Forensic Science International 271 (2017) 59–67

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

A new approach to obtain metric data from video surveillance: Preliminary evaluation of a low-cost stereo-photogrammetric system

Paolo Russo[a], Emanuela Gualdi-Russo[b,*], Alberto Pellegrinelli[a], Juri Balboni[a], Alessio Furini[a]

[a] Department of Engineering, University of Ferrara, Ferrara, Italy
[b] Department of Biomedical Sciences and Surgical Specialties, University of Ferrara, Corso Ercole I d'Este, 32, 44121 Ferrara, Italy

ARTICLE INFO

Article history:
Received 23 June 2016
Received in revised form 2 November 2016
Accepted 16 December 2016
Available online 23 December 2016

Keywords:
Forensic science
Video surveillance
Photogrammetry
Identification
Anthropometry
Height

ABSTRACT

Using an interdisciplinary approach the authors demonstrate the possibility to obtain reliable anthropometric data of a subject by means of a new video surveillance system.

In general the use of current video surveillance systems provides law enforcement with useful data to solve many crimes. Unfortunately the quality of the images and the way in which they are taken often makes it very difficult to judge the compatibility between suspect and perpetrator.

In this paper, the authors present the results obtained with a low-cost photogrammetric video surveillance system based on a pair of common surveillance cameras synchronized with each other.

The innovative aspect of the system is that it allows estimation with considerable accuracy not only of body height (error 0.1–3.1 cm, SD 1.8–4.5 cm) but also of other anthropometric characters of the subject, consequently with better determination of the biological profile and greatly increased effectiveness of the judgment of compatibility.

1. Introduction

Video surveillance systems have become very diffuse in recent years and are now used to monitor a large number of urban areas and public sites. In many cases the images taken by such systems are the major, if not only, source of evidence for the identification of perpetrators. The best solution for this purpose is facial recognition, provided the images are sharp enough. If the image quality is insufficient or the people are masked, other elements must be analyzed, such as stature, gait, behavior, clothing, etc.

In addition to attempted facial recognition, the estimation of stature is often required to assess the compatibility between the suspect and the perpetrator. The most efficient techniques to estimate stature are based either on perspective, knowing the size of objects in the scene [1], or on “inverse photogrammetry”. In this case a 3D virtual model of the crime scene is realized in which a virtual camera is placed and oriented just as the real one. The virtual images provided by the virtual camera are superimposed to the original ones. The height measurements are then taken in the 3D model by means of cylinders [2] or virtual humans [3,4] or by means of a 3D ruler [5]. Another method [6] makes use of a frame placed in the crime at the perpetrator's location in the video in order to calibrate the camera, although the measurement accuracy is too dependent on the correct position of the frame in the scene.

Many factors can reduce the accuracy of stature estimation: location, orientation, optical and electronic quality of the cameras, attitude, posture and camouflage of the subject in the context of the crime scene, operator skill, etc. [2,7–9].

For example, if the perpetrator is walking, his/her height varies approximately 6 cm in the vertical direction, and the maximum size may not correspond to the static value of stature [1]; indeed the search for confidence intervals for this parameter are based on a relationship between a systematic component and a random one [9]. Moreover a significant limitation to the accuracy of the stature estimation is due to the fact that the above-mentioned methods are based on processing of images taken by a single camera. This is because the normal installation of video surveillance cameras aims to limit as much as possible the overlapping of images. In this way the number of cameras is reduced to a minimum, with the consequent decrease of the costs of installation, maintenance and management of the plant. Therefore only rarely is the subject recorded by multiple cameras. Unfortunately this is a drawback because it prevents application of the algorithms of analytical photogrammetry and thus an accurate 3D determination of

* Corresponding author.
E-mail addresses: gldmnl@unife.it, emanuela.gualdi@unife.it (E. Gualdi-Russo).

http://dx.doi.org/10.1016/j.forsciint.2016.12.023

stature. Interesting results have been obtained in forensic studies with the simultaneous use of multiple cameras, but at least three cameras are required to achieve good accuracy [10]. Other multiple-camera applications different from surveillance [11–14] regard the identification of an object moving in the scene. Similarly, in the fields of medicine and sports (mainly athletics, skiing, golf, rugby), there is now a widespread use of multiple cameras to perform anthropometric measurements and reconstruct the trajectories of human gestures (often for the prevention of accidents) [15,16]. Therefore, considering the metrological potential of modern digital stereo-photogrammetry, the low cost of the sensors, the simplicity of use of image processing software, and the promising results of previous tests carried out in a simulated static crime scene [17], we decided to make and test a stereo-photogrammetric video surveillance device (SPVD) based on the acquisition of two simultaneous images of a subject. The surveillance system consists of two simple coupled cameras with parallel axes and a commercial photogrammetric software for the processing of digital images.

