Image and Vision Computing 55 (2016) 22–25

Contents lists available at ScienceDirect

Image and Vision Computing

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

Visible-wavelength iris/periocular imaging and recognition surveillance environments☆,☆☆

Hugo Proença*, João C. Neves

IT: Instituto de Telecomunicações, University of Beira Interior, Portugal

ARTICLE INFO

Article history:
Received 3 February 2016
Accepted 29 March 2016
Available online 8 April 2016

Keywords:
Visual surveillance
Non-cooperative recognition
Iris/periocular recognition

ABSTRACT

Visual surveillance cameras have been massively deployed in public urban environments over the recent years, as a crime prevention and law enforcement solution. This fact raised the interest in developing automata to infer useful information from such crowded scenes (from abnormal behavior detection to human identification). In order to cover wide outdoor areas, one interesting possibility is to combine wide-angle and pan–tilt–zoom (PTZ) cameras in a master–slave configuration. The use of fish-eye lenses allows the master camera to maximize the coverage area while the PTZ acts as a foveal sensor, providing high-resolution images of the interest regions. This paper addresses the feasibility of using this type of data acquisition paradigm for imaging iris/periocular data with enough discriminating power to be used for biometric recognition purposes.

1. Biometrics in surveillance environments

Recent attacks in crowded urban environments reduced the perception of safety in modern societies, while the citizens' tolerance to reasonable risks has been also decreasing. There are now growing needs of assuring the safety of people, particularly in places/events that concentrate large crowds, which are naturally perceived as those with the highest risk (due to e.g., 2001 New York 9/11, 2004 Madrid train bombing, 2013 Boston marathon and 2015 Paris events). To counterbalance this fear, visual surveillance is now deployed massively worldwide. The amount of surveillance cameras running has grown astonishingly in the recent years, with more than 5.9 million CCTV cameras reported only in the United Kingdom [1]. However, contrary to popular belief, there are still no fully automatic techniques to identify subjects without requiring their participation in data acquisition, and the automated understanding of data is most times reduced to action recognition. For every identification attempt, it is still required some kind of human intervention in the process. Even though national/international authorities have lists of potentially harmful individuals, it is particularly difficult for humans to confirm whether such elements are among a crowd. As an example, the *TIDE: Terrorist Identities Datamart Environment* from the U.S. National Counterterrorism Center has over 745,000 people listed in the database which authorities are willing to arrest, but only a small proportion of these was actually detected in visual surveillance systems.

One interesting possibility is using coupled wide-angle and PTZ devices, which are able to acquire high resolution images on arbitrary scene locations. In this kind of configuration, a master–slave paradigm is usually adopted, i.e., the wide-angle camera covers the whole scene and provides data both for detecting and tracking subjects, also supplying 3D cues about the position to where the PTZ camera should be pointed to. While several advantages of this paradigm can be outlined, inter-camera calibration is the major bottleneck of this configuration, since determining the mapping function from image coordinates to pan–tilt parameters requires depth information. A solution to this problem is described in [2] and illustrated in Fig. 1: by inferring the subjects' height h, the depth ambiguity problem can be avoided and a univocal correspondence between positions in the wide-angle image data (x_s, y_s) and in the 3D physical coordinate system (X_p, Y_p, Z_p) can be obtained, enabling to infer the pan–tilt angle (θ_p, θ_t) values required to center the PTZ device at a particular position in the scene.

Another obvious difference between the operating requirements of systems working in visual surveillance scenarios and the traditional stop-and-stare protocol currently used is that in the former type of environments the number of targets usually exceeds the available active cameras, which demands schedule techniques to not only maximize the number of targets imaged but also the number of

---

☆ This paper has been recommended for acceptance by Sinisa Todorovic, PhD.

☆☆ This work was supported by FCT: Fundação Ciência e Tecnologia project UID/EEA/50008/2013.

* Corresponding author at: University of Beira Interior, R. Marques D'Avila e Bolama, 6201-001 Covilhã, Portugal.

