstruction and the rendering of the reconstructed models to generate digital walkthrough. The refined point cloud data generated after the 3D data processing step is surface reconstructed using Poisson surface reconstruction [17] or Ball-pivoting surface reconstruction algorithm [2]. The surface reconstructed model is texture-mapped using image alignment with mutual information [5] and parameterization of the registered rasters. The texture-mapped models are fed to coarse to detail level 3D reconstruction framework to generate the final 3D reconstructed model of the cultural heritage site.

We carry out coarse level 3D reconstruction using methods such as single-view 3D reconstruction [18] or from modeling tools. The models generated using modeling tools and single-view reconstruction do not accurately portray the geometrical complexities of the artwork at the cultural heritage sites. However, the detail level 3D reconstruction of large-scale outdoor objects is not feasible using the data processing techniques. To resolve this issue, we propose a coarse to detail level 3D reconstruction of the cultural heritage sites. The coarse to detail level 3D reconstruction is achieved by registering the coarse level 3D models with the detail level 3D models. The detail level 3D models are superimposed on the coarse level 3D models by interactively selecting the correspondence points in the model. The coarse and detail level 3D models are subsequently registered using the iterative closest point (ICP) algorithm [23] for the corresponding points as shown in Fig. 7a.

The coarse to detail 3D reconstructed models are subsequently rendered using real-time interactive rendering engines to obtain a digital version of the cultural heritage sites. The rendering of the reconstructed models is carried out using either a rendering engine like OGRE 3D or a gaming engine like Unity 3D to provide a real-time interactive rendering experience of the cultural heritage site.
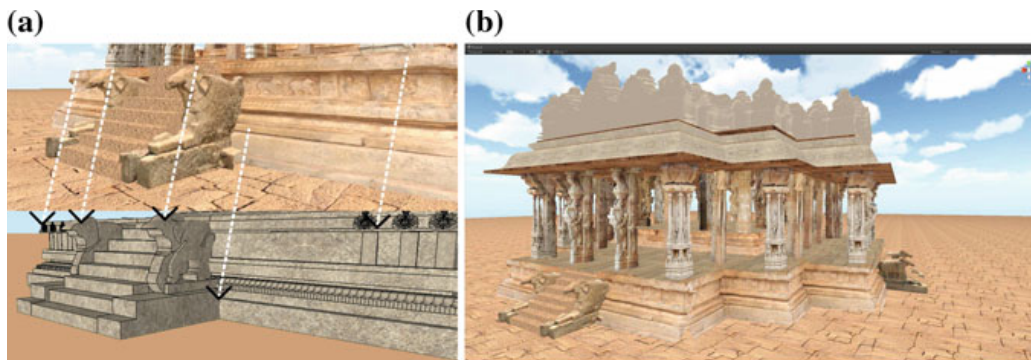

Fig. 7 **a** Coarse to detail reconstruction of 3D model using ICP registration with corresponding points in coarse model and detail model. Upper part of the image shows the detail level reconstruction model and lower part of the image shows the coarse level reconstruction model. **b** Results of coarse to detail 3D reconstructed model of Kalyana Mantapa, Vittala Temple, Hampi

3   Results and Discussions

In this section, we demonstrate the proposed 3D reconstruction framework on Kalyana Mantapa and Stone Chariot artifacts at Vittala Temple, Hampi, India (see Fig. 2 of Chap. 5). The data processing algorithms are implemented on Intel(R) Xeon(R) CPU E5-2665 @2.40 GHz (16 CPU's) and 64 GB RAM with NVIDIA Quadro K5000 graphics, 4 GB DDR3 graphics memory.

3 1   Data Acquisition

The coarse level 3D models at the cultural heritage sites obtained using single-view reconstruction techniques or using modeling tools are as shown in Fig. 8. The detail level models obtained using 3D scanning devices like laser scanner, Microsoft Kinect, and image-based methods like PMVS are as shown in Fig. 9.

3 2   Data Processing

The detail level 3D models are processed using the proposed 3D super resolution, hole filling, and object categorization algorithms. The processed 3D models are then surface reconstructed using the Poisson surface reconstruction algorithm and subsequently textured mapped using image alignment mutual information and registration of rasters. The super-resolved models generated for di erent artifacts of one of the pillars at Main Mantapa, Hampi with a magnification factor of approximately 2 are shown in Fig. 10. We observe that the low-resolution artifacts with 9,324, 11,514, and 10,601 points are super-resolved to 19,293, 23,473, and 20,969.

The hole fillings for a part of the Stone Chariot at Vittala Temple, Hampi and for a part of the artifact of one of the pillars at Main Mantapa, Hampi are as shown in Fig. 11a and b. More experimental analysis on hole filling for synthetically generated missing regions on 3D artifacts from pillars at Kalyana Mantapa is as shown in Fig. 12 and for real missing regions on point clouds with di erent complexities and sizes from Stone Chariot at Vittala Temple, Hampi shown in Fig. 13.

In object categorization, a dataset of pillars is trained to create a learning model comprising of the basic four categories: Yali, sculpted pillar, core pillars, and core with two or more small pillars as shown in Fig. 14. The testing results obtained after the object categorization of pillars into predefined classes are also shown in Fig. 14.
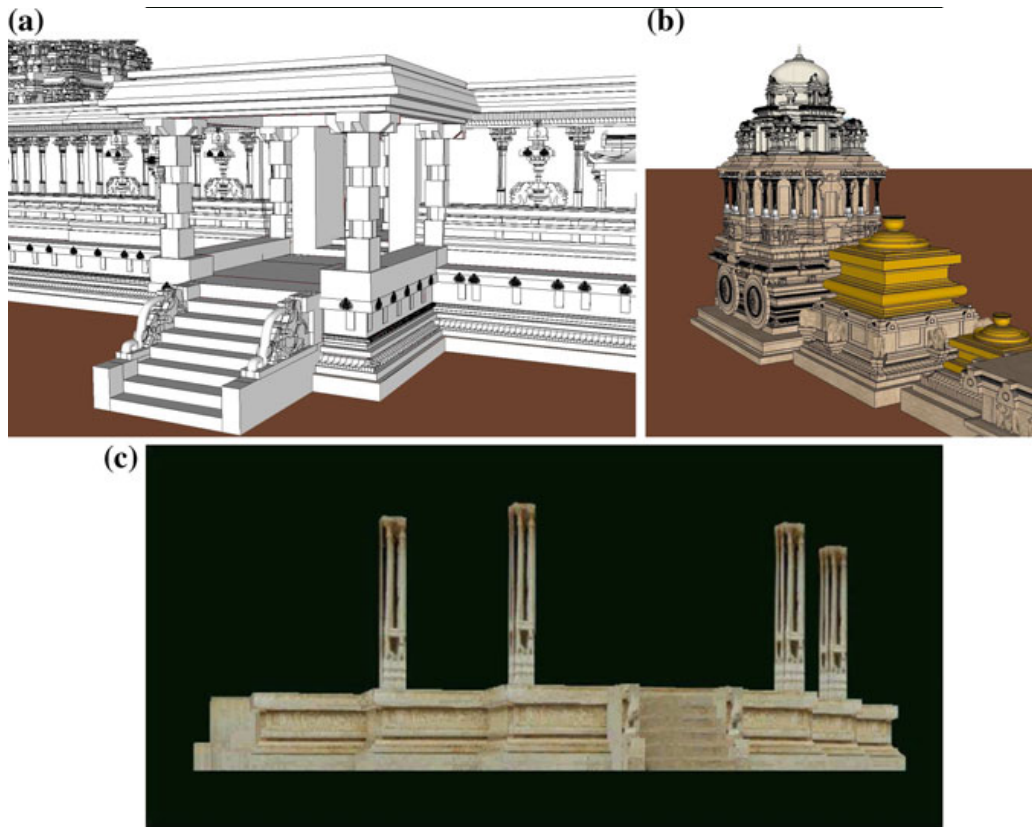
Fig. 8 Coarse level models: **a** CAD model of Main Mantapa, **b** CAD model of Stone Chariot and **c** single-view reconstruction model of Kalyana Mantapa at Vittala Temple, Hampi
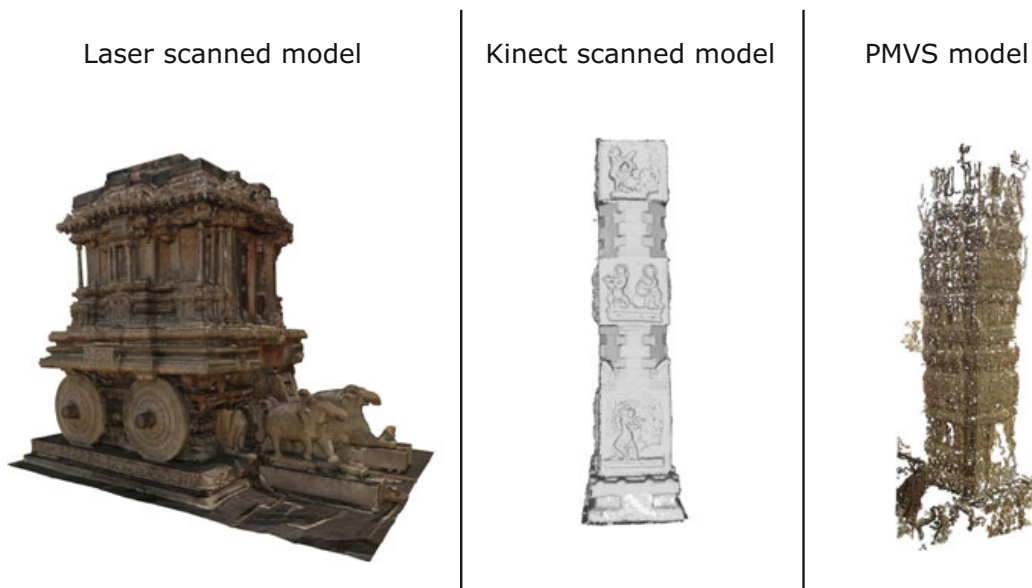

Laser scanned model   Kinect scanned model   PMVS model

Fig. 9 Detail level 3D reconstructed models from laser scanner for Stone Chariot, Kinect model for a pillar at Main Mantapa and PMVS model for a pillar at Kalyana Mantapa, Hampi
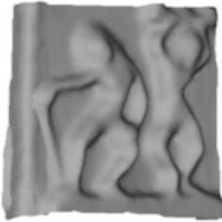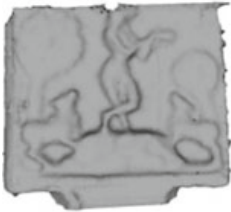
Fig. 10 Results for the proposed super resolution algorithm. First row shows the 3D models of low-resolution (LR) point cloud data. Second row shows the 3D models of super-resolved (SR) point cloud data
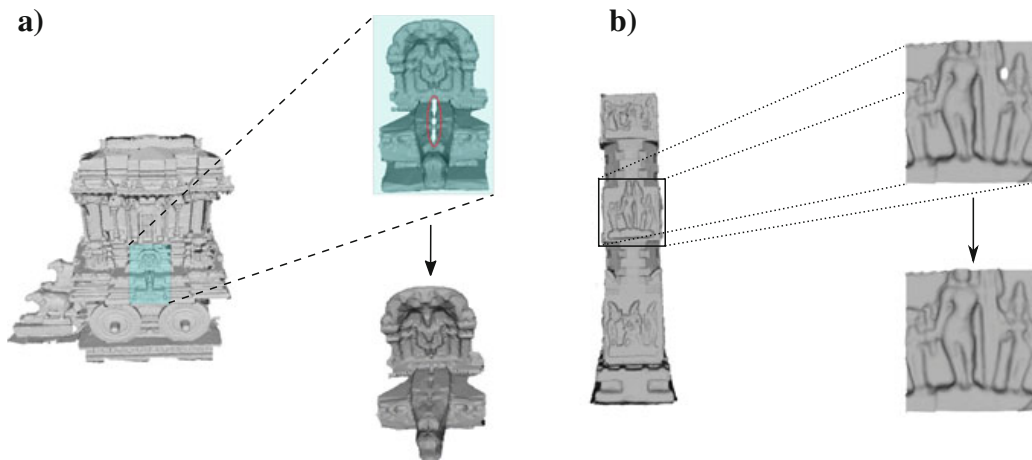

Fig. 11 Results for the proposed hole filling algorithm. **a** hole filling for a part of the Stone Chariot at Vittala Temple, Hampi and **b** hole filling for a part of the artifact at one of the pillars of Main Mantapa, Hampi
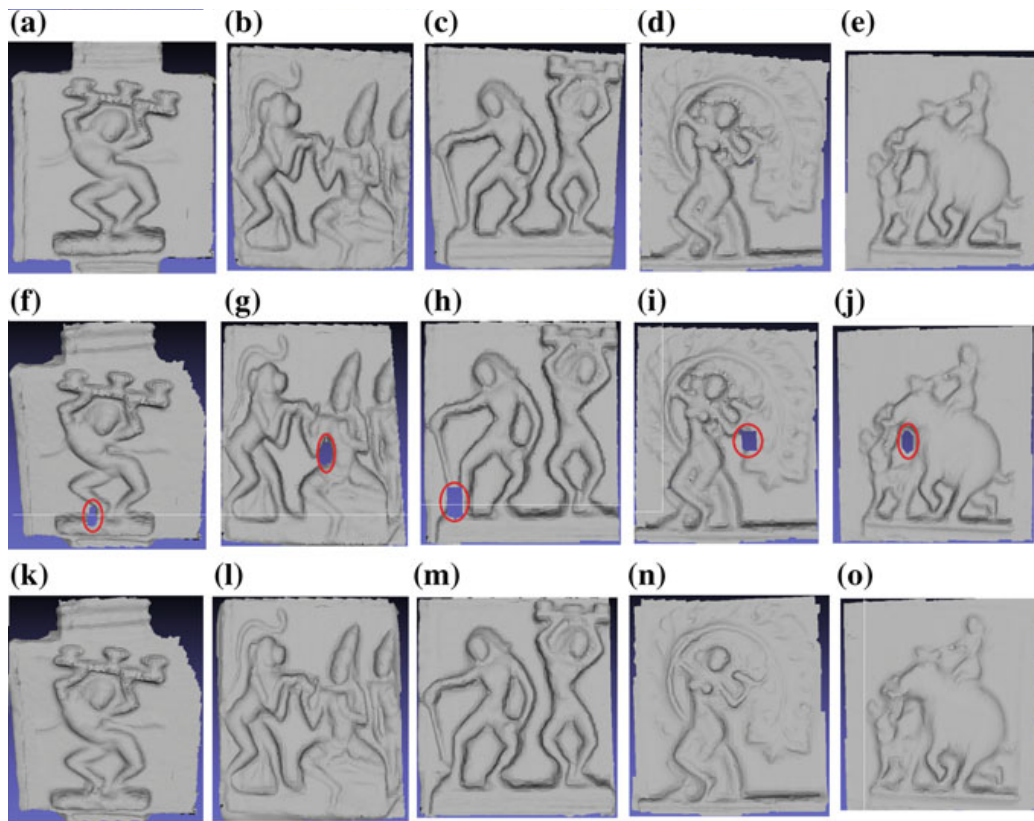
Fig. 12 Hole filling of synthetically generated missing regions on 3D models from pillars at Kalyana Mantapa, Hampi
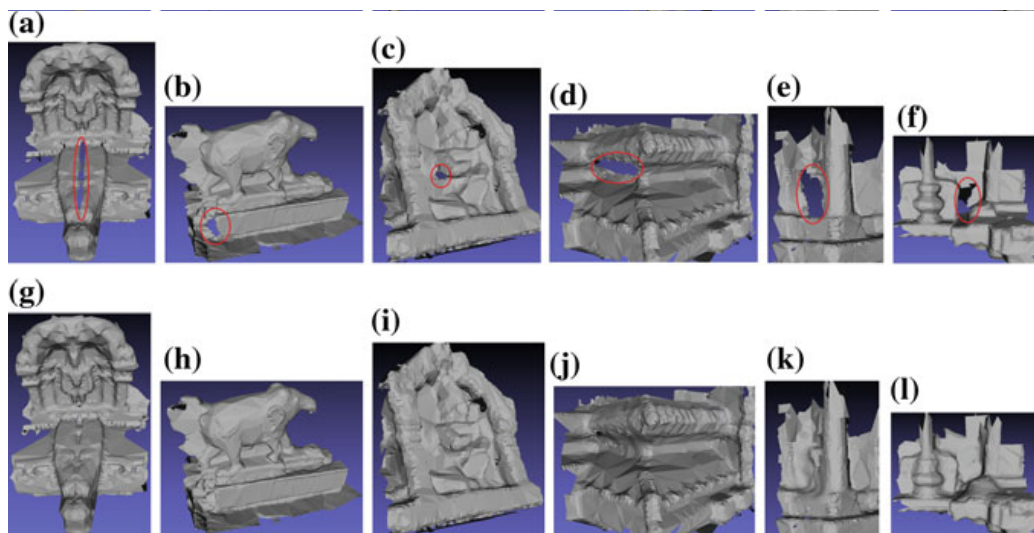

Fig. 13 Hole filling of real missing regions on 3D models from Stone Chariot at Vittala Temple, Hampi

Fig. 14 Categorization results after training four variants of the pillars and testing with similar pillars at Kalyana Mantapa, Hampi
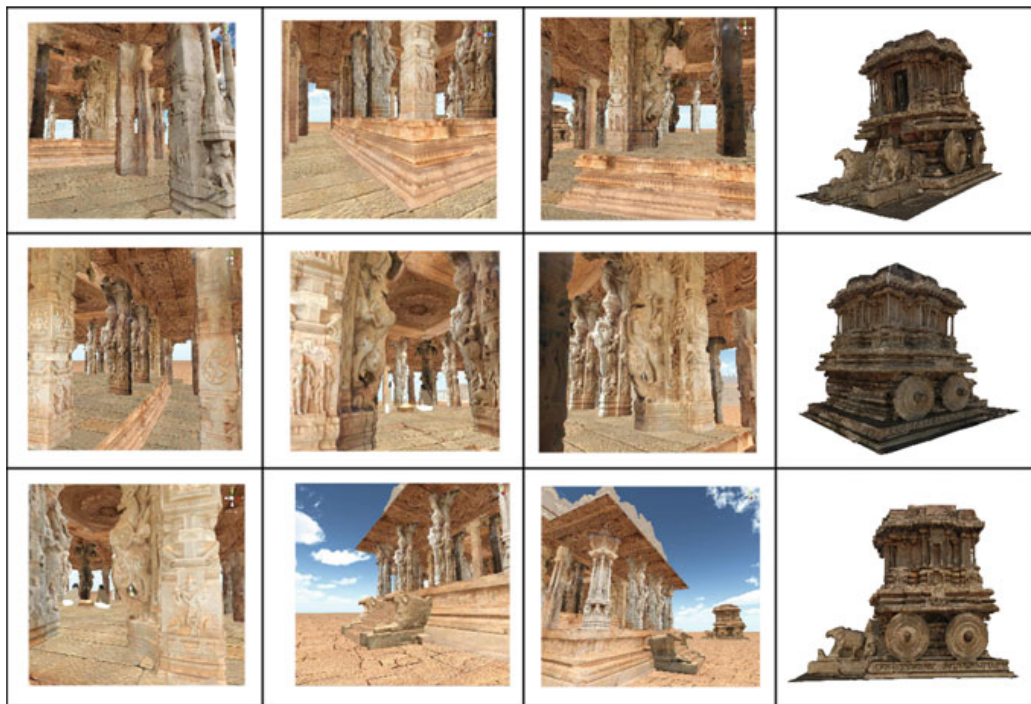

Fig. 15 Rendered views of Kalyana Mantapa (first three columns) and Stone Chariot (last column) using Unity 3D gaming engine and OGRE 3D rendering engine

Fig. 16 Comparison of rendered views and the original images at Vittala Temple, Hampi. Left side of each image is rendered scene and right side of each image is original image

3 3 The Realistic Walkthrough

In the rendering phase, the coarse level models and detail level models are registered using ICP algorithm [23]. The coarse to detail level reconstruction of Kalyana Mantapa is as shown in Fig. 7b. The coarse to detail 3D reconstructed models are interactively rendered for the generation of the *realistic walkthrough*. The coarse to detail level reconstructed models are rendered using OGRE 3D rendering engine and Unity 3D gaming engine and the rendered views are as shown in Fig. 15. The closeups of the rendered scene and the original images at the Vittala Temple are as shown in Fig. 16.

Finally, a gesture-based natural interface is provided for an interactive walkthrough in the virtual arena using Kinect camera. A gesture recognizing application with Kinect is used to determine when a particular gesture has been made by the user. Camera calibration is adjusted to user's body proportions, tracking the user is done in virtual walkthrough, and collision detection is carried out to prevent walkthrough in restricted areas. We demonstrate the realistic walkthrough of Kalyana Mantapa and Stone Chariot at Vittala Temple, Hampi, India.

