Computer Aided Reconstruction of Complex Sites and Architectures. Application to the Grotta dei Cervi and the Broken Frescoes of the Assisi Basilica

Paquet, Eric; Beraldin, Jean-Angelo; Viktor, H.L.; Benedetti, B.

Publisher's version / Version de l'éditeur:

The XXI Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS 2008) [Proceedings], 2008

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

*Computer Aided Reconstruction of Complex Sites and Architectures. Application to the Grotta dei Cervi and the Broken Frescoes of the Assisi Basilica **

Paquet, E., Beraldin, J.-A., Viktor, H.L., Benedetti, B.
July 2008

* published at The XXI Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS 2008). Beijing, China. July 3-11, 2008. NRC 50356.

COMPUTER AIDED RECONSTRUCTION OF COMPLEX SITES AND ARCHITECTURES: APPLICATION TO THE GROTTA DEI CERVI AND THE BROKEN FRESCOES OF THE ASSISI BASILICA

E. Paquet[a], J.-A. Beraldin[a], H. L. Viktor[b] and B. Benedetti[c]

[a] IIT-VIT, National Research Council, Ottawa, K1A 0R6, Canada - (eric.paquet, jean-angelo.beraldin)@nrc-cnrc.gc.ca
[b] SITE, University of Ottawa, Ottawa K1N 6N5, Canada – hlviktor@site.uottawa.ca
[c] Scuola Normale Superiore di Pisa, Pisa, Italy – b.benedetti@sns.it

Commission VI, WG V/2

KEYWORDS: Cave, Cultural Heritage, Descriptor, Fresco, Indexing, Invariant, Modelling, Reconstruction, Restoration

ABSTRACT:

This paper describes a computer-aided method for the reconstruction and modelling of complex sites based on content-based indexing of images. A pose, distortion and composition robust descriptor is automatically generated for each image associated with a given site. These descriptors are subsequently placed in a database, to be used for future querying. The database is interrogated in real-time using an intuitive query by example paradigm, providing visual results which may be easily understood and interpreted. Our approach is applied to the modelling of the Grotta dei Cervi and to the restoration of the broken frescoes of the Assisi Basilica.

1. INTRODUCTION

Image-based 3D modelling of complex sites and architectures and laser scanning of large and complex objects involves the acquisition of numerous digital pictures and range images; each picture or range image being associated with a certain partial view [1, 2]. These views must subsequently be analysed, registered and merged together, in order to create a coherent and accurate model of the site under consideration.

Over the past decade, the number and complexity of such views have been increasing tremendously, both for image-based and laser scanning techniques. An analogous remark may be made about the projects themselves. Most of these projects, involve numerous teams and a multiplicity of data sources and acquisition techniques. Also, the vast majority of projects involve multinational collaborations which mean that participating teams may differ in both their language (e.g. English, Italian, and Chinese) and the methodology by which they annotate the views, which may lead to data inconsistencies. In addition, if the acquisition process is not carefully recorded, data integration problems such as sequential but meaningless naming of files, spelling errors, synonyms and repetitive use of ambiguous file names, to name but a few, may render the exchange of information and models difficult, if not impossible. Furthermore, complex projects may require the use of legacy data which might employ different, outdated methodologies and annotations. Typically complex projects span a long period of time. This implies that during that period, teams may be added or removed, with a potential impact on the homogeneity and the coherence of the methodology. It thus becomes difficult to efficiently model complex sites and architectures with either photogrammetric techniques or laser scanners. Consequently, a substantial amount of time is devoted when aiming to understand the various methodologies and to retrieve similar or related views.

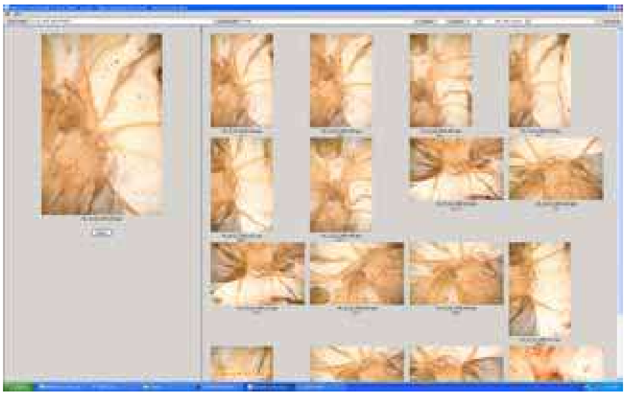

Figure 1. Retrieval of various views of a geological detail of the Grotto dei Cervi; a Neolithic cave near Lecce (Italy)