E-mail addresses: hugomcp@di.ubi.pt (H. Proença), jcneves@di.ubi.pt (J. Neves).

http://dx.doi.org/10.1016/j.imavis.2016.03.015
0262-8856/

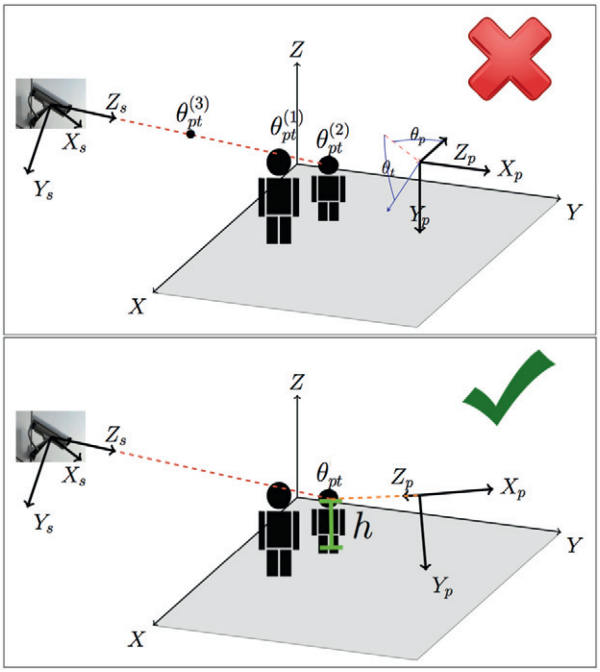

Fig. 1. Top image: illustration of the *depth ambiguity* problem, with several physical positions in the scene correspond to the same pan–tilt angles of the PTZ device ($\theta_{pt}^{(i)}$). Bottom image: by estimating subjects' height (h), it is possible to establish a bijection between physical 3D positions and pan–tilt angles θ_{pt}.

shots taken from each one. This is a variant of the classical *optimal tour* finding problem, which exhaustive solution has time cost $O(n!)$, being n the number of targets in the scene. Although brute-force might be feasible for a reduced number of targets, the real-time nature of the problem prohibits its use for more than six targets. Several works have presented solutions to this problem (e.g., [3] and [4]), where the contextual and dynamic scene information is considered to find the optimal sequence of targets (Fig. 2).

2. Related work

In order to consider an image with *acceptable* quality, the iris recognition standards recommend a resolution of at least 100 pixels across the iris diameter (ISO/IEC 2004) and an in-focus image. Also, sufficient near infrared (NIR) illumination should be ensured (more than 2 mW/cm^2) without harming human health (less than 10 mW/cm^2 according to the international safety standard IEC-60852-1). The space volume in front of the image acquisition system where these constraints met is usually referred as the capture volume of the system. Commercial iris recognition systems achieve extremely low error rates, yet imposing highly restrictive capture volumes that reduce the workability in less constrained scenarios. In recent years, several attempts to relax the constraints of iris recognition systems have been made, exploiting innovative strategies to increase both the capture volume and the stand-off distance, i.e., the distance between the front of the lens and the subjects.

Current strategies to perform the acquisition of iris data in less constrained conditions can be divided into two families, depending of whether they use (or not) magnification devices. In terms of the approaches that make no use of magnification devices, the Iris-on-the-Move [5] system is notable for having significantly decreased the cooperation levels required for image acquisition, allowing subjects continuous movement through a portal equipped with NIR illuminators. Another well known commercial device is the LG IrisAccess4000, where image is acquired at-a-distance, provided that subjects' gaze point at a specific direction.

