Pattern Recognition Letters 68 (2015) 218–230

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

On soft biometrics☆

Mark S. Nixon[a,*], Paulo L. Correia[b], Kamal Nasrollahi[c], Thomas B. Moeslund[c], Abdenour Hadid[d], Massimo Tistarelli[e]

[a] University of Southampton, UK
[b] Instituto de Telecomunicações, Instituto Superior Técnico, Lisboa, Portugal
[c] Aalborg University, Denmark
[d] University of Oulu, Finland
[e] University of Sassari, Italy

ARTICLE INFO

Article history:
Available online 8 September 2015

Keywords:
Soft biometrics
Recognition
Surveillance
Body
Gender
Age

ABSTRACT

Innovation has formed much of the rich history in biometrics. The field of soft biometrics was originally aimed to augment the recognition process by fusion of metrics that were sufficient to discriminate populations rather than individuals. This was later refined to use measures that could be used to discriminate individuals, especially using descriptions that can be perceived using human vision and in surveillance imagery. A further branch of this new field concerns approaches to estimate soft biometrics, either using conventional biometrics approaches or just from images alone. These three strands combine to form what is now known as soft biometrics. We survey the achievements that have been made in recognition by and in estimation of these parameters, describing how these approaches can be used and where they might lead to. The approaches lead to a new type of recognition, and one similar to Bertillonage which is one of the earliest approaches to human identification.

1. Introduction

One of the oldest biometrics approaches derived from the work by Bertillon in the 19th century and one of the newest areas is that of soft biometrics. The work of Bertillon was superseded by fingerprints after showing early promise; the interest in soft biometrics is reflected in the increasing volume of papers that mention it. Unlike Bertillonage, soft biometrics is unlikely to be superseded as it can be used to reinforce biometric identification as well as be deployed alone in the analysis of data that conventional biometrics cannot handle or with invariant attributes that conventional biometrics cannot even approach.

Further, the UK Niche Records Management System has primary features such as a suspect's name, address and date of birth. Other desirable fields include factors used in Bertillon's approach such as gender, ethnicity and height. The FBI requires some demographic information to conduct a fingerprint-based background check. The information includes sex/gender, race, height, weight, eye and hair color. These factors are often included on suspect identification forms, along with other descriptions of a suspect's build and clothing. These factors are primary human factors that we shall see can be identified from (color) image data that are now termed soft biometrics.

As we shall find in this survey, there is a plethora of soft biometrics. The original formulation [42, 43] concerned measures that can be used to aid recognition rather than for identification. The measures were suited to the discrimination between classes rather than individuals, and could be used to buttress recognition performance. Later techniques were to extend the definition, to concern estimation of personal characteristics, such as height and gender. Such factors have been of interest from the earliest days of biometrics, and are of interest given their ability to split populations. Given the pervading need for security in modern environments, there has been a concerted interest for the use of soft biometrics [84] largely to handle the low quality of video images where traditional biometrics cannot be applied. This is reflected in a theme of soft biometrics for identifying people, originally termed semantic biometrics [89].

In this paper we shall review the approaches that have been made in this new field. We shall provide an updated definition of the term soft biometrics and describe its advantages in biometric recognition. There is a contemporaneous review with similar aims [20] that offers an alternate view into this new and exciting topic. We shall concentrate on initial approaches and state-of-art techniques, aiming to establish the main basis of these new approaches and where they have developed to. We shall complete by suggesting new areas for development in this new and fascinating set of approaches.

☆ This paper has been recommended for acceptance by Gabriella Sanniti di Baja.

* Corresponding author. Tel.: +44 2380 593542; fax: +44 2380 594498.
E-mail address: msn@ecs.soton.ac.uk (M.S. Nixon).