4   Conclusions

In this chapter, we have proposed a framework for the realistic walkthrough of cultural heritage sites. Digital restoration and preservation of cultural heritage sites is an important area of research due to the availability of techniques in data acquisition, data processing, and rendering. The main goal of our work is to create a coarse to detail 3D reconstruction framework for the generation of the *realistic walkthrough* of cultural heritage sites. To accomplish this, we acquire data of different modalities by combining different scanning and modeling tools to achieve a better completion of 3D reconstructed models especially in large and complex artifacts. Since there is loss of data in 3D models obtained from these devices, we have developed data processing algorithms. These include super resolution, hole filling, and object categorization using metric tensor and Chirstofel symbols as novel set of geometric features. For reconstruction, we proposed a coarse to detail level 3D reconstruction framework by registering the coarse level 3D reconstruction of the cultural heritage sites and detail level 3D reconstruction of the artworks at the cultural heritage sites. The rendering engine is used for digital restoration of original appearance of cultural heritage artifacts. We have demonstrated the proposed framework on Vittala Temple at Hampi, India.

References

1. Allen PK, Troccoli A, Smith B, Stamos I, Murray S (2003) The beauvais cathedral project. In: Conference on computer vision and pattern recognition workshop. CVPRW '03, vol 1, pp 10–10
2. Bernardini F, Mittleman J, Rushmeier H, Silva C, Taubin G (1999) The ball-pivoting algorithm for surface reconstruction. IEEE Trans Vis Comput Graph 5(4):349–359
3. Bernardini F, Rushmeier HE (2002) The 3d model acquisition pipeline. Comput Graph Forum 21(2):149–172
4. Chang CC, Lin CJ (2011) Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol 2(3):27:1–27:27
5. Corsini M, Dellepiane M, Ponchio F, Scopigno R (2009) Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties. Comput Graph Forum 28(7):1755–1764
6. Fontana R, Greco M, Materazzi M, Pampaloni E, Pezzati L, Rocchini C, Scopigno R (2002) Three-dimensional modelling of statues: the minerva of arezzo. J Cult Heritage 3(4):325–331
7. Furukawa Y, Ponce J (2010) Accurate, dense, and robust multi-view stereopsis. IEEE Trans Pattern Anal Mach Intell 32(8):1362–1376

8. Ganihar SA, Joshi S, Setty S, Mudenagudi U (2014) 3d object decomposition and super resolution. In: SIGGRAPH Asia posters. ACM, pp 5:1–5:1

9. Ganihar SA, Joshi S, Setty S, Mudenagudi U (2014) 3d object super resolution using metric tensor and christo el symbols. In: Proceedings of the 2014 Indian conference on computer vision graphics and image processing, ICVGIP '14. ACM, pp 87:1–87:8

10. Ganihar SA, Joshi S, Setty S, Mudenagudi U (2015) Computer vision—ACCV 2014 workshops, chap. Metric tensor and Christo el symbols based 3D object categorization. Springer, pp 138–151

11. Ganihar SA, Joshi S, Shetty S, Mudenagudi U (2014) Metric tensor and christo el symbols based 3d object categorization. In: ACM SIGGRAPH posters, pp 38:1–38:1

12. Gomes L, Bellon ORP, Silva L (2014) 3d reconstruction methods for digital preservation of cultural heritage: a survey. Pattern Recogn Lett 50:3–14

13. Grun A, Remondino F, Zhang L (2004) Photogrammetric reconstruction of the great buddha of bamiyan, afghanistan. Photogram Rec 19(107):177–199

14. Ikeuchi K, Oishi T, Takamatsu J, Sagawa R, Nakazawa A, Kurazume R, Nishino K, Kamakura M, Okamoto Y (2007) The great buddha project: digitally archiving, restoring, and analyzing cultural heritage objects. Int J Comput Vis 75(1):189–208

15. Izadi S, Kim D, Hilliges O, Molyneaux D, Newcombe R, Kohli P, Shotton J, Hodges S, Freeman D, Davison A, Fitzgibbon A (2011) Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In: Proceedings of the 24th annual ACM symposium on user interface software and technology, UIST '11. ACM, pp 559–568

16. Jost J (2011) Riemannian geometry and geometric analysis. Springer Universitat texts. Springer, Berlin

17. Kazhdan M, Bolitho M, Hoppe H (2006) Poisson surface reconstruction. In: Proceedings of the fourth Eurographics symposium on geometry processing, SGP '06. Eurographics Association, pp 61–70

18. Koutsourakis P, Simon L, Teboul O, Tziritas G, Paragios N (2009) Single view reconstruction using shape grammars for urban environments. In: 2009 IEEE 12th international conference on computer vision, pp 1795–1802

19. Levoy M, Pulli K, Curless B, Rusinkiewicz S, Koller D, Pereira L, Ginzton M, Anderson S, Davis J, Ginsberg J, Shade J, Fulk D (2000) The digital michelangelo project: 3d scanning of large statues. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques, SIGGRAPH '00. ACM Press/Addison-Wesley Publishing Co, pp 131–144

20. Li R, Luo T, Zha H (2010) 3d digitization and its applications in cultural heritage. In: Proceedings of the third international conference on digital heritage, EuroMed'10. Springer, pp 381–388

21. Liepa P (2003) Filling holes in meshes. In: Proceedings of the 2003 eurographics/ACM SIGGRAPH symposium on geometry processing, SGP '03. Eurographics Association, pp 200–205

22. Mudenagudi U, Ganihar SA, Joshi S, Setty S, Rahul G, Dhotrad S, Natampally M, Kalra P (2015) Computer vision—ACCV 2014 workshops, chap. Realistic walkthrough of cultural heritage sites-Hampi. Springer, pp 554–566

23. Rusinkiewicz S, Levoy M (2001) Efficient variants of the ICP algorithm. In: Third international conference on 3D digital imaging and modeling (3DIM)

24. Setty S, Ganihar SA, Mudenagudi U (2015) Framework for 3d object hole filling. In: IEEE NCVPRIPG, pp 1–4 (2015)

25. Snavely N, Seitz SM, Szeliski R (2006) Photo tourism: exploring photo collections in 3d. ACM Trans Graph 25(3):835–846

26. Stamos I, Allen PK (2001) Automatic registration of 2-d with 3-d imagery in urban environments. In: ICCV, pp 731–737

27. Sonnemann T, Sauerbier M, Remondino F, Schrotter G (2006) Reality-based 3d modeling of the angkorian temples using aerial images. Brit Archaeol Rep Int Ser 1568:573–579

28. Vrubel A, Bellon ORP, Silva L (2009) A 3d reconstruction pipeline for digital preservation. In: IEEE conference on computer vision and pattern recognition. CVPR 2009, pp 2687–2694

29. Wasserman J (2003) Michelangelo's florence peita. Princeton University Press

A Methodology to Reconstruct Large Damaged Regions in Heritage Structures

A. N. Rajagopalan Pratyush Sahay and Subeesh Vasu

1 Introduction

The importance of digitization of heritage data cannot be underestimated. In addition to facing degradation over centuries due to both man-made and natural causes, archaeological and heritage sites are being subject to immense pressure due to rapid urbanization and development activities of the modern-day world which threatens to erode, or in certain cases, even eliminate the rich legacy inherited from past generations [27]. Since "heritage once lost is lost forever", it has become critical to devise steps to aid heritage conservation, which is well highlighted in UNESCO's draft charter on the preservation of digital heritage [32]. Given this vulnerability of heritage sites, it has become imperative to develop methods that shall aid in the process of preservation for posterity. While credible initiatives have been launched in several countries to restore damaged heritage ([10] in India), the conservation architects at such organizations would need a reference to perform the restoration task.

Large-scale interest in tangible heritage digitization arose in the vision, graphics, virtual reality, and related research communities with the advent of e orts such as the Digital Michelangelo project [19] and the David restoration project [4], which successfully demonstrated the use of 3D models in the framework of cultural restoration. In addition to preservation, the above e orts, and also the Google Art Project [9], aim to provide a capability to perform a virtual walk-through of the tangible cultural heritage, enabling internet-based access of rich "common" heritage from across the world. However, it is to be noted that these works, in the current state,

A. N. Rajagopalan ( ) P. Sahay S. Vasu
Indian Institute of Technology Madras, Chennai, India
e-mail: raju@ee.iitm.ac.in

P. Sahay
e-mail: pratyush.a.sahay@gmail.com

S. Vasu
e-mail: ee13d050@ee.iitm.ac.in

would show the naturally existing large damaged regions (if present) "as is". The work in [22] performs detection and inpainting of cracks in images and videos but the scope is limited to 2D while assuming the missing regions to be small.

With the introduction of real-time techniques for 3D model generation using low-cost motion sensor cameras such as the Microsoft Kinect [11], creation of 3D world models has now been brought into the realm of public domain. A simultaneous growth has also been witnessed in online services that allow for web-based 3D model generation from user-provided images [12] or 3D point clouds [13]. The need of the hour is to utilize these capabilities for the preservation of digital heritage.

This chapter deals exclusively with the scenario involving naturally existing large damaged regions in tangible cultural artifacts. Missing information due to inaccurate scanning has been deliberately left out from the discussions as it is a different research problem in its own right. Furthermore, several works already exist in the literature that deals exclusively with this scenario. The input data to our algorithm is a 3D mesh, generation of which constitutes a preprocessing step for our algorithm [7]. Our work provides a framework that can be a valuable aid to heritage restoration and visualization applications by providing the ability to perform geometric reconstruction of such large missing regions in the rendered 3D mesh models.

For geometric reconstruction, we utilize one or more self-similar examples and efficiently blend the undamaged regions from them onto the damaged 3D model at hand. Towards this end, we have used two different methodologies: (i) A tensor voting based method when multiple self-similar examples are available, and (ii) a gradient map and dictionary learning (DL) based method for the scenario when only a single self-similar example is available to harness the geometric prior from. In both the cases, the missing geometry is inferred by exploiting the constraint that the underlying "missing" geometry in the hole shows locally smooth variations. For all the examples presented in this chapter, the self-similar structures were picked by visual inspection and suggestions from cultural experts. In these examples, the damaged regions were easy to visually identify and were given manually as input to our algorithms.

Both the geometric inpainting methods (TV-based and DL-based) work on point clouds irrespective of their regularity. The workflow involves obtaining the corresponding depth map of a boundary patch on the surface of the 3D model, performing inpainting in the depth domain, and reprojecting the inpainted depth patch onto the 3D surface. This resampling ensures that the algorithm can handle both regular and irregular point clouds, and at the same time helps to obtain a smoothly varying gradient map. Since the camera poses are known (as we use the structure-from-motion (SFM)-based 3D model generation pipeline), it is straightforward to obtain the corresponding projection of the 3D model, and hence its depth map in each camera. The point cloud is projected in each of the cameras and only the region of the hole that best projects into a camera is filled up in that view. The best visibility of a 3D point in a camera is as defined in [30]. A flowchart of our overall approach to 3D inpainting is given in Fig. 1.
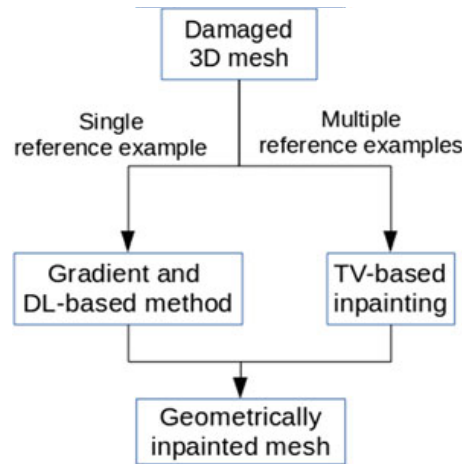
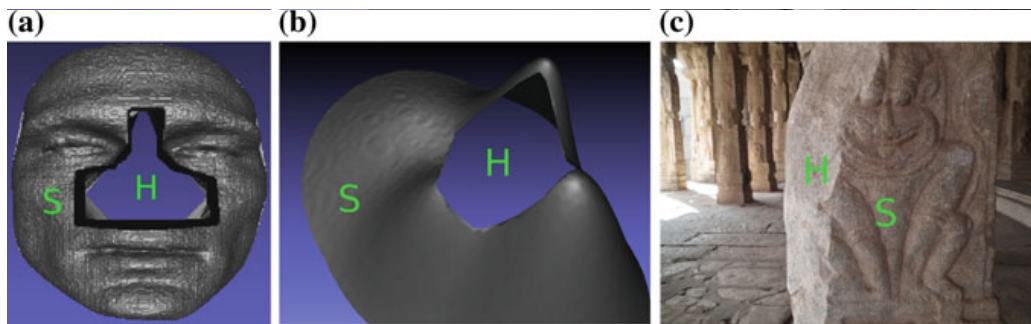**Fig. 1** Proposed workflow for geometric reconstruction

Fig. 2 **a**, **b**, **c** Examples of damaged structures. $\mathscr{H}$ represents the damaged region, while $\mathscr{S}$ represents the known geometry in the 3D model

2 Multiple Self-similar Example-Based 3D Geometry Inpainting

It is not unusual to find similar looking archaeological structures in heritage sites. This is all the more true of structures that have historical significance (for example, the famous Narasimha statue which can be found across different temples in Hampi (see Fig. 2c)). In this section, we discuss the scenario of geometric inpainting of a large hole region in the presence of such multiple self-similar examples. We achieve this by suitably harnessing the object class-specific geometric prior made available by the set of self-similar examples. The term "hole" implies a naturally occurring large damaged region. During the inpainting process, the hole region is manually marked and is labeled as $\mathscr{H}$.

Given a 3D model $\mathscr{S} \in \mathbb{R}^3$ of a real-world object with large missing or damaged regions $\mathscr{H}$ (Fig. 2), it is desirable to estimate the underlying surface geometry in $\mathscr{H}$ for a faithful digital 3D model generation. A perceptually intuitive way is to provide for a natural progression of the surface topology existing in the neighborhood of the

boundary of $\mathcal{H}$, maintaining the local surface curvature and smoothness in different regions in the process [17]. Works such as [6] have attempted mesh completion for the case of small missing regions by using smoothness priors with respect to the local neighborhood. They considered iterative extensions of the neighborhood geometry into the hole using volumetric diffusion. However, such an approach fails to correctly inpaint large holes which tend to have a surface complexity unique to an object class. A related scenario of filling holes in 3D surfaces created due to sensor imperfections, low surface reflectivity or self-occlusion on the object being mapped, etc. has been addressed in [20, 33] which considered filling-up of small- or medium-sized surface deficiencies using local surface topology. Hole filling of smooth surfaces using tensor voting (TV) was carried out in [14], while [16] considered a TV-based inpainting of holes in depth maps using local geometry alone. [3] addressed surface completion for 3D models with repeating "relief structures" (which is a restriction). 3D scan completion from examples is addressed in [24]. However, the requirement of a well-annotated and pre-segmented database and manual marking of landmarks involves considerable manual effort. This calls for the development of more sophisticated hole-filling algorithms that can make use of the available context in order to correctly reconstruct the hole region.

We address this difficult problem by making use of geometric prior harnessed from a few self-similar examples, i.e., we provide class-specific prior information about complex surface variations in the hole region by means of these examples. Such structures can provide an effective prior about the surface complexity that may be unique to the object class in the hole region. In this section, we explain our non-iterative hole-filling algorithm based on an unsupervised computational framework called tensor voting (TV) that judiciously utilizes this geometric prior. We give details on the use of our inpainting methodology for challenging scenarios such as damaged real-world structures which generally have very few self-similar examples. It is to be noted that the non-availability of a large number of examples for such scenarios precludes usage of big database methods such as PCA.

We begin with a 3D model (point cloud) of the target structure to be inpainted. For real-world structures "in the wild", the availability of examples similar to the target structure will typically be limited to a few (about 5), although some heritage sites may have more. These structures, wherever available, too need to be converted to their respective 3D models. Without the imposition of any constraints on the orientation of the acquisition device (optical or laser), large pose variations can exist between 3D point clouds of the damaged structure $\mathcal{S}$ and its self-similar examples $\{\mathcal{M}_i\}$. In order to exploit geometric prior for the purpose of missing surface inference, registration of each example in $\{\mathcal{M}_i\}$ by means of a transform T_i, $T_i : \mathcal{M}_i \quad ^3$, is needed, wherein

$$T_i^* = \arg\min_{T_i} \quad T_i(\mathcal{M}_i) - \mathcal{S} \quad _2 \tag{1}$$

Since the acquisition environment is not a controlled setup, the 3D models may exhibit a large number of outliers too. However, the fact that $\{\mathcal{M}_i\}$ are self-similar
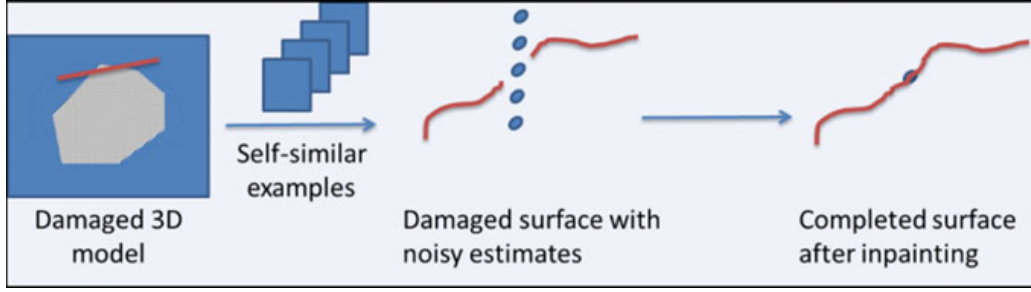
Fig. 3 Summary of TV-based inpainting which is performed on the "height" scanline represented by the red line

to $\mathcal{S}$ works to the advantage of this registration step by providing a large ratio of similar undamaged regions compared to the damaged region and the outliers. Thus, a robust point cloud registration technique that efficiently handles outliers such as CPD (Coherent Point Drift) [2] is preferred to obtain the best possible transform T_i for each self-similar example. Further, we ensure that only the undamaged portions of the 3D model participate in the registration process. Since the ratio of the damaged to the undamaged region is usually small, CPD results in the best possible transform needed to register the point clouds. In this work, we assume that the images are taken from the same camera. If the images are taken from different cameras, there can be resolution differences in the generated point clouds. However, registration using [2] can address these differences to some extent.

Post-registration, the examples in the aligned database $\{\mathcal{M}'_i\}$ provide estimates for the missing region of $\mathcal{S}$ from their corresponding regions. This enables usage of surface point contributions from $\{\mathcal{M}'_i\}$ as noisy estimates within a tensor voting (TV) framework (Fig. 3). To employ tensor voting, the problem of surface completion needs to be reinterpreted. Assume the availability of a set of vertically stacked noisy estimates at a spatial location $\mathbf{p}_i \in \mathbb{R}^3$ in the hole region. In our scenario, these noisy estimates would typically belong to actual undamaged regions on $\{\mathcal{M}'_i\}$. Estimation of the underlying geometric structure requires knowledge of the vectors spanning the normal and tangent space at each point $\mathbf{p}_i = [x_i \, y_i \, z_i]^T$. This is exactly what TV does. By encoding all the 3D surface points in $\mathcal{S}$ and $\mathcal{H}$ (Fig. 2) as second-order tensors, each point $\mathbf{p}_i$ receives a second-order tensor vote $[v_x \, v_y \, v_z]^T$ from each $\mathbf{p}_j$, where $\mathbf{p}_j \in \{\mathcal{S} \cup \mathcal{H}\}$ and $\mathbf{p}_j \in \mathcal{N}(\mathbf{p}_i)$, the neigborhood of $\mathbf{p}_i$. These votes at $\mathbf{p}_i$ are collected into a 3×3 covariance matrix S_{p_i} [25], which allows for estimating the local geometry at $\mathbf{p}_i$ through its eigensystem, i.e.,

$$S_{p_i} = \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \lambda_2 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \lambda_3 \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T \tag{2}$$

$$= (\lambda_1 - \lambda_2)\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + (\lambda_2 - \lambda_3)(\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T) + \lambda_3(\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T) \tag{3}$$

where λ_1, λ_2 and λ_3 are the eigenvalues in decreasing order of magnitude, and $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, and $\hat{\mathbf{e}}_3$ are the respective eigenvectors. Inference of a geometric surface at $\mathbf{p}_i$ is

indicated by the dominance of the scalar coefficient of the first summation term in Eq. (3), i.e., ($\lambda_1 - \lambda_2$), called surface saliency.