The system is based on classical stereo-photogrammetry, widely used in engineering and architecture, which allows you to determine location, size and shape of any object provided two digital images of it are taken from distinct positions with calibrated cameras with known external orientation and arranged with parallel optical axes and about the same distance from the object itself.

After clarifying the anthropometric significance of the terms “stature” and “height” (often incorrectly used as synonyms), we will describe the experimental study aimed at valuating the accuracy of the proposed low-cost sensor in the determination of height and, at the same time, at verifying if it can provide the values of other anthropometric characters for a much more effective assessment of compatibility [18].

In this study we present the first results obtained with the proposed device. To this end, we have realized a simulated crime scene, where the position and orientation of the cameras have been chosen so as to better approximate one of the most recurrent configurations of the video surveillance equipment currently in use in Italy.

2. Materials and methods

The SPVD consists of a pair of digital cameras set up with their optical axes parallel at a distance of 40 cm from each other. The cameras, with a CMOS sensor (2MP, 1/3.2″, zoom 2.8–8.2 mm, F 1.4, video 1600 × 1200 px, 14 fps), were fixed to a horizontal bar that was in turn fixed to the wall with a bracket. An IR projector was located at the center of the bar (Fig. 1a). To assess the accuracy of the measurements, an indoor test site was set up in an atrium of the Engineering Department (Fig. 2).

The stereo device was fixed at a height of 2.65 m above the ground and the optical axes were slightly inclined with respect to the horizontal (Fig. 1a). Thirty-nine Control Points (CPs) were placed on the walls and on the floor and were surveyed with a Topcon GPT-3100N total station in order to determine their 3D coordinates in a local reference system (Fig. 1b).

The two cameras were calibrated using the photogrammetric software PhotoModeler (EOS Systems Inc., v.6) to determine the focal distance, the main point of auto-collimation and the coefficients of the mathematical model of radial distortion. Hence it was possible to correct all the individual frames for the effect of distortion, as shown in Fig. 3.

Subsequently the coordinates of the perspective centers of the cameras and the orientation of the optical axes in the local reference were determined by means of the CPs visible in both frames.

Testing of the video surveillance system was carried out on 11 volunteers (males of different body size, aged between 25 and 50 years) who were asked to walk naturally within the scene while being recorded by the cameras. Each of the volunteers walked along the hallway and stairs and back four times in one minute, providing hundreds of stereo-coupled images to be used for the metrological tests (Fig. 4).

The video recordings provided by the two cameras were synchronized and many stereo pairs of images were extracted. In the selected pairs of images the subjects were framed in front and from behind, in the more upright position possible and in different areas of the scene.

Significant anthropometric landmarks in each pair of images were identified and positioned by means of PhotoModeler (Fig. 5). It is necessary to specify that it is unlikely to detect the individual’s stature in such circumstances. The stature is the distance between *vertex* and *planta* on a subject stretched vertically to the fullest extent and with the head positioned in the Frankfurt plane [19], which passes through *tragion* (above the top of the ear canal) to *orbitale* (the lowest point on lower orbital arch). The same measurement done without extension of the spine can be called “unstretched stature”. When the orientation of the head is lacking,

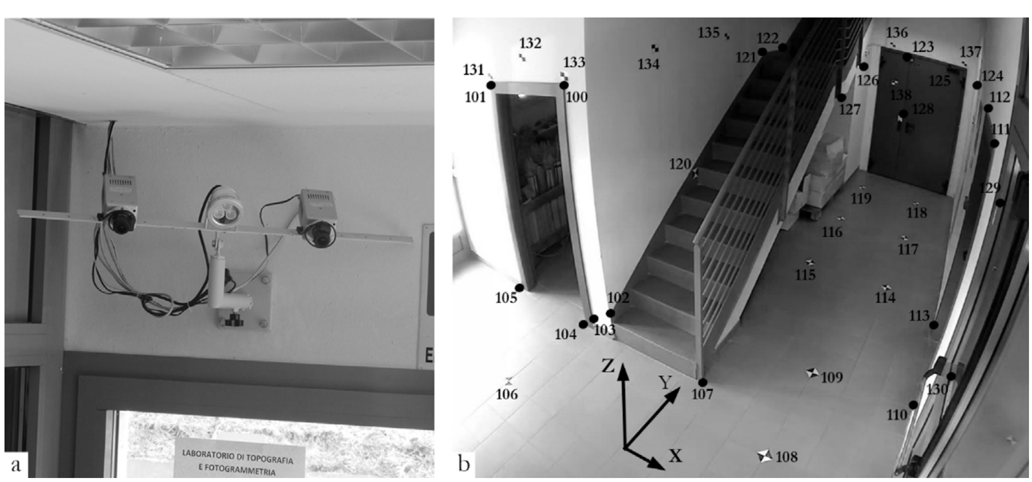

Fig. 1. The test site: the stereo photogrammetric video-surveillance device (first prototype of SPVD) fixed to the wall (a) and an image of the site with the CPs (100–138) provided by one of the cameras (b).