http://dx.doi.org/10.1016/j.patrec.2015.08.006
0167-8655/

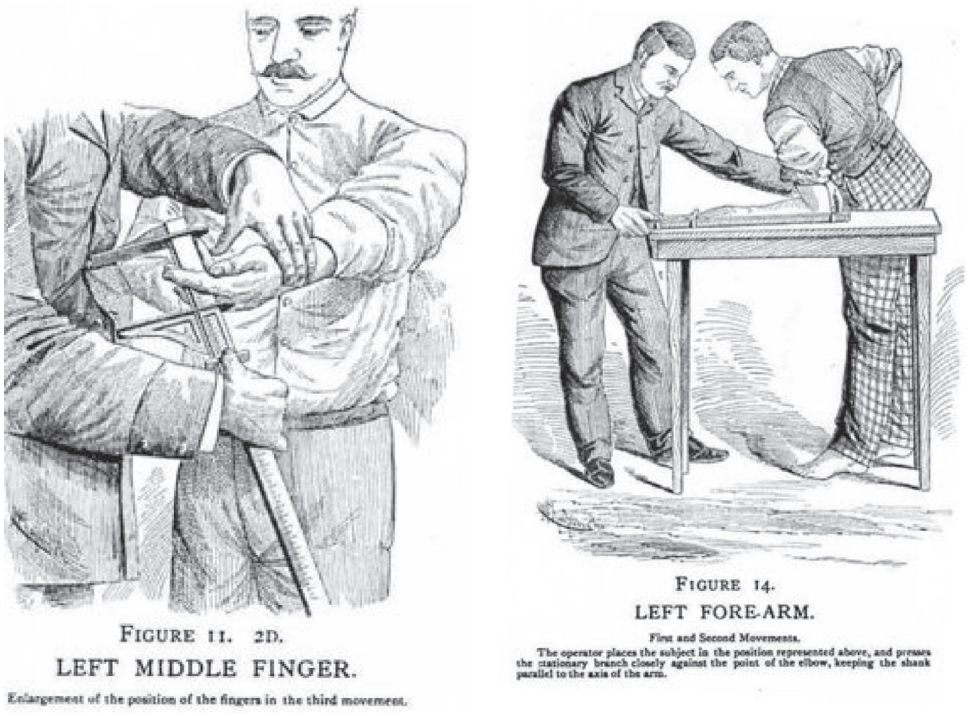

Fig. 1. Examples of Bertillon's gathering of measurements [6].

1.1. Bertillonage

A variety of factors motivated a precise system for identification in 19th century France after the abandonment of branding criminals and a system of deportation. Limitations of the descriptions used to identify people, especially in size of forehead and color of the eye motivated Bertillon to develop an anthropometric system to systematically describe people for identification based on their physiological traits in 1879 [6]. Though the photographs were useful for confirmation of identity, the acquisition techniques were not standardized. A photograph is of use for verification of identity rather than to discover it, when relying on manual search. Bertillon noted the failings of the police identification and cataloguing system and developed his father's anthropological work to a more systematic method of identifying people. His system of anthropometrics, eponymously Bertillonage, outlined the tools and techniques for the careful measurement of:

- Physical features including length/width of head, lengths of certain fingers and the dimensions of the feet, arm, right ear and standing height;
- Descriptions of the dimensions of the nose, eye and hair color; and
- The description and location of notable scars, tattoos and other marks.

The method for gathering these features was outlined in Bertillon's manual [6] along with a set of diagrams (see Fig. 1). He implored 'police authorities … not to introduce special modifications of their own' for this could destroy uniformity, a practice echoed in modern biometrics databases and experimental practices. The measurements for a given individual were held on separate slides along with standardized photographs of the individual. The metrics of the system were chosen primarily to be simple so that they could be gathered accurately. The measurements were taken by a trained individual, though not necessarily a skilled one. Features were chosen to allow easy identification of points to begin and to end measurement. The success of Bertillonage came from its ability to reduce the probability of type 1 errors. Though two individuals may have very similar height, the chance of the same two having similar measurements for the other features is unlikely. Furthermore, Bertillonage inherently allowed for efficient discovery of an individual's existing measurement card and therefore their identity. Cards were stored according to specific range combinations of each metric in a given order. As such that once new measurements of an unidentified individual were taken then the identity of the individual could be easily ascertained.

Achieving great success and popularity in France, Bertillonage progressed too many countries by the late 19th century. Difficulties in cases such as West vs. West [13] where Bertillonage appeared to be unable to distinguish people of similar appearance is often quoted as a reason for it being superseded by forms of identification such as fingerprint analysis (since the fingerprints of identical twins differ). It is of note that recently there has been a study on the potency of body measurements vs those of the face, for identification, in some ways redeeming Bertillon's original approach [60]. The study was based on using anthropometric measurements of 3982 individuals from the US Army Anthropometry Survey (ANSUR) database and concluded that "The body is more variable than the face and should be used in identification" also stating practical advantage.

These systems aim to reduce identity to a representative and measurable set of features, though not using descriptions of the human body as a whole. Measurements are taken in a controlled way, much the same as in modern biometrics, though lacking its sophisticated statistical and recognition techniques.

1.2. On the development of soft biometrics approaches

The earliest approach explicitly mentioning a form of soft biometrics appears to be by Wayman [101] who proposed the use of soft biometric traits like gender and age, to filter a large biometric database. An early motivation of the first soft biometrics was to augment conventional biometric signatures and envisaged that soft biometrics would be obtained separately, perhaps not from images originally used for recognition, and then used to enrich the biometric signature. A database was acquired that comprised of four fingerprint impressions of 160 users (obtained using a Veridicom sensor) of which a