The actual process of inpainting proceeds from *outside in* as explained next. In order to preserve local surface curvature variations, a hole-filling strategy from the current hole boundary $\{h_k\}$ to the center of the hole is followed, wherein the oriented surface points on $\mathscr{S}$ (called tokens) in the neighborhood of the hole boundary infer the missing surface points along the hole boundary $\{h_k\}$ within their voting region using the noisy estimates $\{\mathbf{p}_i\}$ from $\{\mathscr{M}_i'\}$ at every h_k. These newly inferred surface points $\mathscr{S}(h_k)$ along with the original tokens now vote to infer missing surface points along a new hole boundary, and this process is repeated till the hole gets completely filled up. Thus, for a surface point $\mathscr{S}(h_k)$ corresponding to the current hole boundary h_k, we have

$$\mathscr{S}(h_k) = \mathbf{p}_{i*}, \quad \mathbf{p}_{i*} \in \mathscr{M}_{i*}' \tag{4}$$

$$\text{where } i^* = \arg\max_i(\lambda_{1,i} - \lambda_{2,i}), i = 1, 2, \ldots r \tag{5}$$
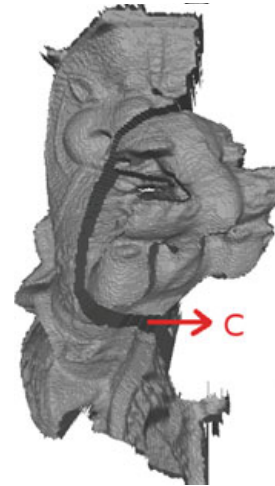
i.e., the best estimate for $\mathscr{S}(h_k)$ is the point $\mathbf{p}_{i*}$ that has received the maximum surface saliency (post-voting from the neighborhood tokens) in comparison to the other points in $\{\mathbf{p}_i\}$ provided by the remaining $(r-1)$ examples in $\{\mathscr{M}_i'\}$. This propagates adaptive surface smoothness in $\mathscr{H}$ from the local geometric information in $\mathscr{S}$ and the noisy estimates $\{\mathbf{p}_i\}$. Once the voting process is over, a smooth structure connecting the subset of $\{\mathbf{p}_i\}$ with high surface saliencies results in completion of the underlying surface geometry in $\mathscr{H}$.

An interesting point to note is that TV provides a highly flexible framework, wherein as many estimates (as are available) can be made use of to infer the missing region, without changing the methodology. Thus, in scenarios where symmetry information is available, even estimates from symmetrical point clouds can be allowed to participate in our 3D inpainting framework.

3 Hole Filling Using a Single Self-similar Example

While we had assumed the availability of at least a few self-similar examples in the previous section, there might be situations wherein this may not be possible or finding several $\{\mathscr{M}_i\}$ may be cumbersome. In this section, we discuss geometric inpainting of large missing regions for situations wherein only a *single* self-similar example ($\mathscr{M}$) is available for use as geometric prior. One could still attempt the TV-based approach of Sect. 2. However, due to the unconstrained nature of the image capturing process, the scale of $\mathscr{S}$ and $\mathscr{M}$ can vary significantly. A global point cloud registration may not resolve local scale changes, which in turn can lead to boundary artifacts in the inpainted result as illustrated in Fig. 4. Our goal is to propagate local region smoothness into the hole, while staying faithful to the geometric prior provided by the single self-similar example.

Fig. 4 The sharp height difference indicated by the boundary marked as shows the effects of local level scale difference between the damaged 3D model and the self-similar example $\mathcal{M}$ on the result of TV-based method

It is well understood that dictionary learning (DL) provides a robust local representation for a given signal class [1]. Since several depth databases are available online, it stands to reason that one can attempt a DL-based approach for hole filling. Surface gradients are known to be resilient to the effects of relative scale differences while at the same time they can exhaustively capture higher order curvatures that may be present on the surface. We solve the problem of performing 3D geometry inpainting (while simultaneously avoiding boundary artifacts) within the framework of DL by harvesting gradients from the self-similar example $\mathcal{M}$ to guide the choice of the sparse representation learnt from online depth databases to infer the missing region that best represents the geometric prior from $\mathcal{M}$.

Incidentally, a related problem in the texture domain is addressed by Poisson image editing [23]. This technique has been shown to successfully blend a given texture patch onto a possibly completely different background images. We also make use of the known gradient field from the registered self-similar example $\mathcal{M}'$, from regions corresponding to the damaged region $\mathcal{H}$ of the broken structure. However, in contrast to their method, we search for a set of sparse representations from an overcomplete dictionary D that best represents the known regions of the current patch at the hole boundary ($\mathcal{P}$). Of these, the sparse representation that has the most similar gradient to $\mathcal{M}'$ at the current boundary point h_k in the unknown region is used to provide an estimate for the missing region.

3.1 Depth Dictionary Generation

With the ready availability of low-cost ranges scanners (such as Microsoft Kinect), several depth databases have been made available online in the last few years [18, 28, 29]. Using the methodology of the k-SVD based DL algorithm [1, 15], overlapping local patches $\{J_i\}$ of size $p \times p$ are extracted from a large number of randomly selected depth maps from online depth databases and arranged into a matrix Y such

that $Y = [\text{vec}(J_1)\ \text{vec}(J_2)\ \cdots\ \text{vec}(J_v)]$, where $\text{vec}(J_i)$ represents an operation that lexico-graphically orders $J_i,\ i = 1, 2, \ldots v$, $Y \quad {}^{m \times v}$, $m << v$ and $m = p^2$. After normal-ization of the columns of Y, a k-atom dictionary $(k > m)$ is learned. This problem is formulated as

$$(\alpha^*, D^*) = \arg\min_{\alpha, D} \parallel Y - D\alpha \parallel_F^2 + \psi \parallel \alpha \parallel_1 \quad \text{s.t.} \parallel \mathbf{d}_i \parallel \quad 1,\ i = 1, 2, \ldots k \quad (6)$$

where $\{\mathbf{d}_i\}$ are the columns of the overcomplete dictionary D, $D \quad {}^{m \times k}$ and α is the matrix encoding the sparsity for the dictionary-based representation, and $k < v$. For our implementation, we set $p = 8, v = 10000$, and $k = 1024$.

For further discussions, we will use D to denote the learned dictionary. In order to establish the representative nature of the learned dictionary, we empirically show the e ectiveness of dictionary-based representation in reconstructing some standard range images. The Middlebury stereo dataset [26] provides several depth maps and stereo pairs (Fig. 5a) which have been widely used in the computer vision commu-nity. The sparse representation of some of these images is first found in the learned dictionary D shown in Fig. 5b, followed by reconstruction of the respective images from this overcomplete dictionary D. This is analogous to KSVD-based DL and image reconstruction using the learnt dictionary [31], but in the depth domain. The original image, the reconstructed image, and the error in reconstruction are shown in Fig. 6 for a few representative examples. The low values of the reconstruction error (as indicated by the legend in Fig. 6c, f) indicate that widely varying depth images admit a sparse yet e ective representation in the overcomplete dictionary D. It is to be noted that none of these images were used for learning the dictionary itself.

3.2 Gradient as a Cue

Returning to the problem on hand, consider the reconstruction of a patch $\mathscr{P}$ (see Fig. 7) such that it contains a large known region and unknown body. Suppose $\alpha_{\mathscr{P}}$ is the column of α corresponding to this patch from the learnt dictionary. Our objec-tive is to find the set of sparse representations $\{\alpha_{\mathscr{P}}(\gamma)\}$ that best explains the known regions in patch $\mathscr{P}$ (γ is the weight used to enforce sparsity of the vector $\alpha_{\mathscr{P}}$, while finding the approximation to $\mathscr{P}$ using learned dictionary D). As discussed earlier, we wish to harness the gradient information coming from the known self-similar exam-ple to guide the sparse representation process in order to achieve better accuracy. Due to the unconstrained image capturing setup, the scale and orientation of the damaged 3D model $\mathscr{S}$ and its self-similar example $\mathscr{M}$ may di er widely. Though there may exist local scale changes in terms of the absolute depth values, a gradient domain rep-resentation of the depths is largely una ected by the relative scale of the 3D models. Also, gradients exhaustively capture local surface variations, thereby serving as a good cue to guide the inpainting process. Generating the gradients in $\mathscr{H}$ by har-vesting the gradients from the corresponding region of the registered self-similar
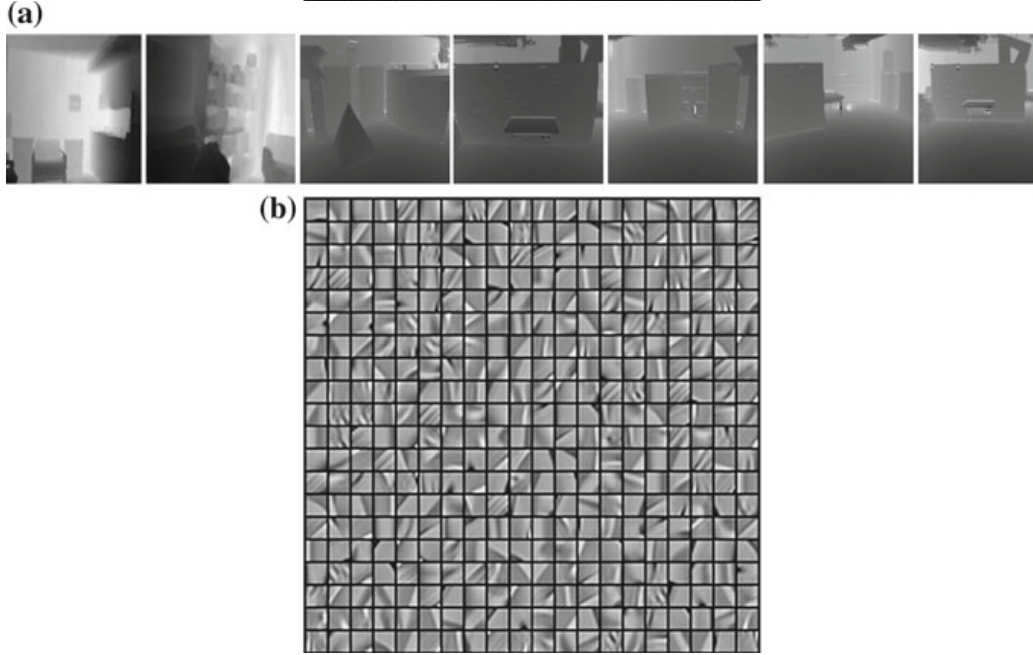
Fig. 5 **a** Some of the depth maps used to learn the dictionary. **b** An example of a $k = 400$ atom dictionary
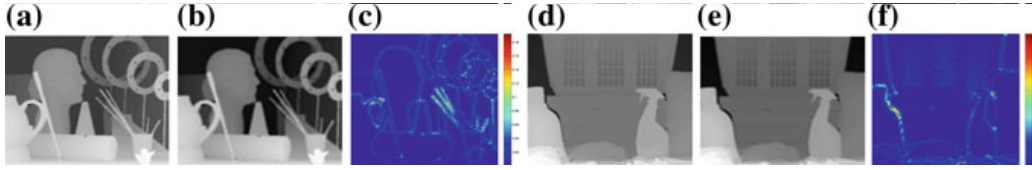

Fig. 6 Error in reconstruction of depth images using the atoms of a depth dictionary. **a**, **d** Input image, **b**, **e** reconstructed image, and **c**, **f** error in reconstruction. The legend values vary between 0 and 0.2, with blue, green, and red roughly corresponding to 0, 0.1, and 0.2, respectively

example $\mathcal{M}'$, we use the sparse representation $\alpha_{\mathcal{P}}(\gamma)$ that has the most similar gradient at the point corresponding to h_k as the best representation for $\mathcal{P}$. This, in turn, provides an estimate for the missing surface h_k. This problem can be formulated in terms of the dot product of the gradients as

$$
\alpha_{\mathcal{P}}^{*}(\gamma) = \arg \min_{\alpha_{\mathcal{P}}(\gamma)} \Bigg(\parallel \left(D\alpha_{\mathcal{P}}(\gamma) - \mathcal{P} \right) \otimes B \parallel_{2}^{2} + \psi \parallel \alpha_{\mathcal{P}(\gamma)} \parallel_{1} +
$$

$$
\gamma \left(1 - \left(\nabla_{n}(D\alpha_{\mathcal{P}}(\gamma)) \mid_{h_k}^{T} \nabla_{n}(\mathcal{M}') \mid_{h_k} \right) \right) \Bigg) \qquad (7)
$$

where B is a mask corresponding to the known region of $\mathcal{P}$ (see Fig. 7), the notation $\otimes$ refers to pixelwise multiplication operation, ∇_{n} calculates the normalized gradient of a function, and $(.) \mid_{h_k}$ indicates the value of a function evaluated at the
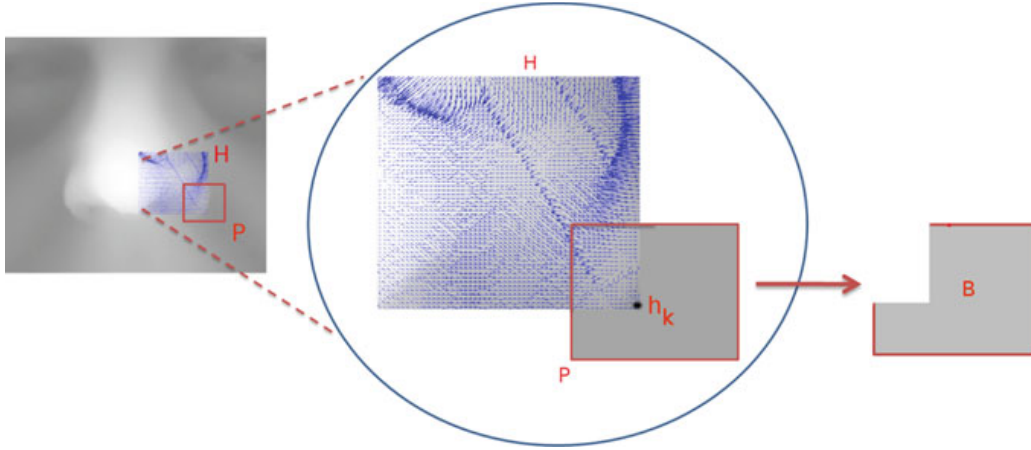
Fig. 7 Using gradients as a cue for inpainting. Consider the region filled with gradients (blue in color) to be a hole. The gradients from the self-similar example are superimposed over the hole. The region bounded in the red box labeled $\mathscr{P}$ is the current patch, while h_k is the current boundary pixel within $\mathscr{P}$. B is the mask corresponding to known regions of $\mathscr{P}$

boundary h_k. The first term in Eq. 7 attempts to reconstruct the known region in the patch $\mathscr{P}$ as close as possible to the original values, while the second term tries to minimize the angle between the normalized gradient vectors. This has the effect of generating a surface that is as similar in orientation as possible to the corresponding surface region in the self-similar example. For solving Eq. 7, we use an iterative greedy algorithm which is as follows: for the current patch under consideration, we obtain a set of sparse representations by varying γ. This is followed by iterating over this set and choosing the atom which minimizes Eq. 7 as the best match for the current patch. Minimizing the error in representing the large known region of $\mathscr{P}$ will ensure local surface smoothness in the region around h_k. We wish to point out that such a sparse representation has been used in image inpainting [21]. While they discuss an optimization technique to obtain the sparse representation "online", our work uses traditional DL and combines the atoms and gradients obtained from self-similar examples into a single formulation to achieve the intended objective.

The example in Fig. 7 shows the depth map of the region around the nose of a human face. A square region ($\mathscr{H}$) is marked as damaged and the blue arrows over the region indicate the harvested gradients from a self-similar example $\mathscr{M}$. The zoomed-in region indicates the positioning of the patch ($\mathscr{P}$) at the current hole boundary h_k. It is to be noted that though gradients encode only the difference of depth values, taking an overlapping patch around the boundary and propagating the information into the hole region ensures a geometrically inpainted result which is both smooth with respect to the boundary of the hole region as well as effectively follows the geometric prior from the self-similar example $\mathscr{M}$. This methodology is succinctly described in Algorithm 1.

Algorithm 1 Fill-in large complex missing region $\mathcal{H}$ using a single self-similar example.

Require: (a) Set of overlapping patches $\{J\ \_i\}$ from several range maps, (b) Self-similar example $\mathcal{M}$

Ensure: $\mathcal{M}$ is registered with $\mathcal{S}$
 1: $\mathcal{H} \leftarrow$ missing regions of $\mathcal{S}$
 2: Obtain the overcomplete dictionary D using Eq. 6
 3: **while** $\mathcal{H} \neq \phi$ **do**
 4:    $\{h_k\} \leftarrow$ boundary($\mathcal{H}$)
 5:    **for** $k = 1$ to No. of elements in $\{h_k\}$ **do**
 6:       Obtain $\mathcal{P}$ containing h_k as shown in Fig. 7
 7:       Get the optimum $\alpha^*_{\mathcal{P}}(\gamma)$ using the cost function in Eq. 7
 8:    **end for**
 9:    Re-estimate the missing region $\mathcal{H}$
10: **end while**

---

4 Experimental Results

We discuss results for geometric completion from multiple self-similar examples followed by the single self-similar example scenario.

4.1 Multiple Self similar Examples

In this section, we demonstrate the effectiveness of our method discussed in Sect. 2 on synthetic as well as real data sets. Complex and large holes that present significantly challenging situations for 3D inpainting are considered. For the real examples, the result of mesh completion using the hole-filling option found in Meshlab [5] is also provided for qualitative comparison with our inpainting methodology. Meshlab is an open-source tool that includes several "state-of-the-art" mesh processing algorithms. The hole-filling algorithm provides a customizable interface using which the best hole filling suitable to a given mesh can be obtained. However, these algorithms try to connect the vertices at the boundary of the hole region using non-self-intersecting patches. As will be evident from the examples, such a flat-fit often vastly deviates from the true underlying surface of the object class. Also, the hole should be bounded by known geometry all around for using this method. Our method, in comparison, can model complex surface curvatures unique to an object class.

4.1.1 Synthetic Data

In the first example (Fig. 8), a 3D model of a human face with significant loss of information is used. For this scenario, Texas 3D Face Recognition Database (Texas 3DFRD) [8] is made use of to provide self-similar examples from which the
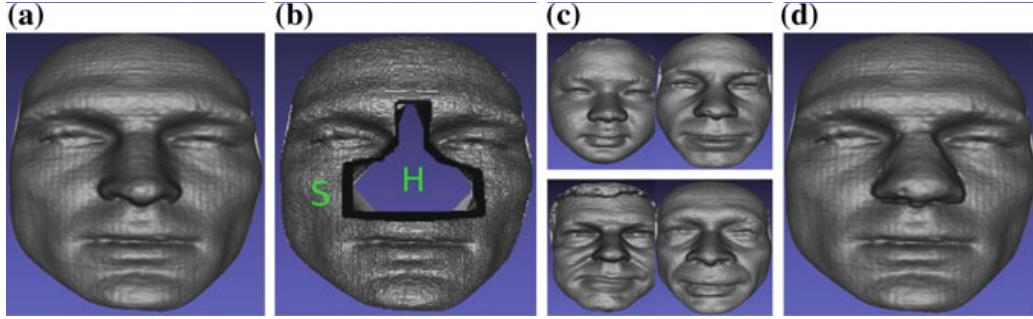
Fig. 8 Complex hole filling: **a** Original face, **b** damaged face model, **c** four of the examples used, and **d** reconstructed face

underlying geometry to be filled-in is inferred using our method. Texas 3DFRD provides 1149 pairs of intensity images and co-registered range images of 118 male and female adult human subjects, captured by 3Q Technologies Ltd.'s MU-2 imaging system, covering five ethnic groups to provide for ample variations in the face space. Face model of subject number 16 is randomly chosen and a fairly complex hole is created corresponding to the region marked $\mathcal{H}$.

The reconstructed face (Fig. 8d) is evidently quite visually satisfying and shows ample local smoothness in the reconstructed regions. Quantitatively evaluating against the ground truth, the reconstruction error is determined using a standard error metric given by

$$ERR = \sqrt{\text{Avg.} \left(\frac{\hat{z}}{z} - 1 \right)^2} \tag{8}$$

where Avg.() refers to the averaging function, $\hat{z}$ is the depth map of the reconstructed face, and z is the depth map of the original face. Although a significant region is marked as missing, a fairly faithful reconstruction is found to be obtained with $ERR = 0.0095$ only.

4.1.2 Real-World Objects

Archaeological sites provide numerous examples of structures showing structural deformities or broken regions due to exposure to several natural and man-made forces of degradation over centuries and thus provide good examples over which the performance of the algorithm can be evaluated. We consider here a few interesting examples of 3D models of real-world structures.