In order to address these problems, we propose a new approach based on content-based indexing and retrieval of digital pictures. In our method, digital pictures are analysed automatically, in order to generate a compact and abstract index that describe their 2-D content [3, 4]. By image, we mean pictures acquired with a digital camera or multi-spectral imaging device, or colour intensity images acquired with a laser scanner. The ensemble of all these indexes constitutes a database that may be searched with a dedicated search engine. The search engine is based on the query by example paradigm: the user chooses an image and the closest images, from the composition point of view, is retrieved by the search engine. In this way, the various views created by different teams may rapidly and efficiently be registered and merged, irrespectively of their origin or of their underlying methodology.

The paper is organised as follow. Firstly, a description of our approach as well as of the related algorithms is provided. Next,

Figure 2. Retrieval of various views of a fragment of "St Francis Giving is Mantle to a Poor Knight" from a database of 10.000 images simulating the virtual reconstruction of a broken fresco of the Assisi Basilica

we apply our method to two real life problems of acquisition, modelling and reconstruction and present our conclusions. In the first case, we address the search and retrieval of similar and related digital images from a database of 10.000 views of two frescoes painted by Giotto and Cimabue in the Assisi Basilica. This case acts as a benchmark for the reconstruction of the frescoes that were broken into a huge number of fragments in the 1997 earthquake. In the second case, we apply our method to the modelling of the "Grotta dei Cervi", a Neolithic cave that may be found near Porto Badisco 49 km southeast from Lecce (Italy). Two typical queries and the associated retrieved results are shown in Figures 1 and 2, for the cave and for the painting, respectively. In the first figure, various similar and related views of a geological detail of the cave are retrieved based on their appearance while, in the second figure, images similar or related to a low-resolution sub-image (i.e. a fragment of the ring) of the Giotto's painting are retrieved from a database of 10.000 images. In the next section, we provide an outline of our method.

2. DESCRIPTION OF THE INDEXING AND SEARCHING APPROACH

Our approach does not involve any assumption about the underlying annotation and/or the methodology. Furthermore, our method is transparent, in that it does not require one to modify annotations or methodologies. Our objective is to assist the user in the reconstruction or modelling process. We aim to achieve this goal by providing a limited number of visual options which may easily be analysed and from which the best reconstruction or modelling option may be inferred.

The two case study presented in this paper are characterised by a large number of pictures. In the case of the cave, each picture corresponds to a digital picture of a section of the cave; most of which were acquired with a digital camera that was registered with a range scanner [5]. For this reason, they may be utilised to find correlated views of 3-D scans (range images) as well. In the case of the frescoes, each picture corresponds to a simulated broken fragment.

Our approach may be described as follow. The content of each image is described according to the algorithm presented in section 3. This algorithm provides a compact (less than 300 bytes) representation of the composition of the images. It is robust against sampling, orientation (or pose), distortion and composition. The descriptors are then placed in a database that may be searched using the query by example paradigm. This example may be an image from the database or an external image provided by the user, which may then described, in real-time, by the system. For a given query, the descriptors or indexes are compared with a Euclidian metric and the most similar results are displayed visually to the user. It is possible to search a database of 100.000 images in less than a second on a standard workstation. The results provide, for instance, a set of correlated views which represent adjacent or overlapping view of the cave. In the next section, we provide a description of the indexing algorithm.

