Journal of Cultural Heritage xxx (2017) xxx–xxx

Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

Review

Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach

Irene Aicardi[a], Filiberto Chiabrando[b], Andrea Lingua[a,*], Francesca Noardo[a]

[a] Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, c.so Duca degli Abruzzi, 24 10129 Torino, Italy
[b] Department of Architecture and Design (DAD), Politecnico di Torino, v.le Mattioli, 39 10125 Torino, Italy

INFO ARTICLE

Historique de l'article :
Reçu le 6 décembre 2016
Accepté le 14 novembre 2017
Disponible sur Internet le xxx

Keywords :
Photogrammetry
Computer vision
Structure-from-motion
Algorithms
3D modelling
Orthophoto

ABSTRACT

The techniques of measuring and 3D modelling based on images, as is typical in photogrammetry, grew in interest again in recent years, since a new generation of software tools has spread. These ones implement in different measure the algorithms developed by computer vision, increasing the automation of the standard photogrammetric process. This made the use of image-based approaches for 3D models reconstruction enormously increase, which is an essential part of the Cultural Heritage documentation and analysis processes. Starting from these assumptions, the aim of the paper is to evaluate what and where it is possible nowadays to find the main differences between photogrammetry and computer vision approaches and how these have to be considered in the choice of the processing technique. The analysis has been performed starting from a theoretical point of view in order to trace the main characteristics of the two methods. Moreover, in order to complete the investigation, an experimental part is reported on two particular cases study, considered as representative of two types of usually surveyed objects. The results allow to enlighten some differences between the two image processing approaches, in terms of accuracy and achieved products.

1. Introduction

The great back in vogue,in the last years, of the use of images for measuring 3D features, made many software have a wide success in many application fields, including cultural heritage documentation and analysis. They offer the possibility of a cheap acquisition, fast and automatic processing and high accuracies in the results. After the skepticism and the enthusiasm are past, the necessity arise for photogrammetrists and geomatics users to know more about the behaviour of the recently spread software tools. These implement techniques and algorithms mainly developed by the computer vision (CV) community, a different but more and more familiar discipline. For researchers or professional actors working with geomatics, it is a strategic topic to know when and how to prefer a photogrammetry-based software or a computer vision-derived technique.

In all the branches of cultural heritage field (mobile museum objects, architecture, archaeological sites, and so on), the 3D survey is an essential support for a number of activities: the object documentation, different kinds of analysis (statistical analysis, historical reconstructions, etc.), the communication and promotion of the sites, and so on. The possibility to generate very accurate and detailed 3D models from imagery is a great opportunity, since a number of tools exist for the method to be employed by a wide community of users, with limited costs. However, the results need to be controlled and critically considered to be reliable. Moreover, it is necessary to contemplate different kinds of processing tools or acquisition procedures based on the kind of object to be surveyed (evaluating the geometry, the texture, the distance, the complexity of its shape, and so on). It is therefore necessary to be conscious of the characteristics of the main two methods for measuring objects and generating 3D models from images (photogrammetry and computer vision derived techniques), in order to choose, each time, the most suitable one.

If we analyze the definition of Photogrammetry (PH) and Computer Vision we can find that:

- "Photogrammetry is the art and science of determining the position and shape of objects from photographs" [1];
- "Computer Vision is a mathematical technique for recovering the three-dimensional shape and appearance of objects in imagery" [2].

* Corresponding author. Politecnico di Torino, c.so Duca degli Abruzzi, 24 10129 Torino, Italy.
Adresse e-mail : andrea.lingua@polito.it (A. Lingua).

https://doi.org/10.1016/j.culher.2017.11.006

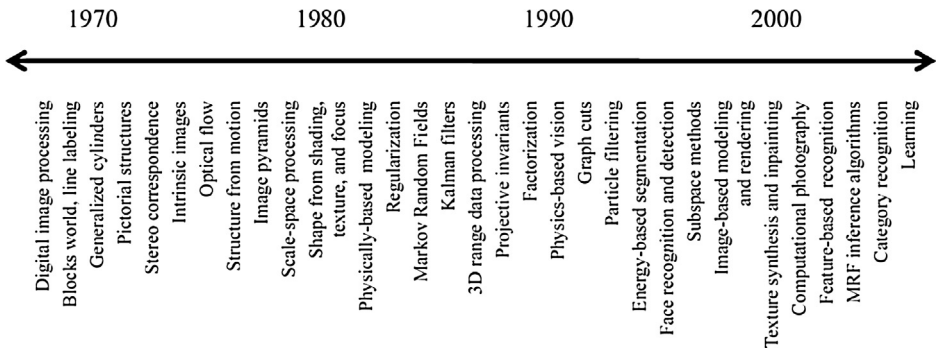

Fig. 1. Timeline of some of the most significant topic in the field of Computer Vision [2].