The first real example is that of a stone pillar (Fig. 9). This real example has considerable symmetry information that can be utilized in addition to using the self-similar examples. The broken structure in Fig. 9 is geometrically inpainted using Meshlab as well as our TV-based method of Sect. 2 to yield the results shown in
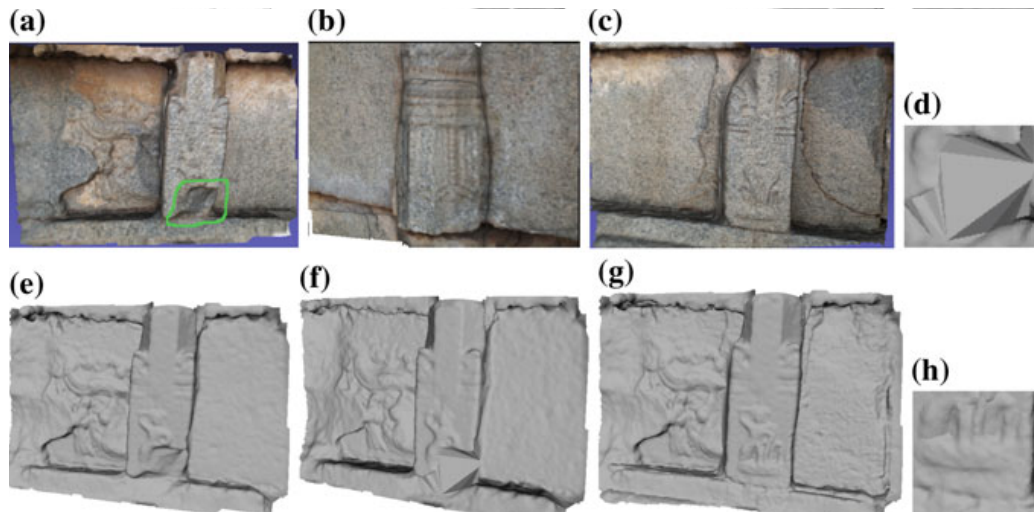
Fig. 9 A real-world structure example. **a**, **e** Image and 3D model of a broken stone pillar, **b**, **c** self-similar pillar structures, **f** reconstructed structure using Meshlab, and **g** reconstructed structure using our method. A zoomed-in inpainted region from **f** and **g** is shown in **d** and **h**, respectively. The damaged region is shown bounded in green

Fig. 9f, g, respectively. The corresponding inpainted zoomed-in regions are shown in Fig. 9d, h. Note that in comparison to the hole-filling algorithm of Meshlab, knowledge of symmetry along with the available information from self-similar examples is better harnessed by our method to infer the most salient surface in the missing region.

The following example is that of a damaged *bas relief* of an elephant where the region comprising one of the front legs, the trunk, and the flat base region is missing (Fig. 10a, d). Here, symmetry information is available for only the flat base region, while no such symmetry exists corresponding to the other missing regions. Though the missing region is quite complex with varying curvatures, our 3D geometry inpainting algorithm provides for salient surface inference in the missing region (Fig. 10f) effectively utilizing the class information propagated by the available self-similar examples (Fig. 10b, c). A local geometry-based hole filling provided by Meshlab fails to correctly recover the true geometry in the underlying hole region (Fig. 10e).

In the next example, a much more complex scenario of a stone carving of a lion found at Mahabalipuram (India) is considered (Fig. 11a, e). Here, the entire face of the lion statue was missing. This is a more difficult scenario compared to the previous real examples as the missing region has no symmetry information to make use of. The underlying missing region is thus inferred based purely on the estimates provided by the self-similar examples (Fig. 11b, c). The geometrically inpainted result shown Fig. 11g indicates that our algorithm is much more effective at handling even such difficult scenarios than the Meshlab-based hole-filling algorithm (Fig. 11f). This is further confirmed by the corresponding inpainted zoomed-in regions (Fig. 11h, d).

Fig. 10 Another real-world structure. **a**, **d** Damaged *bas relief* of an elephant, **b**, **c** self-similar examples, **e** reconstructed elephant structure using Meshlab, and **f** reconstructed elephant structure using our method

Fig. 11 A real-world structure. **a**, **e** Image and 3D model of damaged lion statue, with the face region missing, **b**, **c** self-similar examples, **f** reconstructed statue using Meshlab, and **g** reconstructed statue using our method. A zoomed-in inpainted region from **f** and **g** is shown in **d** and **h**, respectively.

Fig. 12 Real-world object. **a**, **e** Broken horse statue, **b**, **c** self-similar statues, **f** reconstructed 3D model using Meshlab, and **g** reconstructed 3D model using our method. Zoomed-in inpainted region from **f** and **g** is shown in **d** and **h**, respectively
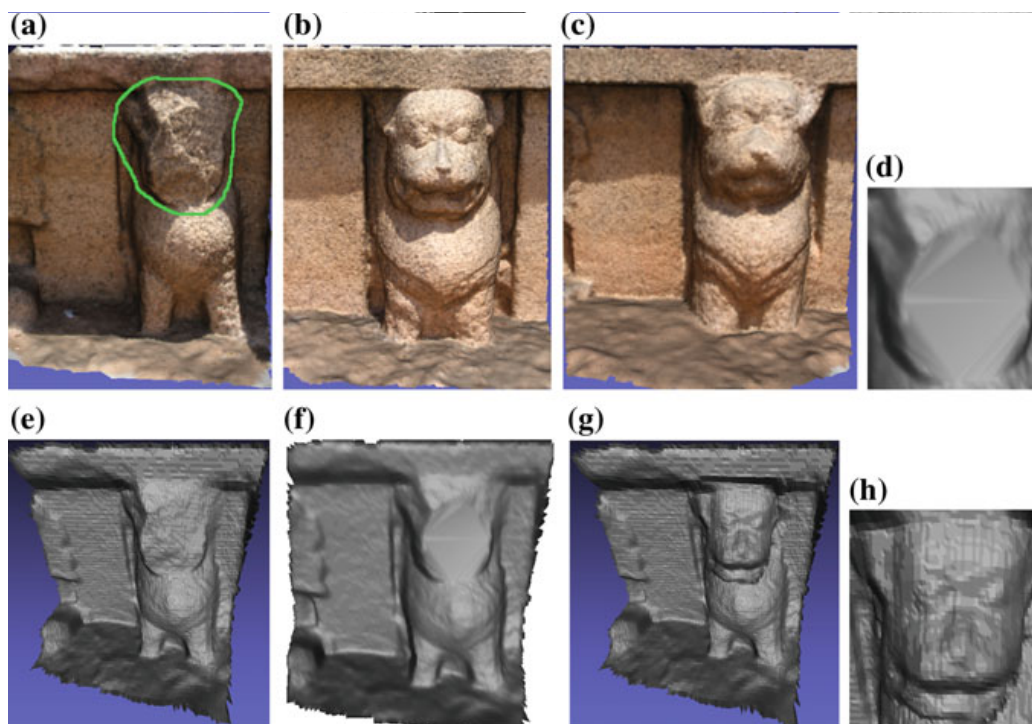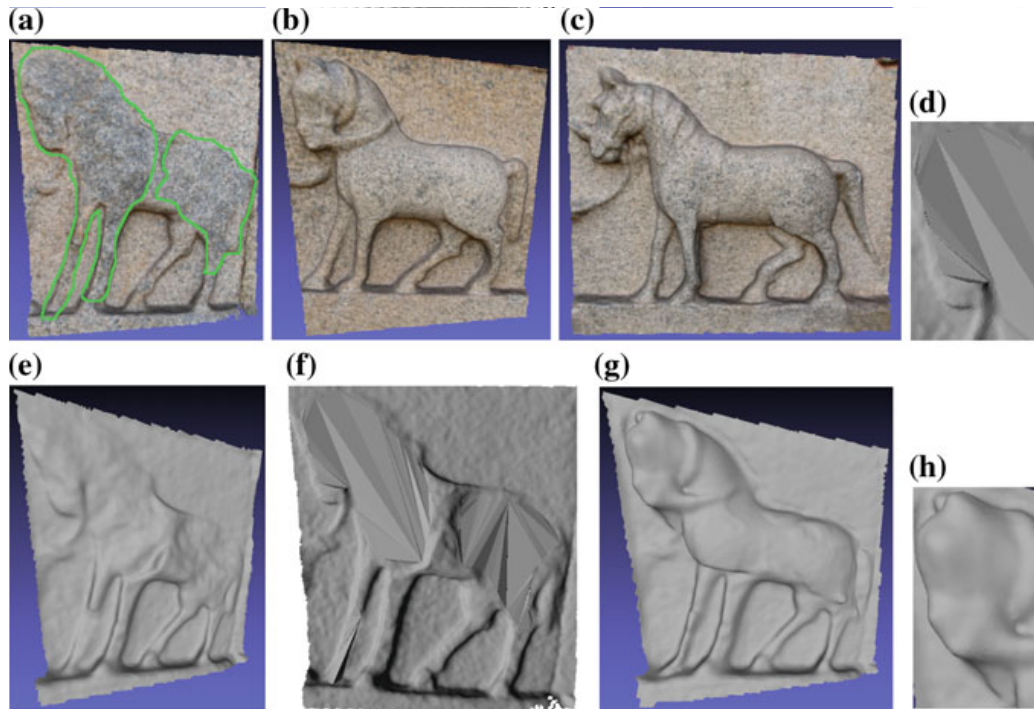
In the final example, the *bas relief* of a damaged horse is considered. The scenario here is quite complex as the whole body area of the horse is damaged. Yet again, our method (Fig. 12g) recovers the underlying geometry while the local geometry-based hole filling found in Meshlab completely fails to do so (Fig. 12f). A part of the inpainted region corresponding to the head of the horse from Fig. 12f, g is shown zoomed in Fig. 12d, h, respectively, to depict the difference in inpainting quality of the two methods.

4.2 Single Self similar Example Scenario

In this section, we discuss the effectiveness of the DL-based method on synthetic as well as real-world objects, followed by a comparison of the results obtained using our DL-based method with those obtained using our TV-based method.

Consider the synthetic example of a statue of a human in a dancing pose in Fig. 13. A large hole region is created as shown in Fig. 13c, and hole filling is performed using a single self-similar example available (Fig. 13b). The reconstructed 3D model (Fig. 13e) is visually quite similar to the original 3D model (Fig. 13d). In addition, the low values of the reconstruction error (as indicated by the mean and legend in
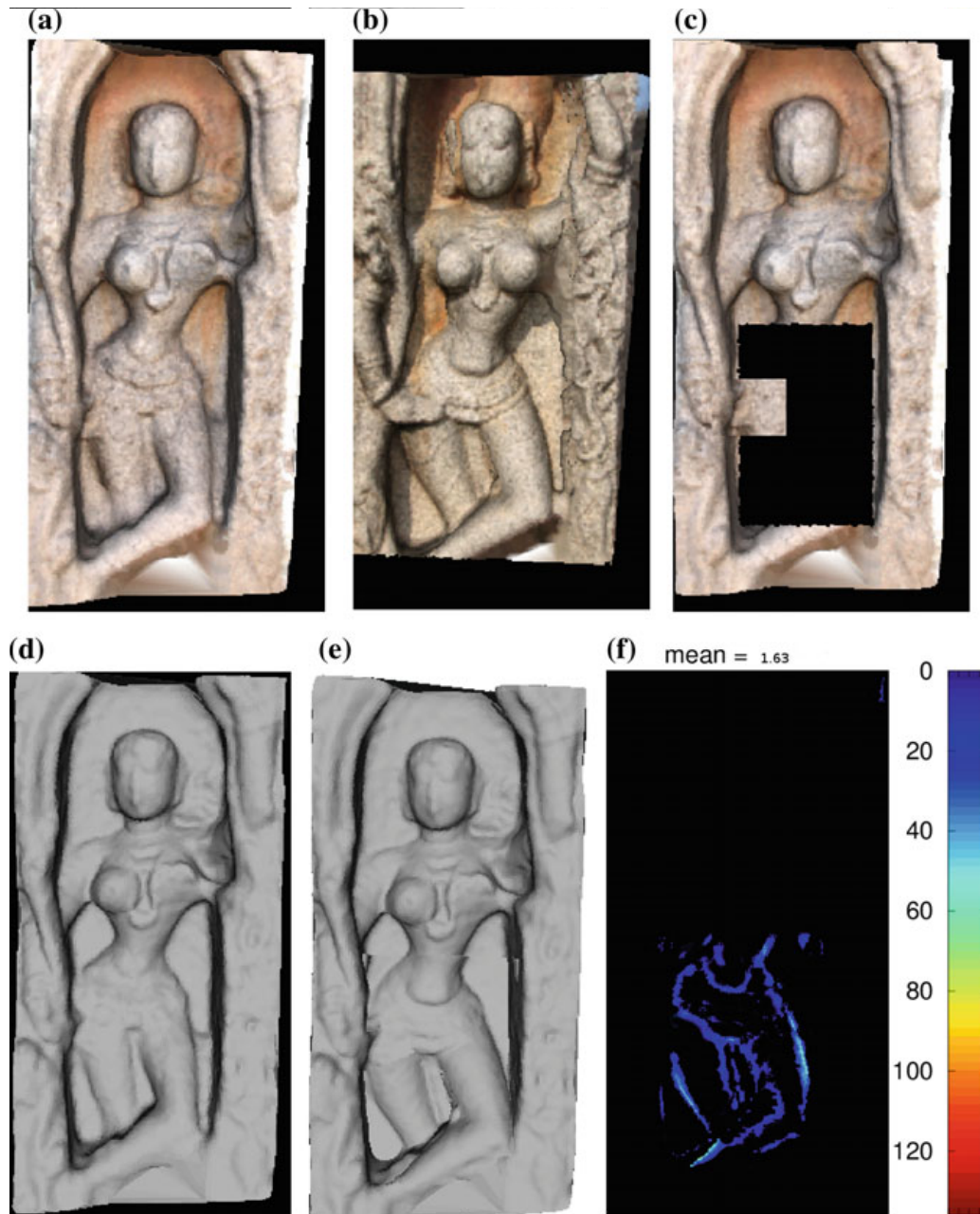
Fig. 13 3D geometry inpainting of a synthetically created hole region using a single self-similar example. **a**, **d** The original structure, **b** self-similar example used, **c** the synthetically created damaged structure, **e** reconstructed 3D structure, and **f** reconstruction error
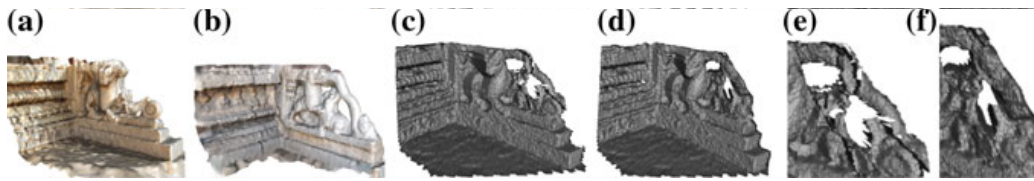
Fig. 14 Another example for 3D geometry inpainting using a single self-similar example. **a** The damaged structure. **b** Self-similar example used. **c** Result using TV-based method. **d** Reconstruction using the DL-based method. Zoomed-in regions corresponding to **c** and **d** are shown in **e** and **f**, respectively
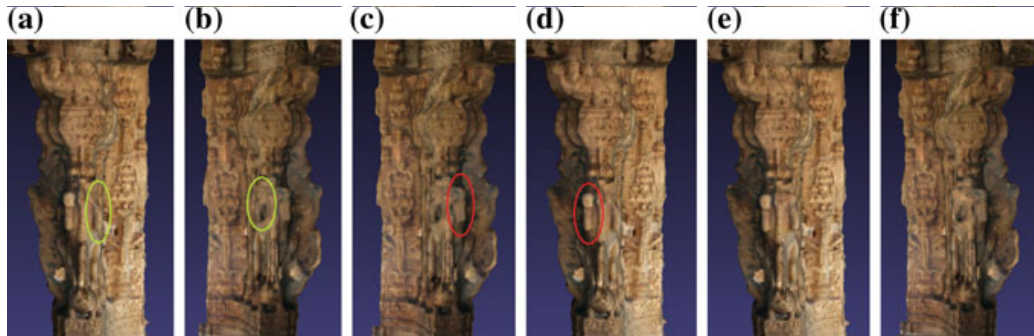

Fig. 15 **a**, **b** Pillar containing damaged parts. **c**, **d** Self-similar examples used for **a**, **b**. **e**, **f** 3D outputs of inpainted structures

Fig. 13f) reflect effective hole filling even when inpainting such large and complex geometric holes using just a single self-similar example within our DL framework.

Among the real examples, we first consider the damaged staircase of a temple structure, where the whole sidearm was missing (Fig. 14a). Using the self-similar example shown in Fig. 14b, the TV-based method results in an inpainting with visible discontinuities at the boundary due to local scale mismatch (Fig. 14c) with the corresponding zoomed-in reconstructed region shown in Fig. 14e. The DL-based method, in comparison, provides a much smoother inpainting (Fig. 14d) with corresponding zoomed-in reconstructed region displayed in Fig. 14f.

In Fig. 15, we showcase our inpainting results on the broken paws of the Yali carved on a pillar inside Vithala temple complex. This pillar had two Yali structures on it with broken left paw on one (Fig. 15a) and a broken right paw on the other (Fig. 15b). We have used one Yali as the self-similar example for another and vice versa (Fig. 15c, d) to perform the DL-based inpainting to get the results shown in Fig. 15e, f. For this particular example, we have done an additional texture mapping to the inpainted region by assigning the texture value of the nearest 3D point to every new point in the inpainted structure. A 3D printout (on a scale of 75 : 1) was also taken for both the broken and inpainted pillars.

5   Further Discussions

The TV-based method discussed in Sect. 2 was shown to infer a smooth underlying 3D geometry upon using a set of self-similar examples, while in Sect. 3, we showed that the DL-based method can provide a smooth, boundary-artifact free inpainting using a dictionary learned from several depth images and a single self-similar example. The result for the TV-based method is obtained by using all the available self-similar examples while the result from the DL-based method is obtained by picking one of the available self-similar examples. Specifically, the self-similar example that provides a lower error upon registration (Eq. 1) is picked as the example for harnessing the gradient information for the broken region $\mathcal{H}$. Though the application scenarios of the two methods may seem different, a qualitative comparison of the performance of the two methods can reveal the extent of visual appeal of the results that can be obtained using just a single self-similar example by means of the DL-based method when compared to the result provided by the TV-based method. A comparison of the results in Fig. 14 validates the effectiveness of our DL-based method to saliently reconstruct large missing regions even when using a single self-similar example. The method generally performs quite well although there are instances (such as Fig. 13) wherein some discontinuities continue to exist if the scale of the examples varies by large amounts, since our method can only infer the best available missing surface from the examples. This could possibly be resolved by a post-processing operation involving surface smoothing.

For the example shown in the 1st row of Fig. 16, the result due to the DL-based method (Fig. 16c) follows the geometric prior well and yields smooth inpainting in the hole region. However, since there is no inference of the best surface being performed in this method, small local level artifacts may exist compared to the result due to the TV-based method (Fig. 16b). This can be seen on the inpainted left leg, which looks asymmetric with respect to the right leg. Similarly, for the example considered in the 2nd row of Fig. 16, the globally smooth inpainting provided by the DL-based method (Fig. 16f) does contain some local artifacts around the head region when compared to the TV-based best-inferred surface shown in Fig. 16e. Thus, except for some minor local level artifacts, the DL-based method performs an appreciable inpainting of the hole region even in the difficult scenario of using the geometric prior from a single self-similar example.