3. ALGORITHM FOR IMAGE INDEXING

Our approach does not involve any segmentation or image understanding "per se". It is a purely statistical approach which was designed in order to be highly robust against pose, distortion and change in relative disposition of pictorial elements. Also, it must be suitable to process a very large amount of data in a short period of time. Furthermore, the generated descriptors are compact in order to be searchable in real-time. We now described our algorithm [3]. A window or structuring element is randomly displaced on the image. The position of the centre of the structuring element is associated to a Sobel sequence which has the double advantage of being both pseudo-random and multiresolution. The randomness is important in order not to introduce any a priori hypothesis in the analysis of the image. The multiresolution requirement is related to the level of details suitable for the concrete application. For each position of the structuring element, two bidimensional histograms are calculated. The first dimension is related to the hue or the saturation associated with the pixels belonging to the structuring element, while the second dimension is related to the relative proportion of the hue or saturation on a quantised scale. The descriptor is obtained by accumulating the obtained results for each position of the structuring elements. The size of the later is less than 300 bytes irrespectively of the size of the original image. As shown by our experimental results, this algorithm provides a robust representation against pose, distortion and relative position of pictorial elements.

4. APPLICATION TO THE RECONSTRUCTION OF FRESCOES FROM THE ASSISI BASILICA: SIMULATION

In 1997, two earthquakes severely damaged the Assisi Basilica. Many irreplaceable frescoes painted by Giotto and Cimabue were broken into thousands of fragments (see Figure 12 in Appendix). Some of them were reconstructed by hand, but many are still waiting to be reconstructed. The problem is twofold: the number of fragments is unmanageable and their extensive manipulation, for the reconstruction, adds to the damage already caused by the earthquakes.

Figure 3. Same fresco as Figure 2. Identification of a detail of the eye of the donkey

In order to address this problem, the following approach is proposed. A digital picture of each fragment is acquired on a neutral background, taking the following observation into consideration. Since the support of the fresco, that is the plaster, has been badly damaged, the thickness of the fragment is likely to be uneven. This implies that the fragments are likely to make an angle with the optical axis of the digital camera when a picture is acquired; meaning that some kind of distortion, relative to the reference image, is to be expected.

Once the pictures of the fragments have been acquired, each picture is processed in order to eliminate the border and the background: the background because it does not carry any information from a reconstruction point of view and the border because it might carry false information. Indeed, a broken fresco is not like a puzzle, in which the borders of one piece match perfectly the borders of another piece. Rather, the border of a fragment is usually damaged and consequently does not correspond to the border of an adjacent fragment anymore. In such conditions, a border-based approach has very little chance of success. The images resulting from this processing are indexed according to the method presented in sections 2 and 3. The next step is, naturally, to find the position of the different fragments with respect to a global view of the fresco.

Figure 4. Same fresco as Figure 2. Identification of a detail of the knight's lower leg

In order to locate the fragments, we start from the global view of the fresco. The later does not necessarily need to have a high resolution nor does it need to match the resolution of the fragments, although a high resolution facilitates the reconstruction process. From the global picture, a very large set of synthetic fragments is generated. The synthetic fragments are generated by acquiring virtual photographs of very small sections of the fresco at various angles, both in altitude and azimuth, with a virtual camera in order to simulate the unevenness of the plaster. The synthetic fragments, as well as their position relative to the fresco, are stored in a database. Each synthetic fragment is then indexed following the approach presented in sections 2 and 3.

It should be noted that no assumptions are made about the synthetic fragments. Rather, they are generated pseudo-randomly both in terms of size and orientation. This implies that, in general, the number of synthetic fragments must be much higher than the number of real fragments unless some hypothesis may be made about their nature e.g. about their size. Furthermore, as the results will show, it is not necessary to have a synthetic fragment that matches perfectly the corresponding real one, in order to find the position of the later. This follows from the robustness of the algorithm against pose and relative position of pictorial elements.

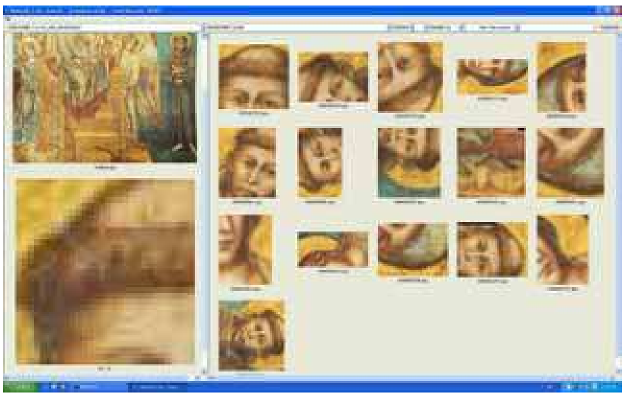

Figure 5. Same fresco as Figure 2. Identification of a detail of the hair (crown)

In order to find the position of a fragment, the user interrogates the synthetic fragment database, as described in section 2. The picture of the real fragment of interest, as well as its descriptor, is input and the most similar synthetic fragments are displayed with their location. Then, the user validates the results and the position of the real fragment may be inferred from the position of the synthetic one. Although time consuming, the validation of the results by an expert is a "sine qua non" condition for any real restoration project to ensure the validity of the results.