Magnification devices, such as PTZ cameras, extend the system stand-off distance while providing enough resolution for reliable iris recognition. Wheeler et al. [6] introduced a system to acquire iris data at a resolution of 200 pixels from cooperative subjects at 1.5 m, using a PTZ camera assisted by two wide view cameras. Dong et al. [7] also proposed a PTZ-based system, that images iris data up to distances of 3 m with more than 150 pixels across the iris diameter. Yoon et al. [8,9] relied on a light stripe to determine the 3D position, avoiding the use of an extra wide camera. The *Eagle Eye* system [10] uses one wide view camera and three close view cameras, for capturing simultaneous images of both irises. This system has a stand-off distance of about 5 m with a operational range of 3 m × 2 m × 3 m. This system uses a bi-ocular setup, that enables to recover the 3D world position of the subject by stereo reconstruction. Depth information cues are used both for pan/tilt angles estimation and for getting focused data.

Despite being considered more reliable, the use of two wide-angle cameras significantly increases the system cost and limits its flexibility. To address this problem, various commercial solutions were introduced: Mitsubishi corporation developed a scheme where depth is estimated using the disparity between facial features [11]. Yoo et al. [12] combined the wide-view and narrow-view cameras with a beam splitter to simultaneously acquire facial and iris images. This integrated dual-sensor enables the same ray to be mapped to same position in both cameras sensors, avoiding the need for depth estimation.

3. Challenges

Most of the current iris recognition systems require that the iris is illuminated in the NIR wavelength band. Although this wavelength has the major advantage (with respect to the visible band) of avoiding corneal reflections from the surrounding light, the use of NIR illuminators highly restricts the workability of iris recognition in less constrained scenarios: the irradiance of the illuminators decreases quadratically as the stand-off distance increases, implying the use of extremely powerful illuminators for acquiring the rich details of the iris from large distances. As such, from our viewpoint non-cooperative iris recognition at such large stand-off distances such as in typical surveillance scenarios should be performed in the visible spectrum. Also, we believe that the use of magnification devices (PTZ) cameras is the most efficient solution to acquire iris with sufficient quality for recognition purposes.

We recently described a system—named QUIS-CAMPI—for acquiring high-resolution face imagery at large distances (up to 50 m) [2,13], but here we discuss its usability for unconstrained acquisition of iris/periocular data. This system uses a PTZ camera with full-HD resolution (1920 × 1080) and 30× optical zoom, corresponding to an angle-of-view of 2.1°. We used this framework to acquire close-up shots of the ocular region at standoff distances of 10 and 15 m, with examples for six subjects being illustrated in Fig. 3. Apart from the typical variability factors of unconstrained scenarios (e.g., occlusions due to eyelids and reflections, poorly focused and off-angle data), the resolution across the iris is a key factor for the reasonability of using a system as QUIS-CAMPI for iris recognition in unconstrained scenarios. Using as baseline the standard that recommends 100 pixels across the iris, it can be seen that we are able to get only about 60% and 40% of that resolution respectively at 10 m and 15 m stand-off distances.

In order to determine the maximum stand-off distance that can be afforded without compromising iris quality, we investigate how

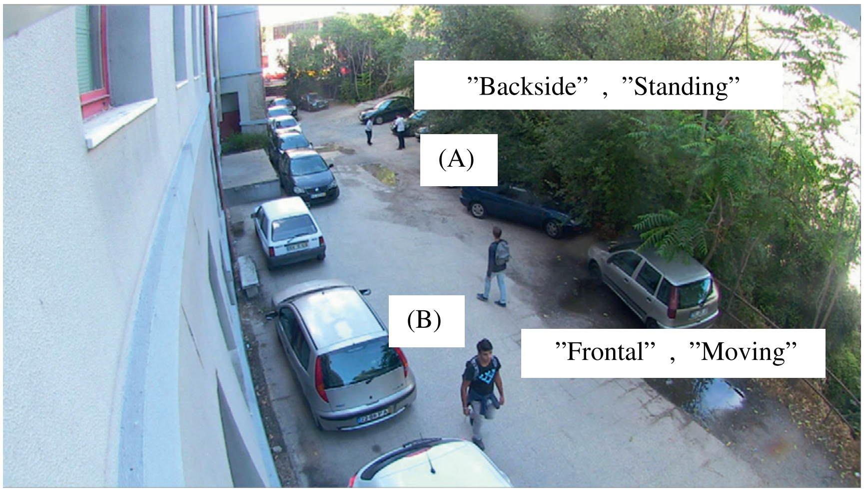

Fig. 2. Target selection: having *n* subjects in the scene, it is important to consider the scene context and dynamics (e.g., number of samples taken previously from each subject, subjects' position, velocity and perspective) to find the *optimal PTZ tour*, i.e., the data acquisition order. In this example, (B) is evidently a better target than (A).