For the synthetic example considered in Fig. 8, we show in Fig. 17a the average accumulation error for every iteration of TV-based hole-filling method, while Fig. 17b shows the average accumulation error for every iteration of the DL-based hole-filling method. With every iteration, the hole filling proceeds from the hole boundary to the center of the hole region. For both the plots, the line plot in blue is for the hole region marked $\mathcal{H}$ in Fig. 8b for which the total number of iterations was 57, while that in red is for hole filling performed over a synthetically created hole over the left eye region of the same face model for which the total number of iterations was 43. For the TV-based method, the accumulation error is of the order of 10^{-3}, while for the DL-based method it is of the order of 10^{-1}. This clearly indicates

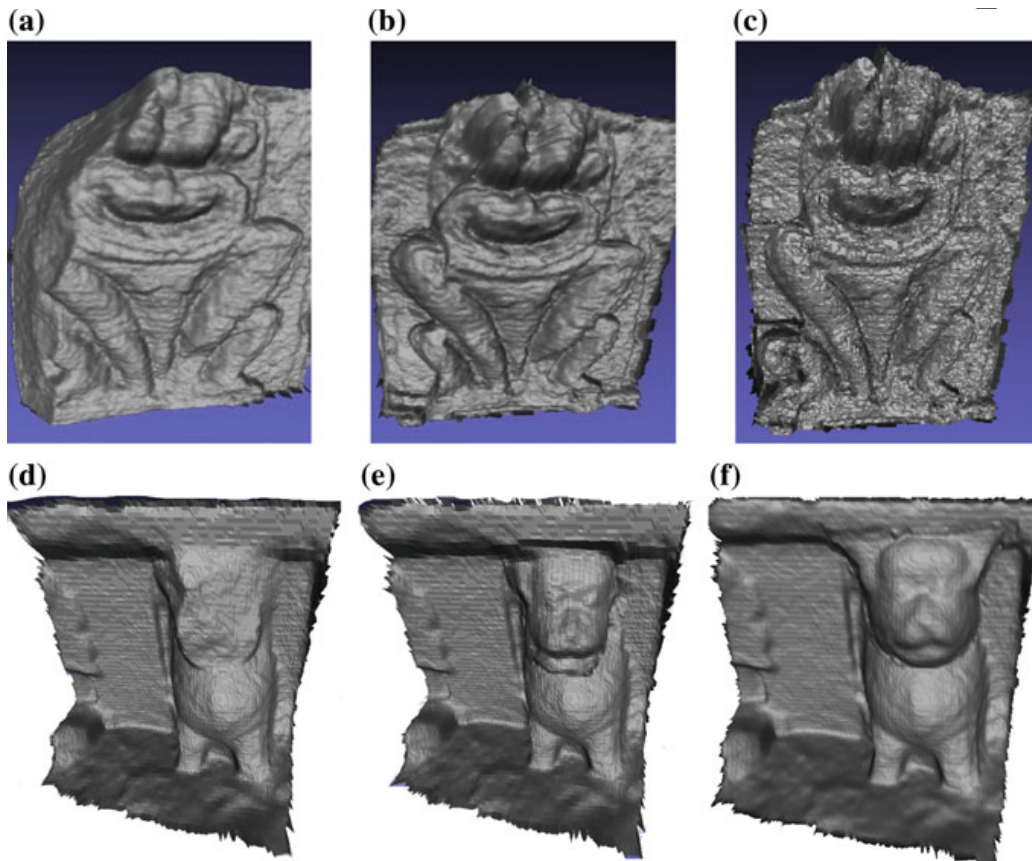**Fig. 16** Comparing the outputs for the examples shown in **a**, **d** as delivered by **b**, **e** the TV-based method, and **c**, **f** the DL-based method
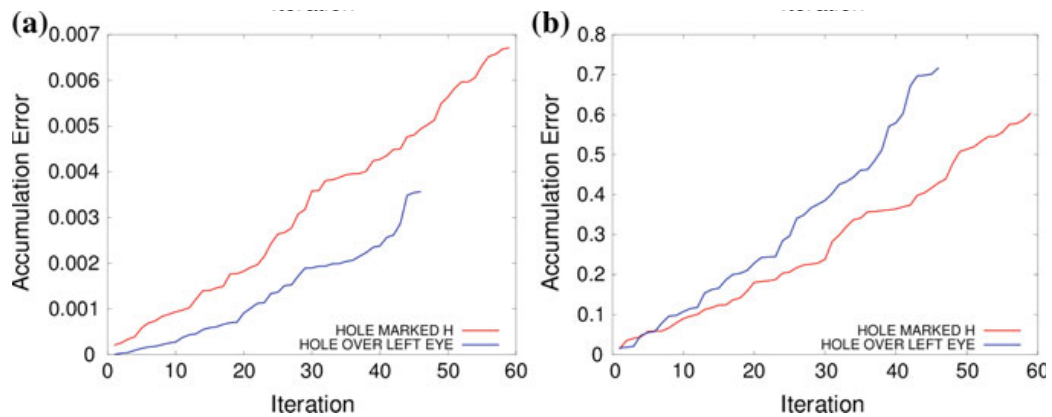

Fig. 17 Accumulation error for **a** TV-based method, and **b** DL-based method. The hole regions considered correspond to $\mathcal{H}$ in Fig. 8b (shown in blue) and the other over the left eye of the same face model (shown in red)

that while the TV-based method results in a finer reconstruction with respect to the original 3D surface, the DL-based method too results in accumulation error that is low enough for it to not introduce any visually jarring reconstruction artifacts.

Although execution time is not a major concern in the scenario that we address, the running time for TV-based inpainting is about 5–10 min, and up to 1 hr for the DL-based inpainting (once the dictionary D is ready). The visually pleasing outputs produced using geometric inpainting for all of the displayed examples indicate the e ectiveness of our methodologies. Rather than attempting an authentic regeneration down to minute details, the goal of our work has been to guide the restoration process. Automated restoration mechanism has the advantage that archaeologists need only acquire scanned self-similar examples corresponding to a damaged structure, and algorithms (of the kind discussed in this chapter) would provide them with completed structures. Such a geometric inpainting pipeline can not only serve as a reference for future restoration tasks but also enable a pleasant virtual walk-through for visualization based on existing evidence. The main limitation of our work is the requirement of at least one self-similar example. At the same time, it is also true that inpainting large holes of the kind we have addressed would be infeasible in the absence of any information about the missing parts.

6   Conclusions

In this chapter, we first discussed a simple, e ective, and non-iterative method to inpaint large, complex missing or damaged regions in 3D models of real-world structures using the neighborhood geometry as a cue along with geometric prior from self-similar examples. The framework involved using corresponding surface points from registered self-similar examples as noisy estimates for the missing region on the damaged model from which an inference of the best underlying geometry was made using tensor voting (TV). Rather than attempting an authentic regeneration down to the minute detail, the goal of our work was to guide and facilitate the restoration process. A high-quality and visually pleasing filling-in or reconstruction of the damaged region was obtained with very low reconstruction errors, verified for scenarios having ground truth information.

Second, we proposed a dictionary learning (DL) and surface gradients-based method to solve the boundary-artifact problem observed upon using the TV-based method when only a single self-similar example (captured in an uncontrolled environment) was available as geometric prior. The results indicate the extent of a visually faithful reconstruction that is possible for even the challenging scenario of e ectively harvesting the geometric prior from a single self-similar example for inpainting large holes. Though the DL-based is e ective in performing smooth inpainting, the TV-based method would be a preferred choice when the

availability of multiple self-similar examples is not an issue. This is because the TV-based method e ectively estimates the most salient surface from the set of noisy data provided by the examples, while the recovered surface in the DL-based method will be biased towards the single self-similar example used.

References

1. Aharon M, Elad M, Bruckstein A (2006) K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Signal Process 54:4311–4322
2. Andriy M, Xubo S (2010) Point set registration: Coherent point drift. IEEE Trans Pattern Anal Mach Intell 32:2262–2275. https://doi.org/10.1109/TPAMI.2010.46
3. Breckon TP, Fisher RB (2008) Three-dimensional surface relief completion via nonparametric techniques. IEEE Trans Pattern Anal Mach Intell 30:2249–2255
4. Callieri M, Cignoni P, Ganovelli F, Impoco G, Montani C, Pingi P, Ponchio F, Scopigno R (2004) Visualization and 3D data processing in the David restoration. IEEE Comput Gr Appl 24:16–21
5. Cignoni P, Corsini M, Ranzuglia G (2008) MeshLab: an Open-Source 3D Mesh Processing System. ERCIM News. 45–46
6. Davis J, Marschner S.R, Garr M, Levoy M (2002) Filling holes in complex surfaces using volumetric di usion. 3D Data processing visualization and transmission. In: First International Symposium on. 428–438. https://doi.org/10.1109/TDPVT.2002.1024098
7. Geuzaine C, Remacle JF (2009) Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. Int J Numer Method Eng 79:1309–1331
8. Gupta S, Castleman K.R, Markey M.K, Bovik A.C (2010) Texas 3D Face Recognition Database. In: Proc. 2010 SSIAI, TX2010, Austin, pp 97–100. https://doi.org/10.1109/SSIAI.2010.5483908
9. In: Google-Art-Project (2011) http://www.googleartproject.com/
10. In: INTACH (1984) http://www.intach.org/
11. In: Kinect (2010) http://www.xbox.com/en-IN/Kinect/
12. In: Photosynth (2010) http://photosynth.net
13. In: p3d (2013) http://p3d.in
14. Jia J, Tang CK (2004) Inference of segmented color and texture description by tensor voting. IEEE Trans Pattern Anal Mach Intell 26:771–786
15. Julien, M, Francis, B, Jean, P, Guillermo, S (2009) Online dictionary learning for sparse coding. ICML. pp 689–696
16. Kulkarni M, Rajagopalan AN, Rigoll G (2012) Depth Inpainting with Tensor Voting Using Local Geometry. In Proceedings, VISAPP
17. Kulkarni M, Rajagopalan AN (2013) Depth inpainting by tensor voting. J Opt Soc Am A opt Image Sci 30:1155–1165
18. Lai K, Bo L, Ren X, Fox D (2013) RGB-D Object Recognition: Features, Algorithms, and a Large Scale Benchmark. Consumer Depth Cameras for Computer Vision: Research Topics and Applications, pp 167–192
19. Levoy M, Pulli K, Curless B, Rusinkiewicz S, Koller D, Pereira L, Ginzton M, Anderson S, Davis J, Ginsberg J, Shade J, Fulk D (2000) The digital Michelangelo project: 3D scanning of large statues. ACM SIGGRAPH 2000:131–144
20. Liepa, P (2003) Filling holes in meshes. In: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing. pp 200–205
21. Mairal J, Bach F, Ponce J, Sapiro G (2010) Online learning for matrix factorization and sparse coding. J Mach Learn Res 11:19–60

22. Padalkar MG, Joshi MV (2015) Auto-inpainting heritage scenes: a complete framework for detecting and infilling cracks in images and videos with quantitative assessment. Mach Vis Appl 26:317–337
23. Patrick P, Michel G, Andrew B (2003) Poisson image editing. ACM Trans Graph 22:313–318
24. Pauly M, Mitra NJ, Giesen J, Gross M, Guibas LJ (2005) Example-based 3D scan completion. In: Proceedings of the third Eurographics symposium on Geometry processing
25. Philippos M, Grard M (2006) Tensor Voting: A Perceptual Organization Approach to Computer Vision and Machine Learning. Synth Lect Image Video Multimed Process. 2:1–136
26. Scharstein D, Szeliski R (2002) A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. Int J Comput Vision 47:7–42
27. Shailaja, T (2013) Monuments gone missing, In: The Hindu. http://www.thehindu.com/features/metroplus/monuments-gone-missing
28. Silberman N, Hoiem D, Kohli P, Fergus R (2012) Indoor segmentation and support inference from RGBD images. ECCV. pp 746–760
29. Sturm, J, Engelhard, N, Endres, F, Burgard, W, Cremers, D (2012) A benchmark for the evaluation of RGB-D SLAM systems. In: Proceeding of the international conference on intelligent robot systems (IROS)
30. Szeliski R (2010) Computer Vision: Algorithms and Applications. Springer, New York
31. Tosic, I, Olshausen, B.A, Culpepper, B.J (2010) Learning sparse representations of depth. arXiv:1011.6656
32. UNESCO (2003) In: Draft charter on the preservation of the digital heritage. UNESCO General Conference 32nd session
33. Verdera, J, Caselles, V, Bertalmio, M, Sapiro, G (2003) Inpainting surface holes. In: Proceedings ICIP 2003, vol 2 II– pp 903–906

Part III
nalysis and Digital Restoration
of rtifacts

Vijayanagara Era Narasimha Bronzes and Sculpture: Digital Iconometry

**Sharada Srinivasan Rajarshi Sengupta S. Padhmapriya
Praveen Johnson Uma Kritika Srinivasa Ranganathan
and Pallavi Thakur**

1 Introduction

The Vijayanagara dynasty rose to political power in 1336 CE under the Sangama brothers Harihara I and Bukka Raya I. The four prominent dynasties of the time were the Sangamas 1336–1485 CE), Saluva 1485–1505 CE), Tuluva 1505–1570 CE), and Aravidu 1570–1646 CE). Vijayanagara inscriptions extensively record the construction of new temples and expansion of existing temples through land grants, and donations of jewelry for the deities and related rituals. The architectural heritage and related religious and cultic practices of the Vijayanagara region were diverse [1–4]. Two important sectarian practices within the Hindu religion thrived at the capital of Vijayanagara or Hampi since the fteenth century which steered the course of the artistic output ranging from sculpture, the making of metal icons and mural art. One was the worship of the river goddess Pampa, a tributary of the Tungabhadra, from which the name Hampi is derived [2]. Her consort, Virupaksha who is considered a form of Shiva, associated with destructive forces) is the tutelary deity of the Kingdom, and is worshipped in the grand Virupaksha temple erected at the Hemakuta hill in the Vijayanagara period. The worship of Vishnu, the lord of preservation, and his avatars or incarnations emerged later to become immensely popular in the region. The mathas or monasteries and temples, dedicated to both Shaiva and Vaishnava deities, were closely associated with the royalty.

The renowned Vijayanagara ruler Krishnadeva Raya 1509–1530), whose capital at Hampi in Karnataka was compared to Rome by the Portuguese Domingo Paes [5] and whose rule extended into the Andhra and Tamil regions, had a set of life-size-inscribed copper alloy portraits of himself and his two wives installed in

S. Srinivasan   )   R. Sengupta   S. Padhmapriya   P. Johnson
U. Kritika   S. Ranganathan   P. Thakur
National Institute of Advanced Studies, Bangalore, India
e-mail: sharasri@gmail.com

the Tirumala temple in Tirupati. Under his rule, the Tamil Dravida style came widely into vogue in southern Indian temples [6]. One of the popular cults was the worship of Narasimha, the man–lion incarnation of Vishnu [7, 8]. The iconic granite colossus of Narasimha at Hampi represents the deity in his most fearsome form, regarded as the protector of the city of Vijayanagara. The original image was greatly damaged following the sack of Vijayanagara in the battle of Talikota at the hands of the Bijapur and Golconda Sultans. This is captured in the nineteenth-century photographs of Greenlaw [9, 10]. The image in its present form had undergone some alterations following a somewhat controversial programme of restoration undertaken in the 80s.

2 Digital Approaches to the Study of Vijayanagara Sculpture and Bronzes

Increasingly, scienti c investigations are coming to the aid of art historical research. For example, Srinivasan [11, 12] has used archaeometric and archaeometallurgical ngerprinting techniques in the study of South Indian bronzes by analyzing about 150 images in leading collections. These studies have shown that the lead isotope ratios and trace elements for Co, As, Bi, Ni, and Sb) of Vijayanagara bronzes fourteenth–sixteenth century) were signi cantly different from Chola and later Chola bronzes ninth–thirteenth century) so as to suggest that their ore sources were different which provided a diagnostic tool for telling them apart [1, 13]. As another example, from the authors' studies, evidence could be found for iconometric shifts in bronzes mirroring the changing ritual practices. Whereas Rama, an avatar of Vishnu emerged more prominently as a major deity by the Vijayanagara period with a full- edged temple of Hazara Nama dedicated to him, in contrast, in the Chola period the iconometric proportions more closely re ected his stature as the prince of Ayodhya and closer to that of an idealized man [14].

This chapter also explores techno-cultural approaches in exploring insights into the signi cance of the Narasimha cult and in retracing the iconographic aspects of the damaged and subsequently altered Narasimha image through comparisons with depictions in various media including bronze. For these purposes, this study made use of laser-scanned imagery of the colossal Narasimha statue at Hampi undertaken by KCST, one of the collaborating teams of the IDH Hampi project. 3-D reconstructions were made by using a 3D geodetic framework through high-resolution terrestrial scanning and global positioning systems [15]. The Archaeological Museum at Kamalapura near Hampi in Northern Karnataka and the Chandragiri Fort Museum in Southern Andhra Pradesh houses several of the stone and metal images from the Vijayanagara period which were studied for comparative purposes. Chandragiri rose to prominence in 1568 CE as a seat of power for the later Vijayanagara kings and as the fourth capital before it was annexed in 1646 CE by the Golconda Sultans. The Raja Mahal Palace has been converted into a museum.

Many images which were found in the immediate vicinity of Hampi and which were originally in the Kamalapura Museum at Hampi were shifted here according to the museum authorities.

3 Iconography of Narasimha in Vijayanagara Sculptures

All over southern India and particularly during Vijayanagara period, a variety of Narasimha icons were made in varied media, the iconographic and iconometric aspects of which are explored further. In a similar vein, Fig. 1 is an image of Lakshmi Narayana from Kamalapura Museum near Hampi, depicting Vishnu or Narayana with consort Lakshmi, with the iconometric design of the traditional *talamana* proportions as indicated in the gure.

At the same the image also sometimes has the attributes usually found in serene depictions of Vishnu, as the lord of preservation, such as the four hands with Sankha conch shell) and Chakra Sudarshana wheel), which establishes his connection with Vishnu and the Vaishnava tilaka on his forehead between two eyebrows and chest [16]. The iconographic depictions of Narasimha images after the ninth century include Girija Narasimha, Kevala Narasimha, Sthauna Narasimha, Yanaka Narasimha, and Lakshmi Narasimha. The earliest non-Shaiva Vijayanagara cult was probably of Narasimha, a popular deity in southern India before the advent of the Vijayanagara Empire [17].

The temples to Narasimha are known predating the Vijayanagara period in parts of Tamil Nadu, Karnataka, Andhra Pradesh and Telangana. Although the site of Vijayanagara city or Hampi is more widely associated with the mythology of Ramayana, a wide range of Vaishnava deities were also worshipped there with Narasimha occupying a special place. The Saluva and Tuluva rulers of Vijayanagara were devotees of Narasimha. This cult was also propagated by the Madhavas, followers of Madhvacharya, who venerated Narasimha, Krishna, and the folk deity Vithala. From the early fourteenth century until the early seventeenth century, Madhava ascetics were said to have been active around Hampi and Anegondi. Narasimha is also worshipped by the Shrivaishnavas at temples such as at Ahobilam in Kurnool district of Andhra Pradesh [17]. According to inscriptions, Krishnadevaraya granted villages to temples and presented precious gold jewelry and gold coins to numerous temples including the Narasimha temples at Ahobilam and Simhachalam ibid.). The important Vijayanagara era temples and those related to the Narasimha cult in Andhra Pradesh are indicated in Fig. 2. As discussed further, the colossal Narasimha statue at Hampi was erected under Krishnadevaraya. This iconic Lakshmi Narasimha image from Hampi represents the deity in his most fearsome form, who was considered as the protector of the city of Vijayanagara. Narasimha images are also widely depicted at numerous temples in Hampi such as in the pillars of the Vitthala temple in the maha mandapa and in stucco in the Virupaksha temple. Some intriguing comparative aspects of Narasimha images made in the Vijayanagara period are highlighted further using digital iconometric studies.

Lakshmi Narayana

A = 28 mm

B = 18 mm

C = 2.5 mm

D = 18 mm

E = 18 mm

F = 18 mm

Image Details

Lakshmi Narayana: Medium Stone, Vijayanagara Period
Location: Kamalapur Museum, Bellary District, Karnataka

Image Size = Width 127.5 mm X Height 170.67mm
Actual Image Body Height = 148 mm

Divisions (Narayana)
A = 28 mm, B = 18 mm, C = 2.5 mm, D = 18 mm, E = 18 mm, F = 18 mm

Average Divisional Height ADH = (B+D+E+F) / 4 = 72/4 = 18 mm
Crown (A) = 28 mm, which is **155.55%** of **ADH** (18 mm)
B = 2.5 mm, which is **13.88%** of **ADH** (18 mm)

Fig. 1 Lakshmi Narayana image, Kamalapura Museum, Hampi

Fig. 2 Places of Narasimha cult in Andhra Pradesh

3.1 Lakshmi Narasimha Hampi) and Iconometric Studies Using Laser Scanning

The magni cent 6.7 m high Lakshmi Narasimha colossus and shrine in Krishna-pura was erected by Krishnadevaraya. An inscription in front of the temple mentions that the monolithic statue of Lakshmi Narasimhadeva was consecrated on April 2, 1528 CE on the order of Krishnadevaraya by the priest Arya Krishna Bhatta [18]. In the following year, the king gifted two villages for food offerings to the deity. This shrine is one of the three main temples in the city built entirely by the monarch Krishnadevaraya. This Lakshmi Narasimha monolith is his last construction demonstrating his special af nity to this deity [17]. Despite its sadly

Fig. 3 Photograph of Narasimha Monolith by Alexander Greenlaw. Modern Positive 2007) from Waxed-paper Negative, Photographers Ref. 29–56, 1856, 445 × 402 mm, ACP: 99.01.0044. Image source accredited to the Alkazi Collection of Photography

mutilated condition, the giant Lakshmi Narasimha monolith remains one of the most impressive monuments in the city. It has mistakenly been identi ed as Ugra Narasimha as it is popularly known today associated with his ercer depiction. As seen from Greenlaws earlier wax-paper negatives of 1856 [9, 10, 19], the image has both the knees broken off, right at the level of the pelvis itself Fig. 3). Thus the state that the image is currently in, depicting Narasimha sitting cross-legged with a yogapatta is a feature added during the later restoration attempts in the 80s; on the basis of which the image is sometimes wrongly identi ed as Yoga Narasimha.