The main feature is that both the techniques start from the analysis of 2D images to discover 3D shape information, even if the employed approach is sometimes different: originally, the goal of photogrammetry was the measurement of the position of a set of 3D points, while the computer vision aimed at the final appearance of the model.

The main goal of PH was for mapping purposes; consequently, the technique was linked to the achievement of the best reachable metric accuracy. Instead, the CV idea was related to the automation of the process, in particular regarding the possibility to make the data understandable to computers for automatically extracting information from images (image segmentation, individuation of lines with similar value or depth, individuation of the steepest gradient, pattern or face recognition, semantic network and so on) [3].

The birth of PH could be traced back to the 1480, starting from the studies of Leonardo da Vinci [4]. It was then improved through the years, using as base the central projection concept and developing new techniques to reconstruct the 3D model (interior and external parameters problem). The PH had a revival in the 1990s thanks to the digital images [5].

On the other hand, the CV started from the 1970s with the aim to endow robots with the human visual perception, in order to improve their intelligent behaviour. The history of this science can be easily summarized in the Fig. 1 below [2].

By the time, in the field of 3D model reconstruction, PH and CV came to have the same goal and purpose, approaching the problem from two different points of view.

First of all, they start from the same mathematical model (the central projection), but the first CV algorithms use a linear approach to solve the problem, while the PH theory generally considers a non-linear solution that must be linearized according to the approximation of the initial parameters. For this reason, in many cases, for the photogrammetric processing is preferable the interior orientation parameters of the used camera to be known, or, at least, stable, while the CV approach starts from the concept of not calibrated cameras (unknown interior orientation parameters) to reconstruct the 3D shape of the area of interest.

As it is mentioned before, the processing steps for the two techniques could be considered different. Simplifying the basic workflow of the two techniques, the PH starts from the knowledge of the camera parameters (focal length and principal point) and then moves to the extraction of homologous points. Instead the CV starts from the analysis of the images to discover common points (homologous points) between the images and then it reconstructs the geometry of the acquisition.

However, even if sometimes they seem to be two very different approaches, nowadays it is difficult and sometimes impossible to define the border between the two applications in the definition of the used algorithms and the obtained results. The frontier is even fainter when considering that some CV techniques are currently implemented in up-to-date photogrammetric software. For example, feature-extraction algorithms, dense reconstruction techniques and so on. It is thus important to investigate the tools that regulate these processes, in order to reach a higher consciousness of the methods used by the software for suitably choosing one of them for the specific application needs.

The purpose of this paper is to analyze and compare the digital photogrammetry technique with the computer vision method in the field of 3D metric reconstruction from images [6–8]. As a first step, the mathematical structure must be understood as a basis for the following phases. Moreover, the goal of each approach is defined and the procedures characterizing the techniques are described. In order to show the differences of the achieved results, finally, an experimental part reports two representative case studies in the Cultural Heritage field.

2. A brief foreword about the algorithms

In PH and CV, images (stereo pairs or sequential images) are used in order to extract further information. The aim of photogrammetry can be pursued by means of well-established algorithms to extract the geometry and the physical model information. In computer vision the great advantage is the implementation of a human-like capability to recognize 3D information in the image data in not-human entities (computers). However, a main difference rests.

The photogrammetric concept starts from a physical model of the image creation	Instead, the CV has a strong connection with the mathematics and the computer science application and it somehow loses its connection with the physical models

Especially in CV, the 3D reconstruction from the images starts from the image matching techniques and the recognition of interest points on 2D images. Digital image matching can be defined as an establishment of automatic correspondences among the primitives extracted from two or more digital images depicting, at least partly, the same scene (Heipke, 1996).

Over the years, sophisticated algorithms have been implemented by researchers, but a fully automatic and always effective image matching method can still be improved. This is due to the information lost during the image acquisition. Indeed, for a given point in one image, its correspondences on other images may not exist due to occlusion. Moreover, it is possible to have more than one match, due, for example, to repetitive texture patterns or even no matches, according to image noise or lack of textures; semi-transparent object surfaces could also give similar problems. For these reasons image matching techniques can be ranked in the ill-posed problem class (there is no guarantee that a solution exist and that it is unique and stable respect to small variations of the input data, Terzopoulos, 1986).

Pour citer cet article : I. Aicardi, et al., Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach, Journal of Cultural Heritage (2017), https://doi.org/10.1016/j.culher.2017.11.006