the number of pixels across the iris ($\psi(d)$) is affected by the stand-off distance (d), angle of view (α) and camera resolution horizontal resolution (ω). Assuming that the human iris has an average diameter of 1.2 cm, the number of pixels across the iris is given by:

$$\psi(d, \alpha, \omega) = \frac{1.2 \times \omega}{2d\tan(\frac{\alpha}{2})}. \quad (1)$$

This relation is depicted in Fig. 4 when using the PTZ camera in QUIS-CAMPI (blue) and when using the next generation of PTZ cameras (4K). For comprehensibility, the chart is divided into three regions according to the quality of data with respect to the resolution factor. It is evident that state-of-the-art PTZ cameras are not sufficient to image iris at large stand-off distances, but it is worth noting that 4 K PTZ cameras should be available soon, and that their maximum optical zoom is also expected to increase. As illustrated in Fig. 4 such type of devices should have an obvious impact in the resulting image resolution, and allow the iris imaging with reasonable quality up to stand-off distances of over 15 m.

Another noteworthy possibility is the use of periocular region as the main biometric trait, which has been advocated as an interesting possibility to increase the robustness of iris recognition in visible-light data. The idea is to compensate for the degradation in the iris data by also considering the discriminating information in the surroundings of the eye (eyelids, eyelashes, eyebrows and skin texture). Even though further empirical validations are required to confirm the reasonability of using the periocular region as biometric trait in this type of data, it is known that periocular recognition is much less demanding in terms of data resolution than iris recognition.

4. Conclusions

Developing automata able to perform biometric recognition in crowded scenes and without explicitly requiring any active human

Fig. 3. Examples of facial data acquired in surveillance environments with the corresponding iris/periocular regions. The upper row contains samples acquired 10m away from the subjects, while the bottom row illustrates images acquired from 15m away.

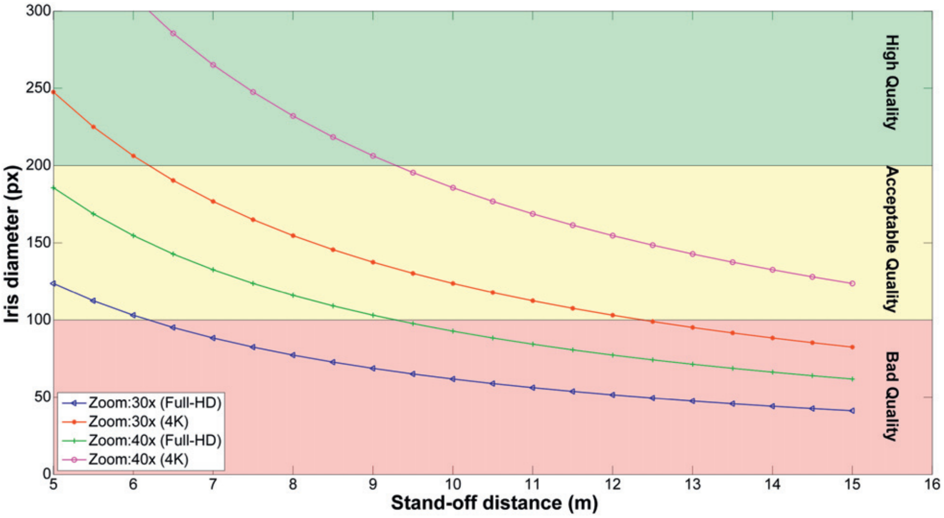

Fig. 4. Relation between the number of pixels across the iris and the stand-off distance when using a different magnification zoom and resolution for the PTZ camera. Considering the current standards as baseline, current PTZ cameras (blue line) do not appropriately image iris data at large stand-off distances. On contrast, 4K resolution cameras (pink line) extend the maximum stand-off distance of the system up to 15 m away from the subjects.