As seen in the Fig. 4 and the laser-scanned image Fig. 5, the granite sculpture of Narasimha at Hampi is depicted seated in a frontal posture framed by a prabhavali or arch with a kirtimukha or face of glory). Normally it is dif cult to go to the back

Fig. 4 Granite monolith of Narasimha at Hampi

of the Narasimha shrine and there is not enough space to take photographs. However, laser scanning came into use for the documentation of these less accessible areas. The laser-scanned rear view Fig. 6) thus gives an unusual perspective. Unusually, unlike other Narasimha images, Narasimha is depicted seated on the three coils of the snake Adishesha, canopied by his seven hoods, taking on the attributes of the major deity Vishnu. The rear laser-scanned view shows the lovely depiction of the serpent on the three coils.

The iconometric analysis on the laser-scanned Narasimha image shows that the image was built on the basis of silpasastras, and generally followed the talamana system of traditional iconometry. According to the talamana canon [20], one tala is said to consist of 12 angulas, spanning the measurement of the head from hairline to chin. From the base of the neck to the middle of the chest at the level of nipples it counts as one tala; from the chest to navel another one tala, and from navel to bottom for the seated image there is yet another tala. These four talas spanning the

Height of Head   =67 MM

Shoulder to Chest=67 MM

Chest to Navel   =66 MM

Navel to Bottom =66 MM

Image Details
Lakshmi Narasimha
Location: Hampi ,Vijayanagara Period ,16th Century
Photoshop Image Size = 1058.33 MM X 846.67 MM
Actual Image Body Height = 602 MM
(ADH)Average Divisional Height = 66. 5 MM

Fig. 5 Laser-scanned image of Narasimha monolith at Hampi

body from the hairline to the base are equally spaced. Taking into account a part of
the crown and the three coils of Adisesha, the proportions come to six talas.

3.2   Iconometric Comparisons with Other Narasimha Images in Bronze and Stone

Interestingly, somewhat similar proportions are also followed in two other
Vijayanagara Narasimha images discussed further in terms of the main bodies of the
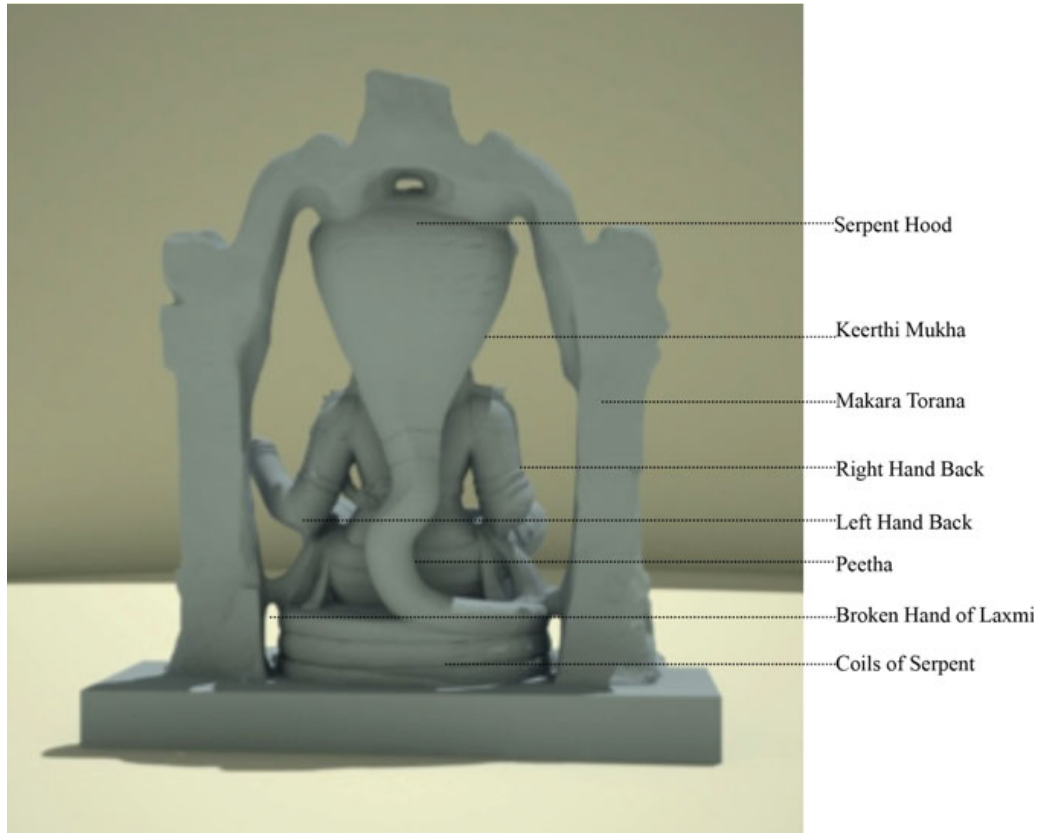
Fig. 6 Rear view of laser-scanned Lakshmi Narasimha image

images being roughly distributed into four talas from the top of the forehead to the base of the seated gures. However, it must be mentioned that in these cases iconometric study was done on digital photographs since laser scanning could not be done which would have given a more accurate picture as in the case. One of these is a Narasimha image, now preserved at Kamalapura Museum, Hampi, made of the medium rather than hard stone. This image is of Kevala Narasimha, i.e., Narasimha without consort Lakshmi Fig. 7). An interesting aspect is that is the image is made in high relief and not totally a 3-dimensional one. But as the image is almost intact, except for the back stele and right arm, the features and attributes are observed clearly. Narasimha is seated in padmasana i.e., the seated posture of a yogi) on a high pedestal. He holds the sankha or conch and the chakra—sudarshana, or discus and his left hand is in varada mudra, granting boons. These gestures are akin to the major deity Vishnu, whose incarnation he is. The right hand is broken, which is most probably in Abhayamudra, as seen in other images. The iconometric analysis shows some similarities with the Lakshmi Narasimha image discussed before. The measurement of one tala of the head, consisting of 12 angulas, is repeated over the chest and abdomen and slightly lower than navel. According to talamana conventions, the length of the palm is same as the length of head, which is also generally followed here. The colossal image of Narasimha from Hampi does not have any of the hands intact so such a comparison is not

Image Details

Narshimha: Medium Stone, Vijayanagara Period
Location: Kamalapur Museum, Bellary District, Karnataka

Image Size = Width 56.92 mm X Height 85.33 mm
Actual Image Body Height = 36 mm

Divisions
A = 10 mm, B = 9.5 mm, C = 9.5 mm, D = 6.5 mm, E = 8.5 mm

Average Divisional Height ADH = (A+B+C) / 3 = 27/3 = 9 mm
Hand (B) = 8.5 mm, which is **87.928%** of **ADH** (9 mm)

Fig. 7 Kevala Narasimha stone image, Kamalapura Museum

possible here. Comparisons with a Vijayanagara period bronze image from the Chandragiri Palace Museum Fig. 8) of Lakshmi Narasimha with the two previous images are also relevant. Lakshmi is depicted in the bronze gurine sitting on

Fig. 8 Chandragiri Palace Museum

Narasimha's lap Fig. 9); which gives an insight into what the intact monolith of Lakshmi Narasimha at Hampi would have looked like before it was damaged.

The small bronze gurine is intricately done with all necessary details and ornamentation. Narasimha is represented here with four arms, holding sankha and chakra—sudarshana, one in Abhayamudra and another embracing devi. He wears a high kiritamukuta and ornaments as necklaces, bracelet, and armlets, girdles, and anklets. His lower garment touches ankle and intricately done with the suggestion of folds in drapery. He sits on a full bloom lotus and another small lotus is placed under his right foot. A comparison of the lion faces of the three images show the similarities and diversities in the expression of Narasimha depicted in different media in the stylized eyes and grimace. This image also shows that the body of the Narasimha is divided into about four talas like the other two from the forehead to the region below the folded leg. Lakshmi who sits upright as the consort of Narasimha Vishnu) wears a stanapatta or breast-band which is speci c to Lakshmi or Sridevi's iconography as consorts of Vishnu. Her towering crown is different from the one worn by Narasimhas. Though both can be categorized as kiritamukuta, the crown of the god is conical and the goddess one is stepped. She holds a lotus in her left arm and her consort the right arm. Her lower garment is detailed with circular ower motifs and rows of parallel line suggesting folds. A lotus is placed under her feet, similar to her consort.

A Vijayanagara Parvati bronze Fig. 10) follows the Chola style of bronzes in the stance with a right hand in kapittha mudra and left in Dola hasta. The Kevala Narasimha stone image discussed before also shows some similarities in the

Lakshminarasimha -Chandragiri

A = 84mm

B= 62mm

C= 5 mm

D= 63mm

E=62mm

F=63 mm

G=24mm

Image Details

Lakshminarasimha ,Medium Bronze,Vijayanagara Period 17th century

Location : ASI museum ,Chandragiri,Andra Pradesh

Iconomety Details

Photoshop Image size = 764.8 mm X 905.9 mm
Image Body Height   = 500 mm

Divisions

A = 84mm, B=62 mm, C= 5 mm ,D =63mm , E= 62 mm ,F = 63 mm ,G =24mm

Average Divisional Height (ADH) = B+D+E+F / 4 = 250/4 = 62.5 mm

Crown A = 84 mm which is   134.4 % of ADH

(C+ G)= 5+ 24 = 29 mm which is 46.9 % of ADH

Fig. 9 Lakshmi Narasimha bronze image, Chandragiri Museum

A= 13.5mm

B= 7mm

C= 2mm

D= 7mm

E= 7mm

F= 7mm

G= 7mm

H= 7mm

I = 7mm

J= 2mm

Image Details

Parvati - Medium Bronze ,Vijayanagara Period
Location : Achaelogical Museum Chandragiri ,Andhra Pradesh

Photoshop Image Size = Width 81.78 mmX 111.42 mm Height
Actual Image body height = 66.50 mm

Divisions
A=13.5 mm,B=7mm,C=2mm,D=7mm,E=7mm,F=7mm,G=7mm,
H=7mm,I=7m,J=2mm

Average Divisional height (ADH) =(B+D+E+F+G+H+J)/7 = 42/6 = 7mm
(C+I) =2mm+2mm= 4 mm which is 57.14 % of ADH (7mm)

Height of Makuta (crown) = A = 13.5mm which is 192.85 % of ADH (7mm)

Fig. 10 Parvati, Chandragiri Museum

decoration to the Parvati image, in the loops on the shoulders, indicating some of the stylistic af nities cutting across stone and bronze during the Vijayanagara period. This image is seen to follow the navatala proportions akin to those for a goddess image Fig. 10).

4   Conclusions: Insights on Talamana Canon

Study of the Narasimha and other Vijayanagara era stone and bronze images reveal some of the features of the iconographic and iconometric conventions followed in the making of images during the Vijayanagara period. The three Narasimha images examined show similarities in the use of four talas for the main body of Narasimha with an additional two talas for the crown and seated base, although they range in sizes from the miniature bronze Lakshmi Narasimha in Chandragiri Museum to the colossal monolith from Hampi. The bronze Lakshmi Narasimha and the iconometric and iconographic study gives an idea of the comparable features in the damaged monolithic Lakshmi Narasimha image from Hampi. The depictions of the bases vary in that the granite monolith of Lakshmi Narasimha has the coils of the serpent in the base, the bronze Lakshmi Narasimha has a lotus base, while the Kevala Narasimha stone image from Kamalapura has no base and a stellate platform below, almost recalling to Hoysala imagery. Thus although the images generally followed sastraic prescriptions in iconographic and iconometric conventions, some interesting examples of creative diversity, syncretism and shared in uences across various regions and media are seen. The use of laser scanning and digital iconometry were found to be particularly helpful in examining these aspects.

References

1. Srinivasan S 2016) Tamil Chola bronzes and Swamimalai Legacy: metal sources and archaeotechnology. J. Metals 68 8):2207–2221
2. Settar S 1990) Hampi: a medieval metropolis. Kala Yatra, Bangalore
3. Fritz J, Michell G 2014) Hampi, Jaico, Mumbai
4. Srinivasan S on behalf of IDH team) 2016) Hampi: splendours of a World Heritage Site, NIAS Report Oct 2016 BRICS YSF Bangalore Conclave. http://www.nias.res.in/publication/hampisplendors-world-heritage-nias-report-october-2016
5. Sewell R, Nunes F, Peas D 1900) A forgotten Empire, Vijayanagara: a contribution to the history of India
6. Fritz JM, Michell G eds) 2001) New light on Hampi: recent research at Vijayanagar. MARG Publication, Mumbai
7. Ramachandra Rao SK 1990) Pratima Kosha: encyclopedia of Indian Iconography, vol 3. Kalpatharu Research Academy Publication, Bangalore

8. Meister MW 1996) Man and man-lion: The Philadelphia Narasimha. Artibus Asiae 56
   3 4):291–301
9. Ware M 2008) Greenlaw's Calotype process. In: Michell G ed) Vijayanagara: splendour in
   ruins. Mapin Publishing Pvt. Ltd. and Alkazi Foundation of Photography, Ahmedabad
10. Gordon S 2012) Monumental visions: architectural photography in India. Unpublished
    thesis, SOAS, University of London, 1840–1901
11. Srinivasan S 1996) The enigma of the dancing pancha-loha   ve-metalled) icons: art
    historical and archaeometallurgical investigations on south Indian metal icons. Unpublished
    PhD thesis, Institute of Archaeology, London
12. Srinivasan S 2015) Bronze image casting in Tanjavur district, Tamil Nadu, Ethnoarchae-
    ological and archaeometallurgical insights. In: Metals and civilisations, proceedings of the
    seventh international conference on beginning of use of metals and alloys. National Institute
    of Advanced Studies, Bangalore, pp 209–216
13. Srinivasan S 1999) Lead isotope and trace element analysis in the study of over a hundred
    South Indian metal icons. Archaeometry 41:91116. https://doi.org/10.1111/j.1475-4754.1999.
    tb00854.x
14. Srinivasan S 2013) Iconographic trends in Rama worship: insights from techno-cultural
    studies of bronzes. In: Conference on The Ramayana in literature, society and the arts, Feb
    1–2, CPR Publications, CP Ramaswamy Aiyar Institute of Indological Research, Chennai,
    pp 345–362
15. Prithviraj M, Vijay UT, Ajay Kumar GC et al 2012) Geo-spatial data generation and
    terrestrial scanning for 3D reconstruction. Int J Adv Res Comput Commun Eng 1 9):601–604
16. Welankar V 2009) The iconography of Kevala Narasimha: a reappraisal. South Asian Stud
    25 1):113–130. https://doi.org/10.1080/02666030.2009.9628702
17. Verghese A 1995) Religious traditions at Vijayanagar. Manohar-AIIS, New Delhi
18.   1882) Epigraphia Indica, vol I. ASI, New Delhi
19. Michel G ed) 2008) Vijayanagara: splendour in ruins. The Alkazi Collection of
    Photography, Mapin Publications
20. Gopinatha TA 1914) Elements of Hindu iconography, vol I. Part II, Motilal Banarsidass,
    New Delhi

Digitizing Hampi and Lepakshi Temple Ceiling Murals: Towards Universal Access and Tangible Virtuality

Uma V. Chandru, Namitha A. Kumar, C. S. Vijayashree and Vijay Chandru

1 Introduction

Paintings, frescoes and murals are a significant and valuable part of our ancient history and cultural heritage. Digitization of these historical treasures gives us a chance to document and preserve these heritage artefacts using techniques like image processing and analysis and computer graphics. Such digital technologies can also augment these artefacts with additional information to enhance user enjoyment and immersive experience. Digital imaging and virtual restoration contribute immensely to historical studies, art stylization studies, heritage experiences through 3D reconstruction and further engagements with ancient art and historical sites [1].

The Indian Digital Heritage (IDH) research project is a pioneering, multi-institutional, multidisciplinary project bringing together diverse experts including historians, art historians, folklore researchers, designers, a design anthropologist, an architect, a filmmaker, photographers and several digital technologists. The interdisciplinary Vijayanagara and post-Vijayanagara murals digital heritage project (2011–2016) at the International Institute for Art, Culture and Democracy (IIACD) was one of over twenty projects supported by the Department of Science and Technology, Government of India under the IDH project. The aims of this project were to research, capture, digitize and archive the tangible and intangible heritage of the ceiling murals in two important sites of Vijayanagara and post-Vijayanagara period, namely Virupaksha temple at Hampi, Karnataka and Veerabhadraswamy temple at Lepakshi, Andhra Pradesh.

Hampi being the capital of Vijayanagara Empire was the epicentre of all cultural activities. Many glorious temples and other monument were constructed during the extensive royal patronage of the Vijayanagara rulers. This world heritage city hence was chosen as the focus of the Indian Digital Heritage projects. The focus of our research was on the Virupaksha temple which is the principal sanctuary in Hampi. The ceiling of this temples rangamantapa built during the reign of Krishnadevaraya (1509–1530) in 1510 CE is painted with images of varied religious and social themes.

U. V. Chandru ( ) · N. A. Kumar · C. S. Vijayashree · V. Chandru
International Institute of Art, Culture and Democracy, Bangalore, India
e-mail: uma@iiacd.org

As the style and antiquity of these paintings have been the source of intense academic debate, we focussed on the study of established examples of Vijayanagara paintings for comparative analysis. Lepakshi in Ananthapur district of Andhra Pradesh was selected as one of the key sites for our comparative research as the murals in the Veerabhadraswamy temple are the best examples of Vijayanagara paintings. The previous experiments of Vijayanagara artists seem to culminate here to produce a well-composed visually supreme balanced art. The continuity of the living art tradition of the hand-painted Kalamkari textiles of Kalahasthi in Andhra Pradesh has close resonance with the murals in Lepakshi village which was a vital trade centre during the Vijayanagara Empire. Epigraphic records indicate that the Veerabhadraswamy temple was patronized by two chieftains, Virupanna Nayaka and his brother Viranna (1537 CE) during the reign of Achyuta Deva Raya (1529–1542), younger brother of Krishnadevaraya.

The mythological themes found in Hampi paintings also extend to Lepakshi. What distinguishes Lepakshi murals from Hampi is the division of the pictorial space. The ceiling of the Hampi rangamantapa (Fig. 1a) is without any architectural elements. The barren ceiling is divided into vertical and horizontal zones of action through ornamental bands with varied motifs. A particular episode is restricted to its allotted space. The paintings rarely overflow into another space. The ceiling of the Lepakshi natyamandapa (Fig. 1b) is interspersed with multiple pillars and beams. The long band that emerged from this demarcation was used as the visual plane for the paintings.
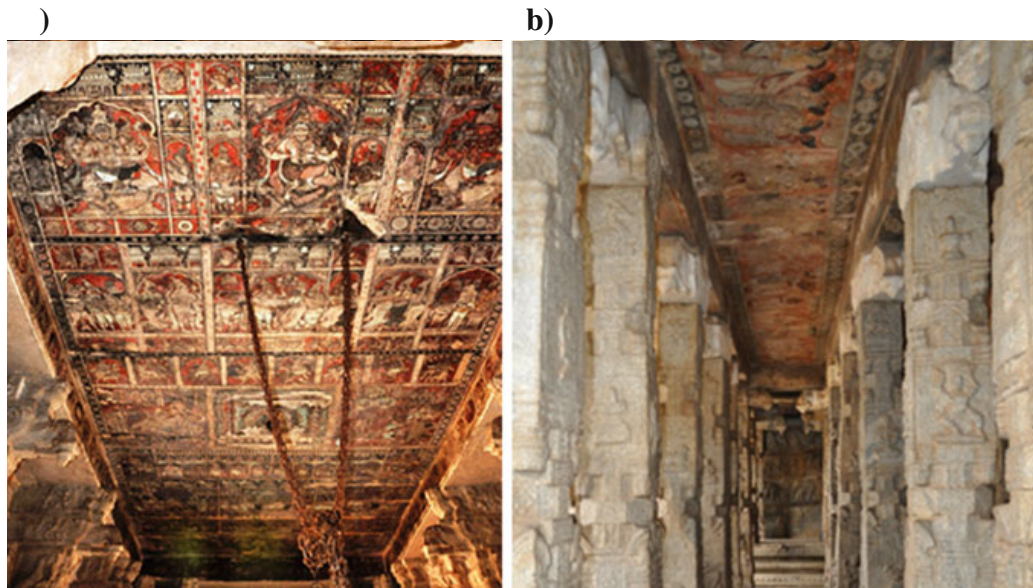

Fig. 1 a Hampi Virupaksha rangamantapa temple. **b** Lepakshi Veerabhadraswamy natyamandapa temple

The layout of the ceiling murals in these two sites substantially influenced the design of their respective digital representations that our team co-created in the form of interactive ceiling mural plans that we discuss below.