Figure 6. The Virgin in Majesty by Cimabue. Identification of a detail of a feather of one of the angels

Figure 7. Same fresco as Figure 6. Identification of St Francis' stigmata

Figure 8. Same fresco as Figure 6. Identification of a detail of the upper mosaic

For the simulation, two frescoes were selected: one by Giotto (1267-1337) which represents St Francis giving his mantle to a poor Knight, and one by Cimabue (1240-1302) which represents the Virgin in Majesty. This choice is based on the observation that Giotto and Cimabue are recognized as the two foremost artists involved in the painting of the Basilica. In order to make the task more challenging, only low resolution pictures were used; which models the worst possible restoration scenario. In the present simulation, we are using a relatively small number of synthetic fragments, ten thousand (10.000) which naturally increase the difficulty of the task. Some of our results are shown in Figures 2 to 8. For each figure, the closest twelve (12) synthetic fragments are displayed. The "real" fragment appears at the bottom-left and the entire fresco at the bottom right. The reader is invited to zoom into the figures in order to see the details. Figures 2 to 4 refer to the fresco by Giotto, while Figures 5 to 8 refer to the fresco by Cimabue. For all the examples, the most similar fragments are easily retrieved in less than a second and the position of the real fragment may be inferred from the position of the synthetic one.

We now describe some results. Figure 2 shows the identification of a fragment of St Francis' ring. Figure 3 shows the retrieval of a fragment of the donkey's eye. Figure 4 shows the identification of a detail of the knight's lower leg. Figure 5 shows the identification of a fragment of St Francis' hair (crown) while Figure 6 shows the retrieval of a fragment of one of the angel's feathers. In figure 7, the identification of St Francis' stigmata is depicted while in Figure 8, a detail of the upper mosaics is retrieved. A point of interest is that not only the position of the fragment may be found but also information related to its surrounding. All the synthetic fragments were generated at very low resolution which again demonstrates the efficiency of the approach. In addition, it is shown that similar fragments may be retrieved, despite of the presence of distortion and independently of the pose.

5. APPLICATION TO THE MODELLING OF THE GROTTA DEI CERVI

The Grotta dei Cervi, which may be literally translated as Cave of the Deer, is a Neolithic cave that may be found near Porto Badisco 49 km southeast of Lecce (Italy). This cave was recently digitized by a team from the National Research Council Canada and the University of Lecce [5].

Once more, we apply our approach, which was introduced in sections 2 and 3, against a database consisting of more than 5.000 digital pictures of the cave. In contrast to the results presented in the previous section, the objective here is to find correlated views in order to assist the scientist in the modelling process. In this particular case, the digital camera was registered with a laser scanner, which implies that correlated pictures correspond to correlated range images (3-D views) as well.

Not only may correlated views be found, but we are also able to find related views. This is important from a scientific point of view. For instance, one would like to find similar pictograms in the cave or similar rock formations in order to perform a comparative study. A sample of our results is shown in Figure 1 and in Figures 9 to 12. The correlated views may also be used in order to facilitate the navigation through a virtual model of the cave.

Figure 9. Retrieval of various views of a group of pictograms; first 11 results from left to right and top to bottom. The retrieval task was successful despite the interference of foreign bodies like the calibration plates and a scientist

We describe some results. The reader is invited to zoom into the figures in order to see more details. Figure 1 shows various views of a rock formation within the cave. Figure 9 shows the retrieval of various views associated with a group of pictograms. The retrieval task was successful despite of the presence of foreign bodies like the colour checking plate and one of the scientist (and co-author!). Figure 10 shows the retrieval of various views a particular pictogram. As shown by the results, partial views are retrieved as well.