effort in the data acquisition process is an ambition that dates back—at least—to 1949, as a result of the widely famous George Orwell's *Big Brother* character. Even though such type of machine raises evident concerns from the ethical/privacy protection perspectives, it is also obvious that it will constitute a valuable law enforcement/security tool. Among several alternatives, one interesting possibility for such kind of system is to use coupled wide-angle and PTZ devices, that not only cover large outdoor areas, but are also able to acquire high-resolution data from moving subjects and large distances. In this paper we discussed some of the major differences between the processing chains of such type of non-cooperative recognition systems and of the current biometrics operating mode. Also, we illustrated the variations in the resulting facial/iris data with respect to the subjects stand-off distance factor and speculate about the suitability of using the periocular region as main biometric trait in such conditions.

References

[1] D. Barret, One surveillance camera for every 11 people in Britain, says CCTV survey, 2014. the Telegraph (, July 10, 2006. Retrieved June 17, 2014)http://www.telegraph.co.uk/technology/10172298/.

[2] J.C. Neves, J.C. Moreno, S. Barra, H. Proença, Acquiring High-resolution face images in outdoor environments: a master–slave calibration algorithm, IEEE Seventh International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2015. pp. 1–6.

[3] N. Kariotoglou, D. Raimondo, S. Summers, J. Lygeros, A stochastic reachability framework for autonomous surveillance with pan–tilt–zoom cameras, 50th IEEE Conference on Decision and Control and European Control Conference, 2011. pp. 1411–1416.

[4] J.C. Neves, H. Proença, Dynamic camera scheduling for visual surveillance in crowded scenes using Markov random fields, 12th IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS), 2015. pp. 1–5.

[5] J.R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D.J. LoIacono, S. Mangru, M. Tinker, T.M. Zappia, W.Y. Zhao, Iris on the move: acquisition of images for iris recognition in less constrained environments, Proc. IEEE 94 (11) (2006) 1936–1947.

[6] F.W. Wheeler, G. Abramovich, B. Yu, P.H. Tu, et al. Stand-off iris recognition system, IEEE International Conference on Biometrics: Theory, Applications and Systems, (BTAS), 2008. pp. 1–7.

[7] W. Dong, Z. Sun, T. Tan, A design of iris recognition system at a distance, Chinese Conference on Pattern Recognition, (CCPR), 2009. pp. 1–5.

[8] S. Yoon, H.G. Jung, J.K. Suhr, J. Kim, Non-intrusive iris image capturing system using light stripe projection and pan–tilt–zoom camera, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007. pp. 1–7.

[9] S. Yoon, H.G. Jung, K.R. Park, J. Kim, Nonintrusive iris image acquisition system based on a pan–tilt–zoom camera and light stripe projection, Opt. Eng. 48 (3) (2009) 1–15.

[10] F. Bashir, P. Casaverde, D. Usher, M. Friedman, Eagle-eyes: a system for iris recognition at a distance, IEEE Conference on Technologies for Homeland Security, IEEE 2008, pp. 426–431.

[11] G. Guo, M. Jones, P. Beardsley, A system for automatic iris capturing, MERL TR2005-044, 1. 2005.

[12] J-H. Yoo, B. Kang, A simply integrated dual-sensor based non-intrusive iris image acquisition system, IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015. pp. 113–117.

[13] J.C. Neves, G. Santos, S. Filipe, E. Grancho, S. Barra, F. Narducci, H. Proença, Quis-Campi: extending in the wild biometric recognition to surveillance environments, ICIAP Workshops, 2015. pp. 59–68.