2   Digital Preservation of Murals

Murals from the past, especially those that express the religious, social and political views of their period have intrinsic artistic and cultural value. They play a vital role as material evidences for construction and reconstruction of history. Unfortunately, murals are more vulnerable than any other art form to damage from natural as well as human sources. Digital archival for preservation of the heritage of temple and other murals in South India is critical as a large number of the paintings have been damaged due to unethical or improper methods of conservation, neglect, vandalism, white washing, sandblasting, water seepage, sunlight, insects, bats and other reasons. Murals in several living temples including Hampi Virupaksha and Lepakshi Veerabhadraswamy temples are also subject to damage from oil and soot.

The recently formulated National Policy for Conservation of the Ancient Monuments, Archaeological Sites and Remains (NPC-AMASR), Archaeological Survey of India (ASI), February 2014 explicitly states that Elements of a monument, such as murals, sculptures, inscriptions and calligraphy should not be restored. Monuments must be restored on the basis of documentary, archaeological or architectural evidences, and not on the basis of any conjecture [2]. Although ASI conservation guidelines do not permit repainting, we find sites where murals have been touched up or repainted. The urgent need for documenting, interpreting and conserving murals, especially those that are endangered was highlighted by scholars, conservators and other participants in the 2008 seminar Painting Narrative: Mural Painting Tradition in the 13th and 19th centuries (January 23–27, 2008) at Dakshina Chitra, Chennai. A digital repository of murals to enable preservation was discussed at this meeting.

While there has been increasing interest in the field of documentation of paintings over the past two decades, most of the art institutions, museums and other cultural institutions in India continue to encourage the traditional methods of documentation of murals like sketching, drawing, photography and videography. The drawings, sketches, photographs, slides and videos remain under the control of the museum authorities and cultural institutions and are often difficult to access. No open access digital repository of murals in India was available when we began our IDH project in 2011. While the IGNCAs South Regional Centre has done the documentation of around 20 mural sites and is preparing a database for the same, the documentation is not available to the public and research scholars need to take permission to access this repository.[1]

An interdisciplinary team led by Professor Baluswamy was working on documenting Nayaka murals in Tamil Nadu when we began working on the Hampi and

---

[1]http://ignca.nic.in/src_projects.htm.

Lepakshi Murals. It was encouraging to find that some of the researchers working on Nayaka murals were departing from the conventional documentation methods and were starting to utilize innovative approaches to documentation of murals. MV Bhaskars Moving Murals project funded by India Foundation for the Arts, for example, focus on the digital restoration and replication of the 17th century Ramayana murals at the Venugopala Parthasarathy Temple at Chengam, Tiruvannamalai District, Tamil Nadu through the living art form of Kalamkari.[2] The project explores alternative methods of mural conservation, reconstruction and replication by experimenting with media like Kalamkari and digital animation. Bhaskar photographed the murals in small parts and then digitally stitched them together into the whole panel. As the high resolution images could not be shared online, to facilitate sharing, Bhaskar creates a digital tracing image of each mural with a stylus and touchpad which results in Scalable Vector Graphics (SVG). This line art can be shared and used to study rich details of the murals.

Internationally, there were attempts to engage with more interactive and experiential projects with multidisciplinary team engaged in digital humanities. Most notable amongst these is the well-funded Virtual Caves project of Digital Dunhuang[3] at the world cultural heritage in Dunhuang, China. Located in China's Gansu province, this ongoing project has been focused on the Mogao Grottoes, which have over 45000 m^2 of incredible Buddhist mural paintings dating back from the fifth to the fourteenth centuries and has enabled viewers to experience the exquisite murals without entering the grottoes. This project involved laser scanning of the caves and ultra-high-resolution digital photography of the paintings. The state of art technology used for photography enabled them to display murals which are invisible in natural light or obstructed by structural members in a few caves.[4]

A fully-interactive virtual 3D exhibit titled Pure Land: Inside the Magao Grottoes at Dunghuang was directed by Professor Jeffrey Shaw and Sarah Kenderdine of City University of Hong Kong, in collaboration with Dunhuang Academy, is an innovative 3D digital representation of Cave 220, which is one of the Caves of the Thousand Buddhas. Through a custom app on an Ipad mini, wearing 3D glasses and standing in the centre of a darkened tent equipped with six projectors and innovative 3D technology, viewers can explore and interact with the 360 virtual cave filled with richly detailed paintings like the seven figures known as medicine Buddhas. This extraordinary exhibits virtual repainting of the dull pigments of the seven Buddhas which brings out the original vivid colours researched at the Dunhuang Academy. Viewers can also explore the ceiling of the cave, focus on the details of the paintings by zooming in and bring dancers and musical instruments out of the paintings to hover and perform. This is an example of virtual reality applications in cultural heritage allowing visitors to engage with a digital facsimile of the caves [3].

Given our interest in working towards tangible virtuality of the Lepakshi and Hampi temple murals, we wanted to use laser scanners and work with the technology

---

[2]http://chengammurals.org.

[3]http://en.dha.ac.cn/.

[4]http://idp.bl.uk/.

teams in the IDH project on walkthroughs of the Virupaksha temple murals in the world heritage city of Hampi and also in Lepakshi. However, the IDH Hampi technology teams working on walkthroughs and other virtual technologies were focused on the Vittala temple and our project's budget did not enable us to work on walkthroughs independently. Hence, after researching, capturing and stitching the high-resolution images of Lepakshi murals, inspired by the ceiling mural plans in Pachener [4] and Dallapiccola [5], publications on Hampi, our design anthropologist suggested that we digitally archive the tangible and intangible heritage of murals in Veerabhadraswamy temple at Lepakshi and Virupaksha temple at Hampi Virupaksha temple in the form of interactive ceiling mural plans that are web-enabled and accessible to all. In this chapter, we explain our process of creating the interactive ceiling plans and how this project can be a precursor to a further engagement with a combination of virtual reality, augmented reality and tangible virtuality.

2.1   Image Capture for the Interactive Ceiling Murals Plan

IIACDs interdisciplinary murals project team included art historians, technologists, a design anthropologist, photographers, designers, artists, art and folklore researchers, conservation and heritage tourism experts. After conducting primary and secondary textual research, several field visits were made to study the murals in Hampi and Lepakshi as well as to other mural sites in Karnataka, Andhra Pradesh, Kerala and Tamil Nadu. During these visits, our research and photography team captured low and medium resolution images of the murals. In Hampi and Lepakshi, they interviewed and gathered narratives from temple priests, tour guides, historians and folklorists for further study.

In June 2011, after obtaining permission from the ASI Hyderabad Circle, Vinod Raja photographed the Lepakshi Natyamandapa ceiling murals with a Canon EOS 5D Mark II camera, a Zeiss 50 mm lens, kinoflo lights, reflectors, tripods, a custom-made monorail slider and other professional equipment to capture the high-resolution mural images. RAW images were converted to TIFF images and stitched into panels on Panorama Tools Graphic User Interface (PTGUI). Stitched panels were edited in Photoshop with the help of an expert followed by a minor-color correction.

In September 2012, we were invited to present our work on the Lepakshi murals to the Lepakshi community and other visitors at the Lepakshi Utsava when we were working on the digital capture of the remaining areas of remaining areas of mukhamandapa, rangamandapa, cave area, pradakshana, ardhamandapa and prakara murals. We only had low and medium resolution images of the areas. We used these images and the PTGUI stitched images of the Natyamandapa images to quickly develop a beta version of the Lepakshi Interactive Ceiling Murals Plan (LICMP), based on Raos [6] plan of the Lepakshi temple which had some inaccuracies.

In June 2015, Vinod Raja and his team completed the high-resolution digital capture of the ardhamandapa, mukhamandapa, pradakshana area, cave area, Raghunatha shrine and prakara murals. They worked with two Canon Mark ID cameras with Zeiss

50 mm, Canon 24 mm Tilt Shift lens L series, Canon 24–70 L series with a slider. Two battery operated LED lights replaced the Kinoflo. Laser pointers were used for alignments to ensure that they were exactly below the centre point of the mural in the ceiling to avoid any distortion or foreshortening. Raja (in private communication) points out that he and his team literally worked like the masons who constructed the temple carrying all the old alignment tools and the modern spirit level.

No manipulation was done in the images except stitching as the alignments were precise and he did not need to warp the images. The challenge, however, was to correct each stitching error using RAW files without warping or distorting the images and ensure that even the corners and the borders are seamless. Only slight colour correction was done. There were others challenges that needed to be addressed during the capture of the Lepakshi murals. The prakara and other areas were not in a straight line and they hence had to follow the curve and re centre the image on each setup and work creatively with the slider. The longer panels in Lepakshi were also particularly challenging, especially the Manuneethi Cholan panel which is the longest panel in the natyamandapa.

In July 2011, a team from the Digital Hampi lab at National Institute of Design (NID) went to Hampi to photograph the Virupaksha temple bazaar to recreate the bazaar and social life of the Vijayanagara period digitally [7]. As the shops and residences in the bazaar were being demolished on that day (In 2011 because of haphazard development of commercial establishments near the Virupaksha temple, the deputy commissioner of Bellary district ordered the demolition of the bazaars [19]), they photographed the Virupaksha temples rangamantapa ceiling murals with a Hasselblad H3DII-39 of focal length 50 mm. The images were later stitched into a full rangamantapa painted ceiling plan using Giga Pano software. No image manipulations were carried out.

We shared our Lepakshi natyamandapa images with the NID, Bangalore IDH project team in exchange for their Hampi rangamantapa ceiling murals images in our Hampi Interactive Ceiling Murals Plan (HICMP) which is described below.

After comparing and analyzing the mural paintings captured at Lepakshi with the Hampi murals obtained from NID, our research team also made further field visits to temples with murals in Hiriyur, Sibi, Sompalem, Srirangam, Kanchipuram, etc. to capture and study the murals.

2.2 Codesign and Development of the Interactive Ceiling Murals Plans with Users

Data generated during cultural heritage research and knowledge related to its subject must be brought out to the public domain. Access to heritage should not be just for the privilege of a few it must be universally accessible. The democratization through open-access and open-source interactive virtual or digital platforms can serve the dual purpose of preservation and the dissemination of tangible and intangible heritage of

the murals. As more users access these open platforms, knowledge about the murals gets shared, thereby increasing the information in circulation in the public sphere.

Our users are from a diverse background and their needs vary. Platform designs must account for the diversity of users and their rationale for accessing the platform. Driven by the firm belief in fostering cultural democracy for all, our inter-disciplinary team at IIACD worked with various users to codesign and develop the open source and open access user-driven Hampi interactive ceiling murals plan (HICMP) in 2014. The HICMP features the rangamantapa ceiling murals at the Virupaksha temple (Fig. 2). Users can access the HICMP online[5] [8].

In August 2015, the Lepakshi interactive ceiling murals plan (LICMP) was developed (Fig. 3). These feature the ceiling murals at the natyamantapa of the Veerabhadraswamy temple in Lepakshi. Users can access the LICMP online.[6] We are currently working on the interactive ceiling mural plan of the full temple in Lepakshi. The architectural drawing of the ceiling murals plan of this temple was shared by George Michell in February 2016 (Fig. 4).

Users who are most likely to access these interactive platforms could be artists, art historians, historians, designers, conservators, virtual tourists, students and heritage enthusiasts. The aim of these web-enabled interactive ceiling mural plans is to not only make them available to scholars and conservators, but to also engage students, virtual tourists and wider public with the tangible and intangible heritage of these temples murals in a democratic manner, where the user can experience and appreciate the murals, zoom, pan and browse image details, explore further and learn more about the history, layout, themes and characters, and also access and obtain scholarly annotations and rich textual, audio and video narratives at their own pace.

Through user research we found that not every potential user understands English and hence the application required a translation feature. The Hampi and Lepakshi interactive ceiling plans include a Google translate (major Indian languages) option to make the site accessible and inclusive for non-English users (Fig. 5). We will soon add international language translation options to both plans.

2.3 Open Source Technologies for Interactive Ceiling Mural Plans

Open-source technologies were used for the development of the interactive plans. The Hampi Interactive Ceiling Murals Plan (HICMP) version 1 was implemented using Mouchak [9], a JavaScript open-source framework for building websites quickly. Mouchak uses HTML5 Boilerplate and Bootstrap project as a boilerplate code for the website and provides a visual editing interface to create a website and edit content, primarily for non-technical users. It leverages powerful libraries like Backbone.js

---

[5]http://iiacd.org/South-Indian-Murals/Hampi/Virupaksha-temple/Rangamantapa/Ceiling/.

[6]http://iiacd.org/South-Indian-Murals/Lepakshi/Veerabhadraswamy-temple/Natyamandapa/Ceiling/.

Fig. 2 Homepage, Hampi interactive ceiling murals plan

Fig. 3 Homepage, Lepakshi interactive ceiling murals plan

and Underscore.js to manage and render content. It abstracts the content of the website into a JSON structure and uses Backbone model and views to render them dynamically. Given the high-resolution images, we had captured and were using, with Mouchak abstracting all the content of the site into JSON the website was loading very slowly unless a cache copy is already present in the browser. Also, after testing the HICMP with diverse users, IIACDs user research team found that the Mouchak framework worked well only when the Google Chrome browser was used. The full range of navigation features of HICMP did not work with Mozilla or Internet Explorer, which are more commonly used. Mouchak was not compatible with mobile

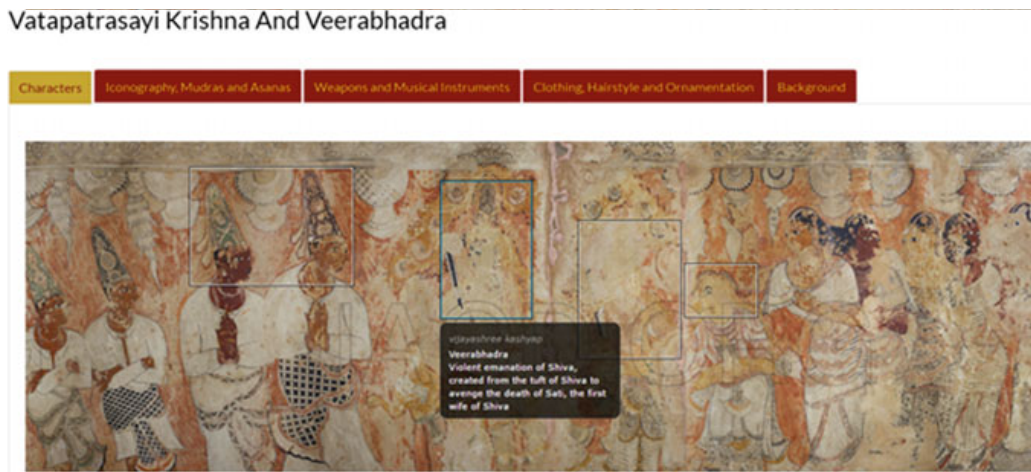**Fig. 4** Homepage, Hampi interactive ceiling murals plan in Kannada

Fig. 5 Different annotation categories

devices or tablets. Users found the number of steps needed to navigate in HICMP to browse image details cumbersome. Given these concerns and the technological obsolescence of Mouchak, which had not been updated for nearly 2 years, IIACD software team decided to search for and apply alternate open source frameworks for building web applications for the Lepakshi Interactive Ceiling Murals Plan (LICMP).

The web application of LICMP was developed using Googles open source Angular JS and Twitters Bootstrap frameworks for creating dynamic web applications. AngularJS has easy routing methods and two-way data binding which other Java scripts lack [7]. Two-way data binding helps in the performance and speed of transferring from the model to the view. AngularJS and Bootstrap permit seamless integration for feature-rich viewing on mobile and other devices. Regular updates for both are available. All features of LICMP work across browsers, operating systems and platforms.

We used OpenSeaDragon, an open-source, web-based viewer for high-resolution zoomable images, implemented in pure JavaScript, for desktop and mobile[7] that uses tiles of high-resolution images and loads the images dynamically, taking the burden off from the browser and making the image viewing experience smoother and faster. User research with art historians, artists and other experts who wanted to see the details of the images and lay users revealed that they favoured the OpenSeadragon image browsing tool to the jQuery plugin used in version 1 of the Hampi interactive ceiling murals plan. Hence, in early 2016 we decided to develop version 2 of the HICMP using AngularJS [10–12], bootstrap and OpenSeaDragron script, which has been recently developed into a library.

2.4 Narratives

Moving beyond the beta version of the LICMP developed for the Lepakshi Utsava in 2012, which permitted mere browsing of the ceiling mural panels, the more recent versions of the LICMP and HICMP have been codesigned and developed with art historians, folklore researchers, artists and designers as well as lay users to provide users a rich, layered experience with information, stories and narratives about the ceiling murals. The paintings at Hampi and Lepakshi have a strong narrative purpose. The religious and social themes depicted here would be hard to interpret for someone who lacks cognizance of Hindu mythology. To enhance the understanding of such visitors/users we incorporated narratives into the interactive ceiling mural plans. Narratives providing scholarly knowledge on the mythological background, themes, composition, iconographical details, aesthetic elements and further reading suggestions were prepared by the art historian and folklore researchers.

Our aim is to not only disseminate mainstream knowledge but to also encourage the transmission of multiple narratives and histories. These interactive plans hence serve as a platform not only for historians, scholars but for temple priests, local guides and local knowledge bearers in the community to add their knowledge of the mural paintings. Other users can also send us their narratives and histories which we will upload. User research revealed that some users preferred auditory narratives over lengthy textual narratives. We have begun supplementing textual narratives with both audio and video narratives for such users. Our future plans including ensuring greater accessibility for the visually challenged [13].

2.5 Annotations

Some lay users found reading lengthy narratives discouraging, thus annotations were chosen as easy pointers for laypersons to understand and engage with the heritage of

---

[7]https://openseadragon.github.io/.

these paintings. User research also showed that users interests ranged from simple character identification to in-depth iconographic elements. To deliver multiple layers of information without overlapping annotations were separated into different categories such as Characters, Iconography, Mudras and Asanas, Weapons and Musical Instruments, Clothing, Hairstyle and Ornamentation, and Background (Fig. 4).

After experimenting with various annotation tools such as szoter, annotorious and annotate we discovered that these tools either lacked several required functionalities or were unaffordable. An alternative solution was developed by our open-source technologists who incorporated the open-source Demon Image Annotation plugin for Wordpress to add textual annotations to images by selecting a region of the image and attaching a textual description to it. They built a User Interface suitable for image annotations which provides better user experience and facilitates enhanced quality of images and image annotations. The Wordpress annotation platform also enables users to comment on the existing ones. Scholarly narratives prepared by our Art Historian and Folklore researchers with rich multi-layered annotations are a distinct feature of our interactive ceiling plans.

2.6   Interactive Plans UI Features

Any user regardless of location can use the HICMP or LICMP urls to view the home page of each plan, which is designed to provide a quick view of the project details such as the team, the temple location, history and layout of the ceiling murals. There is a numbered ceiling plan of the murals with user instructions. Upon clicking any of the numbered boxes in this plan, the user can access the digitally captured image of that mural and use the pan and zoom function to browse the details of the image. A text box in the image page provides the theme, subject, key characters, time period and location of the painting. This information gives the user an idea of the historical background of the mural and enhances his/her understanding.

The icons used for user instructions are universally accepted symbols used in websites. These include zooming in and out, refresh and full screen. Any user with prior experience in browsing websites will find it easy to use these options. He/she can move the image around, zoom in and out and get a close view of the mural. The page also provides links to narratives for the selected panel, annotations and a line drawing. The annotations link takes the user to a page with details about the panel to include characters, iconography, mudras and asanas, weapons and musical instruments, clothing, hairstyle and ornamentation and a brief background. All these features make the interactive experience effective, efficient and satisfying (http://www.w3.org).

We have shown above that the HICMP and the LICMP are both user-driven interactive plans. Usability was a key objective in building these interactive plans. Our design anthropologist, user research and UX (user experience) team have continued to engage with diverse user groups to better understand and meet their unmet needs and desires in improving user experiences and interactivity.

3 Towards Tangible Virtuality

A key challenge for digital heritage scholars and other digital humanities experts is to co-create virtual environments that support immersive interactions that help the user access and retain the information associated with an artefact. Even more impressive would be virtual environments that actually enhance the tangible experience of an artefact in real life.

We have described the interactive virtual environments that represent the ceiling plans of mural paintings of temples in Hampi and Lepakshi. The next logical step would be to take this project forward with a composite blending of elements from virtual reality (VR), augmented reality (AR) and tangible virtuality (TV). VR is an illusion of participation in a synthetic environment rather than external observation of such an environment. VR relies on three-dimensional (3D), stereoscopic, head tracked displays, hand/body tracking and binaural sound. VR is an immersive, multi-sensory experience [14]. As against VR which is a totally synthetic environment, AR is the embedding of synthetic supplements into the real environment. AR integrates synthetic information into real-world environments [15]. TV is an evolution of the both the AR and VR paradigms. TV aims at providing material interaction with the synthetic objects. Synthetic objects can be physically detailed artefacts embedded in the real environment allowing users to touch and feel. TV synthesizes the real environment along with virtual objects [16].