Figure 10. Retrieval of various views of a pictogram; first 6 results from left to right and top to bottom

Figure 11. Retrieval of related images: a man with a bow and an arrow: results 1, 2 and 12 from left to right and top to bottom

Figure 11 is an illustration of the second application described at the beginning of this section, i.e. the retrieval of related items. The reference image shows a man with a bow and an arrow. Retrieved images 2 and 12 (left to right, top to bottom) show two very similar pictograms also representing a man with a bow and an arrow in a different attitude. For the images displayed, the orientation and the surrounding are not the same which make the searching task more challenging.

Figure 12. Retrieval of 5 views of a pictogram (results 1-5) and two views of a related pictogram (results 9 and 10)

Figure 12 shows a combination of both types of retrieval. The first five results show various views a pictogram. In this figure, results 9 and 11 show two views of a similar pictogram. Again, the pictograms are retrieved despite of their variations in terms of pose and lighting.

6. CONCLUSIONS

We have presented a computer aided approach for the reconstruction of complex sites and architecture that is also suitable for the reconstruction of broken frescoes. This approach is based on pose, distortion, composition robust automatic indexing of images. It has been applied with success for the modelling and study of the Grotta dei Cervi and for a simulation for a restoration of frescoes of the Assisi Basilica. For the Assisi Basilica restoration project, further work will involve an attempt to recover some of the frescoes that have been destroyed in the 1997 earthquake, with the hope that it may lead to their restoration.

7. ACKNOWLEDGEMENTS

We would like to thank V. Valzano and A. Bandiera from the University of Salento who are leading the overall Grotta Dei Cervi project. The 3D recording of the Grotta dei Cervi was realized within I18 of the "Piano Coordinato delle Università di Catania e Lecce" co-financed by the European Union (FESR, PON Ricerca 2000-2006). Financial support was also made by the town Council of Otranto (LE), the CEDAD and the Department of Beni Culturali of the University of Lecce, the Museo Provinciale of Lecce and the CASPUR of Rome.

8. RERERENCES

1. Blais, F., 2004. A review of 20 years of range sensor development, *Journal of Elect. Imaging*, 13 (1), pp. 231-243.

2. El-Hakim, S.F. and Beraldin, J.-A., 2007. Sensor Integration and Visualisation. Chap. 10 of *Applications of 3D Measurement from Images*. John Fryer, Harvey Mitchell, Jim Chandler (Editors). Whittles Publishing and CRC.

3. Paquet, E., Lahanier, C., Pitzalis, D., Peters, S., Aitken, G. and Viktor, H. L., 2006. Content-based Indexing and Retrieval of Cultural Heritage Data: an Integrated Approach to Documentation with Application to the EROS Database, *The 7th International Symposium on Virtual Reality, Archeology and Cultural Heritage – VAST 2006*, 30 October - 4 November, Nicosia, Cyprus, pp. 270-275.

4. Paquet, E. and Viktor, H. L., 2005. "On the application of 3-D technologies to the framework of cultural heritage", *Annales des télécommunications*, 60, pp. 1311-1325.

5. Beraldin, J.-A., Blais, F., Cournoyer, L., Picard, M., and Gamache, D., 2006. Multi-Resolution Digital Imaging System Applying to the Recording of Grotto Sites: the Case of the Grotta dei Cervi. *The 7th International Symposium on Virtual Reality, Archeology and Cultural Heritage – VAST 2006*, 30 October - 4 November, Nicosia, Cyprus, 8 pp.

6. BBC, "1997: Earthquakes rock central Italy", London, UK. http://news.bbc.co.uk/onthisday/hi/dates/stories/september/26/newsid_2538000/2538651.stm (accessed 10 April 2008)

9. APPENDIX: NOTE ON THE EARTHQUAKE AND THE PROPOSED APPROACH

Figure 12. Collapse of the ceiling of the Assisi Basilica during the earthquake of 27 September 1997 [6]

The extend of the damage caused by the 1997 earthquake may be understood from this picture of the collapsing ceiling. It follows that the borders of the fragments were badly damaged and consequently do not match, due to the intensity of the impact to which they were submitted during their fall. Furthermore, many fragments have been totally destroyed which means that a border-based approach has little chance of success, since one has to deal with empty zones. For this reason, our approach focuses on the inner regions of the fragments and tries to put them in correspondence with a global view of the original fresco.