While the more obvious definition of tangible is palpable or material, a more subtle definition would be a quality realized by the mind. We believe that the image qualities achieved in the virtual environments of mural paintings described herein, along with user interaction with focused zooming of details, annotations and narratives would provide users with a tangible experience of these exquisite mural paintings. We have also begun to explore the possibility of a more immersive experience for users of these virtual environments using augmented reality renderings of the murals in the temple environment and immersive personal environments such as Samsung VR Gear. This imagining of a future project possibility follows the lines of the design of the Pure Land, a digital installation which allows people to virtually experience the Mogao caves [17].

4 Future Scope and Conclusion

Future scope for this project has several dimensions which we discuss below.

4.1 Accessibility

The concept of accessibility aims at providing equal access and equal opportunity (http://www.w3.org) to use websites, technologies and tools. People with diverse abilities must also be able to access the web and be able to easily view, navigate and interact with the HICMP and LICMP interactive plans. Accessibility supports social

inclusion and also provides a strong case towards translating these plans into responsive mobile sites, multimodal interaction and search engine optimization (SEO).

While striving to make the interactive plans as accessible for a range of users, we are also working on making them accessible to people with visual challenges and disabilities. This implies implementing the key principles of Web Content Accessibility Guidelines 2.0 [18] to include:

- Inclusive perception—where all content including images, multimedia and video will include a textual description for easy scanning and readability with visual assistance software. The site will include options to change colour contrast and increase the font size to enable people with low vision to easily read and view images.
- Operability—where functionalities will be accessed not only through the mouse but also with the keyboard. In this case, it is necessary to complete robust web markup to enable users to find content. All web content should be marked up thoroughly to minimize HTML and CSS validation errors. This will help assistive technology software to interpret content easily.

4.2 Development of Apps

The use of mobile and tablet apps is slowly gaining momentum in the cultural heritage field. An app can serve multiple uses which include providing mobile content on the heritage site/artefacts/museums/galleries. Such content can either serve as a teaser to the actual experience or take the user through a virtual mobile tour (as will be the case of our interactive ceiling plans when we develop such an app). The content for such apps are multimodal to include text, still images, moving images, videos and interactive maps. Interactive maps maximize the capacity of smartphones to provide location-based information. While we have already included the feature of Google map to provide location information, this feature can be maximized as a content feature in a mobile app. While we debate future possibilities and building and extending features into the project, a key question is sustainability. Sustainability involves the availability of funding which has to go into maintaining digital sustainability in terms upgrading IT and software infrastructure as well as sustain collection of digital cultural content across the participating multiple institutions. It is hoped that such funding will be made available to further strengthen and sustain the project which will impact public democratic access to our shared cultural heritage.

4.3 Further Engagements with Tangible Virtuality

One of the key techniques which should be available in the near future is multi-sensational rendering (MSR) (Horvath, Rusak, van der Vegte and Opiyo, 2008) which will provide a tangible form to digital information so that as human beings our senses

can experience artificial environments and objects as it appears in reality. For this project, this can spell out a variety of options including a total, complete experience of the temples along with the murals; a virtual 3D walkthrough with physical, haptic, tactile and kinaesthetic feedback; 360 rendering of multiple views of the temple and ceiling murals; detailed experiences of surfaces, textures, designs and sculpted surfaces. All these experiences can be further enhanced with AR technologies to add value to the users experience through additional audio-visual narrations and audio-visual annotations.

To conclude, in this chapter we have shown that the interactive ceiling plans provide a rich browsing and informative experience to the user. The project has been a democratic endeavour with our working with user groups to arrive at the best possible UX. It is envisioned that the next level of the project will involve a more immersive, multi-sensory experience for users thereby enhancing their aesthetic enjoyment of cultural heritage artefacts without actually being present at the physical site.

References

1. Stanco F, Battiato S, Gallo G (2011) Digital imaging for cultural heritage preservation: analysis, restoration, and reconstruction of ancient artworks. CRC Press
2. http://asi.nic.in/national_consrv_policy_ancient_monu.asp
3. Kenderdine S (2013) Pure land: inhabiting the Mogao caves at Dunhuang. Curator: Museum J 56(2):199–218. Wiley Online Library
4. Pachner R (1985) Paintings in the temple of Virabhadra at Lepakshi. Vijayanagara–City and Empire: New currents of research, vol 1. In: Dallapiccola AL, Zingel-Ave Lallement S
5. Dallapiccola AL (1997) Ceiling paintings in the Virupaksha temple, Hampi. South Asian Stud 13(1):55–66. Taylor Francis
6. Rao DH (2004) Lepakshi temple. Bharatiya Kala Prakashan, Delhi, p 254
7. Wu J, Yu TX (2013) Virtual walk-through of Dunhuang Panorama-based on web. Adv Mater Res 718:2175–2183. Trans Tech Publ
8. Paramasivan S (1936) The Vijayanagara paintings: late Vijayanagara paintings in the Brihadisvara temple at Tanjore. Vijayanagara Sex Centenary Commemoration Volume, Dharwad
9. https://github.com/janastu/mouchak/blob/master/README.md. Accessed 25 June 2016
10. https://docs.angularjs.org/guide/introduction/. Accessed 25 June 2016
11. https://code.tutsplus.com/tutorials/5-awesome-angularjs-features--net-25651/. Accessed 25 June 2016
12. https://www.w3schools.com/angular/angular_databinding.asp/. Accessed 25 June 2016
13. Chandru U, Vijayashree CS, Rathod D, Vishwanatha V, Pradeepa C, Chandru V, 2015. Interactive ceiling plans of mural paintings in Hampi and Lepakshi. In: CIDOC, (2015) Documenting Diversity Collections. Catalogues Context, New Delhi
14. Earnshaw RA, Gigante MA, Jones (1993) Virtual reality systems. Virtual reality: enabling technologies. Academic Press, London, pp 15–25
15. Bimber O, Raskar R (2005) Spatial augmented reality: merging real and virtual worlds. CRC press
16. Horváth I, Rusák Z, de Smit B, Kooijman A, Opiyo EZ (2008). From virtual reality to tangible virtuality: an inventory of the technological challenges. In: ASME-AFM 2009 world conference on innovative virtual reality. American Society of Mechanical Engineers, pp 45–57
17. Chan LKY, Kenderdine S, Shaw J (2013) Spatial user interface for experiencing Mogao caves. In: Proceedings of the 1st symposium on Spatial user interaction. ACM, pp 21–24

18. Narasimhan N, Sharma M, Kaushal D (2012) Accessibility of Government websites in India: a report. The Centre for Internet and Society
19. Campbell M et al (2015) Hampi Bazaar demolition II: how maps alienate people. Econ Polit Wkly 50(29). Sameeksha Trust
20. https://www.jisc.ac.uk/blog/sustaining-digital-collections-finding-future-funding-27-mar-2014/. Accessed 25 June 2016
21. https://canada.pch.gc.ca/eng/1443540740963/. Accessed 25 June 2016

A Patch-Based Constrained Inpainting for Damaged Mural Images

Pulak Purkait, Mrinmoy Ghorai, Soumitra Samanta and Bhabatosh Chanda

1   Introduction

Social, cultural and political history of a nation is reflected in its various monuments and artefacts of the past. So we should respect and care for the art and artefacts of the past. This motivates Heritage Preservation mission in India [3] and other countries [18, 25]. A mural is a piece of artwork painted or applied directly on wall or ceiling or any other large permanent surface. The significant shares of murals or reliefs in temples in India are religious motifs and paintings. Particularly distinguishing characteristics of such murals are symmetry, repetition and harmony in themselves. Over time, the quality of these murals has been deteriorating rapidly and need restoration. However, many of these paintings and reliefs are not allowed to access physically due to religious and other reasons. Thus, it becomes a necessity to restore them digital space so that one may have an impression of original appearance in future. Digital restoration of archaeological objects is more useful and efficient because of two reasons. First, only a few expert artists restore digital artworks manually with great effort, and it takes too long to cater such a huge treasure of heritage. Second, permanent change of the artwork from its present form may not be allowed.

In this work, we propose an algorithm to restore the digital image of mural paintings with occasional intervention of artist/user. This semi-automatic system is more useful because of large variation in topology of distortion as well as style and content of painting. Here we try to develop a sufficiently fast algorithm so that the user can execute it in almost real time and observe the effects of parameter values and the

P. Purkait
University of Birmingham, Birmingham, UK
e-mail: pulak.isi@gmail.com

M. Ghorai · S. Samanta · B. Chanda (✉)
ECSU ISI Kolkata, Kolkata, India
e-mail: chanda@isical.ac.in

M. Ghorai
e-mail: mgre04@gmail.com

S. Samanta
e-mail: soumitramath39@gmail.com

constraints imposed on the input image under study. This in turn helps to select, in an intuitive way, the near optimal parameter values that produce a coherent visual result. Note that, in our experiment, we have used only subjective quality measure as original mural images are not available for comparison.

In classical image processing, image restoration mainly includes image deblurring and denoising. Recently, image inpainting techniques have attracted attention of researchers due to huge demand for removing undesired objects [7, 13, 31] and filling up damaged regions automatically. These algorithms work well for most of the natural images, but for paintings/mural images containing either smooth brush strokes or uniform region with only a few curves/lines, they may not be able to recover the pattern the way it should be [21]. Second, it is expected that the algorithm should response in almost real time so that it can help building a user friendly interactive tool.

Here we develop a coherent image inpainting algorithm that would be effective for the paintings containing repetitive patterns under constrained environment where users need to mark the damaged or missing (target) region along with the intact (source) region. On the other hand, in unconstrained situation, where target region is small or thin and surrounded by smooth texture, the user needs to mark only target region. To remove spurious noises as well as to highlight the edges, we require an edge enhancing diffusion scheme to synthesize more realistic sharp paintings. For this purpose, here, a patch-based diffusion technique and a novel patch-based high-frequency enhancing method are used alternately. In this chapter, our contributions are multifolds as listed below.

Propose a novel patch similarity measure based on spatial coherence.
Devise a patch-based algorithm for automatic coherent texture synthesis under constrained environment.
Introduce a patch-based edge sharpening diffusion technique for denoising.
Customize the proposed method for interactive mural painting restoration.

The preliminary results of the proposed method are reported in [21]. Rest of this chapter is organized as follows. Section 2 describes some pioneering works in the field of image inpainting and texture synthesis. A constrained image inpainting method is proposed in Sect. 3. Section 4 describes our pyramid-based anisotropic diffusion procedure that preserves edge sharpness. Experimental results are given in Sect. 5. Finally, Sect. 6 presents concluding remarks along with direction of future research.

2 Related Works

For automatic restoration of digital images of paintings, Giakoumis and Pitas [15] have proposed a fully automated two-step system. In the first step, they detect the cracks in the digital paintings by some morphological operators followed by identification of brush strokes that have beed misclassified as cracks. In the second step

the cracks are filled using region growing approach from seed. The second step is dependent on the first one and would produce artefacts in case of false detection. Moreover, the algorithm is limited only to the paintings distorted by the thin cracks and there should not be any blob of missing pixels. Later, they appreciated the need of user interaction and modified their algorithm to make it semi-automatic and interactive [14]. Arora et al. [1] have followed the same strategy for detecting and filling the cracks. They have also applied a colour transformation for restoring the degraded color. The said colour mapping is determined based on a database of original and degraded colour image pairs. Though exhibited results appear reasonably good, this algorithm is useful only for the paintings with thin cracks and similar kind of colour distortion. So we need to develop algorithms that can fill both large blob type and thin crack type damaged regions by synthesizing texture from known region(s).

A pioneering classic texture synthesis method proposed by Efros and Leung [11] uses a simple non-parametric sampling. This is further improved by modifying the search technique as well as the sampling method resulting in better structure preservation [10, 26]. The greedy approach adopted by these algorithms may introduce inconsistencies while completing the large holes with complex structures. Wexler et al. [30] formulated image inpainting as a global optimization problem, and obtain a globally consistent solution for filling-in large missing regions. Though this method usually produces good results, it is relatively slow compared to other methods. A fast and randomized patch search algorithm called PatchMatch [4] is developed to exploit natural coherence in the image using nearest neighbor field (NNF). Later, Darabi et al. [8] propose *image melding* to improve PatchMatch by incorporating geometric and photometric transformations.

Another important step in this work, as mentioned earlier, is denoising while keeping the edge strength unaltered. Note that denoising technique usually reduces high-frequency components in the image which consequently blurs the edges. Second, because of ageing some mural paintings are not only noisy but also look blur due to inter-region colour mixing. So preserving edge strength may not be always su cient, it may be sharpened or enhanced. Anisotropic diffusion [20, 24] was originally designed for removing noise without reducing the edge strength. Perona and Malik [20] probably are the first to address the anisotropic diffusion for edge preserving smoothing. Later in [22, 29] edge preserving smoothing was modified for patch space following the success of the non local means (NLM) denoising methods [6].

3 Proposed Image Inpainting Method

In this section, we discuss patch-based image inpainting which consist of several components such as patch representation, coherent texture synthesis, patch filling and local similarity measure. These methods have their own contributions to improve overall image inpainting in visual appearance.

3.1 Patch Representation

Patch-based image processing techniques are popular nowadays as they are able to capture local information e ciently. The main idea is to consider a window surrounding the pixel of interest and accumulate all the pixels within the window for feature representation. Some statistical property of local pixels are considered to compute the feature vector (p) at pixel *p*. For this, at each pixel position *p* of patch $\mathbf{P}(p)$ some local features are determined by statistical measure of the intensity values. Suppose these features are defined as

$$a_i(p) = \frac{1}{n_i}(\quad p_{i,j}) \quad \text{and} \quad g_i(p) = \frac{1}{n_i}(\quad |p_{i,j} - p_{i,j-1}|) \tag{1}$$

for $i = 1, 2, \ldots, r$, where average intensity a_i and average gradient g_i are computed along the *i*th closed path around the pixel. They together represent textural property at pixel *p* and *r* is the radius of the window. So at each pixel *p* of the patch $\mathbf{P}(p)$, a vector (p) of size $n = c(2r + 1)$ is formed concatenating the average intensity and average gradient, where *c* is the number of colour channels. Thus the vector (p) not only contains the colour information at the pixel but also a summarized information about spatial arrangement of colour over its neighbourhood.

The distance between two patches $\mathbf{P}(p)$ and $\mathbf{P}(q)$ is defined by Minkowski distance of order *m* (*m*-norm distance) as

$$d(\mathbf{P}(p), \mathbf{P}(q)) = (\quad |\ (p) -\ (p)|^m)^{1/m}. \tag{2}$$

The 1-norm distance (i.e., $m = 1$) is the sum of absolute difference (SAD), also called Manhattan distance. Here we have chosen SAD as the distance measure.

3.2 Coherent Texture Synthesis

Based on coherence property of nearby patches Simakov et al. [23] proposed a method to represent the visual data by using bi-directional similarity. They used a global objective function to remove the pitfalls of local inconsistencies. Here we adopt a similar idea to devise a novel method for texture synthesis to fill the missing regions so that the filled region is coherent with the known regions of the image.

Proposed method assumes that the missing or damaged region *H* of the input image *I* is coherent with the remaining part of the image. In other words, the missing region *H* should be filled with estimated data H^* such that the inpainted image *I* has to be visually coherent with the known (undistorted) image region *I* *H*. To achieve this we look for a solution, in the form of patches, that maximizes the following objective function

$$Coherence(I \mid I \quad H) = \max_{\substack{q \in I \ H \\ p \in H^*}} s\left(\mathbf{P}(p), \mathbf{P}(q)\right) \tag{3}$$

where $\mathbf{P}(p)$ denotes the patch extraction operator at pixel location p, and coordinates p and q run over all the points of unknown and known regions respectively. The term $s\left(\mathbf{P}(p), \mathbf{P}(q)\right)$ represents the local similarity measure between patches at p and q. In our experiment, we define the similarity measure as:

$$s\left(\mathbf{P}(p), \mathbf{P}(q)\right) = \exp\left(\frac{-d\left(\mathbf{P}(p), \mathbf{P}(q)\right)}{\sigma^2}\right), \tag{4}$$

where $d\left(\mathbf{P}(p), \mathbf{P}(q)\right)$ is the distance between the patches $\mathbf{P}(p)$ and $\mathbf{P}(q)$ defined in Eq. (2).

3.3 Patch Filling

Since our objective function Eq. (3) is a non-linear one, we solve it by an iterative method using Expectation and Maximization (EM) algorithm, where in each step two functions are carried out: a solution is guessed and the guess is updated. The coherence between the patches in H^* and rest of the image $I \quad H$ as given by Eq. (3) is maximized if for every pixel p of the patch $\mathbf{P}(p) \in H^*$ all its surrounding patches $[\mathbf{P}(p_1), \mathbf{P}(p_2), \ldots, \mathbf{P}(p_k)]$, which have come as corresponding similar patches from the known region of the image $I \quad H$, agree with corresponding feature values at p. Therefore, the iterative E-step tries to satisfy this condition for every patch $\mathbf{P}(p)$ surrounding pixel $p \in H^*$, and the M-step searches for the best similar patches from the known region $I \quad H$ of the image I. Let each of $\mathbf{P}(q_1), \mathbf{P}(q_2), \ldots, \mathbf{P}(q_k)$ denotes a patch of $I \quad H$ that is most similar to $\mathbf{P}(p_1), \mathbf{P}(p_2), \ldots, \mathbf{P}(p_k)$, respectively. Hence, at each iteration, for each point $p \in H^*$ and each corresponding surrounding patch $\mathbf{P}(p_i)$, we need to select best possible patch $\mathbf{P}(q_i)$ in $I \quad H$. Then (r, g, b) pixel value at p is replaced by the weighted average of the colour values of the pixels of the surrounding similar patches $\mathbf{P}(q_i)$. That means

$$I(p) = \sum_{p \in P(p_i)} w_i I_{P(q_i)}(q) / \sum w_i \tag{5}$$

The weight value is taken as the similarity value s_i between the corresponding patches at p_i and q_i. This method still requires significant amount of computation. To reduce the required computation for finding out the nearest neighbour we compromise to approximate nearest neighbourhood (ANN) instead [4].

To speed up further and to preserve global consistency, we perform the iterative process in multiple scales. At each scale the resolution of image is a fraction of the resolution of the upper scale. Through our experiment we have seen that usually a scaling factor of 1.25–2.00 produces good result. However in rest of our experiment,

downscaling resolution factor is taken to be 1.5. As mentioned optimization is achieved by using EM technique starting at coarsest scale and the solution obtained at each scale is propagated to upper scale for further refinement. Let the unknown region H is initialized by some random values at the coarsest scale followed by a few EM iterations. Resultant image is then resized by the same resolution factor until the original dimension is obtained. At any resolution level the current solution is obtained by filling the missing region by the best matching patches. A formalism similar to Eq. (3) was already used in [30] for summarizing the visual data.

Further, the performance of the proposed method is improved by imposing geographical constraint on the search space. For this, user can mark a relevant part of the input image as the source region for searching similar patches. Here the underlying assumption is that the synthesized texture would be coherent with the marked source region only. By this interactive approach, the user can exercise more control over the synthesizing process, and as a result the performance of texture synthesis improves significantly. Note that, if the synthesized texture is not of desired type, then the source region is altered. In Fig. 1, we have shown some examples of constrained texture synthesis. Compared to this, the earlier method may be referred to as unconstrained texture synthesis.

In case of unconstrained inpainting, the whole image is considered as the source region. Though constrained image inpainting method produces better result compared to the unconstrained one (see Fig. 2), we adopt the latter method when we are not sure about the coherent source region and let the algorithm to search the whole known region of the image for coherent patch.

3.4 *Local Similarity Measure*

Nearest neighbour search is an important method that finds variety of applications in number of domains including pattern recognition and classification, knowledge discovery and data mining. In the context of present problem the objective is to develop an algorithm that can generate an approximate nearest neighbourhood field for each of the patches $\mathbf{P}(p)$ of an image based on the similarity measure $s(\mathbf{P}(p), \mathbf{P}(q))$ as expressed in Eq. (4). For n points in a d-dimensional space the said problem may be solved through simple brute-force search in $O(dn)$ time. With our goal to build an interactive image completion system we need to develop a fast algorithm even at the cost of accepting a near optimum result or, in other words, approximate nearest neighbour (ANN). There exist a number of e cient ANN search algorithms [2, 17] that can do the task in $O(d \log n)$ or less time. Here we employ a specific ANN search algorithm called "PatchMatch" [4] [1] which works well in the space of patches. One of the main reasons for adopting this algorithm is that it is developed for interactive structural image editing. The PatchMatch algorithm [4] has three main

---

[1]Source code can be downloaded from http://gfx.cs.princeton.edu/pubs/Barnes_2009_PAR/index. php.