Int. J. Human-Computer Studies 86 (2016) 32–47

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

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

Transformations of Gaussian Process priors for user matching☆,☆☆

Shimin Feng*, Roderick Murray-Smith

School of Computing Science, University of Glasgow, Sir Alwyn Williams Building, 18 Lilybank Gardens, Glasgow G12 8QQ, United Kingdom

ARTICLE INFO

Article history:
Received 17 April 2014
Received in revised form
26 July 2015
Accepted 1 September 2015
Communicated by E. Motta
Available online 11 September 2015

Keywords:
User matching
User identification
Proxemic interaction
Gaussian processes
Sensor fusion
Mobile sensing

ABSTRACT

We describe the use of transformations of Gaussian Process (GP) priors to improve the context sensing capability of a system composed of a Kinect sensor and mobile inertial sensors. The Bayesian nonparametric model provides a principled mechanism for incorporating the low-sampling-rate position measurements and the high-sampling-rate derivatives in multi-rate sensor fusion which takes account of the uncertainty of each sensor type. The complementary properties of these sensors enable the GP model to calculate the likelihood of the observed Kinect skeletons and inertial data to identify individual users.

We conducted three experiments to test the performance of the proposed GP model: (1) subtle hand movements, (2) walking with a mobile device in the trouser pocket, and (3) walking with a mobile device held in the hand. We compared the GP with the direct acceleration comparison method. Experimental results show that the GP approach can achieve successful matches (with mean accuracy $\mu > 90\%$) in all 3 contexts, including when there are only subtle hand movements, where the acceleration comparison method performs poorly ($\mu < 20\%$).

1. Introduction

Identification of people and their positions in a room plays a key role in proxemic[1] interactions. Greenberg et al. operationalised the concept of proxemic interaction within ubiquitous computing and proposed five proxemic dimensions including distance, orientation, identity, movement and location for proxemic interaction (Ballendat et al., 2010; Marquardt et al., 2011; Greenberg et al., 2011). The knowledge about the identity of a person or a device is critical in proxemic-aware applications (Ballendat et al., 2010).

When several users are in a sensor-augmented room (e.g. using a Microsoft Kinect depth sensor) and each of them carries a motion sensor enhanced personal mobile device (e.g. with accelerometers), it is possible to find the matching relationship between individual users and the mobile devices. A personal device can then provide the means to associate an identity with a tracked user (Ackad et al., 2012), implicitly providing a way for user identification through user matching, i.e. finding the correlation between the multiple skeletons (users) and the mobile devices. In practice this can be challenging because the different types of sensors have different noise and sampling properties, as well as measuring different physical quantities. In this paper, we apply a novel and improved Gaussian Process prior model to fuse the low-sampling-rate position measurements sensed by the Kinect and the higher frequency acceleration measured by the mobile inertial sensors. The sensor fusion combines data from multiple sensors (Hall and Llinas, 1997), and can be applied to improve the accuracy and speed of measuring the match between a set of users' skeletons and a set of candidate mobile devices.

1.1. Matching scenario

To illustrate this, we propose a scenario of two people using a proxemic interaction system in a room, as shown in Fig. 1. The system can display the users' favorite books and also make personalised recommendations for them (Funk et al., 2010). Two people (Jim and Tom) walk into the room. Each carries a mobile device in the trouser pocket or in the hand. The Kinect starts tracking and assigns a user ID to each person. Jim is user 1 and Tom is user 2. As a personal device can provide the means to associate an identity with a tracked user (Ackad et al., 2012) and the system can detect the identities of the personal devices via the wireless network, we know who the user is if we can link a particular skeleton with one of the mobile devices. This enables the

☆This paper has been recommended for acceptance by E. Motta.
☆☆This document is a collaborative effort.

* Corresponding author.
E-mail addresses: shiminf@dcs.gla.ac.uk (S. Feng), Roderick.Murray-Smith@glasgow.ac.uk (R. Murray-Smith).

[1] *Proxemics* is the theory proposed by Edward Hall about people's understanding and use of interpersonal distances to mediate their interactions with others (Hall and Llinas, 1997).

http://dx.doi.org/10.1016/j.ijhcs.2015.09.001
1071-5819/

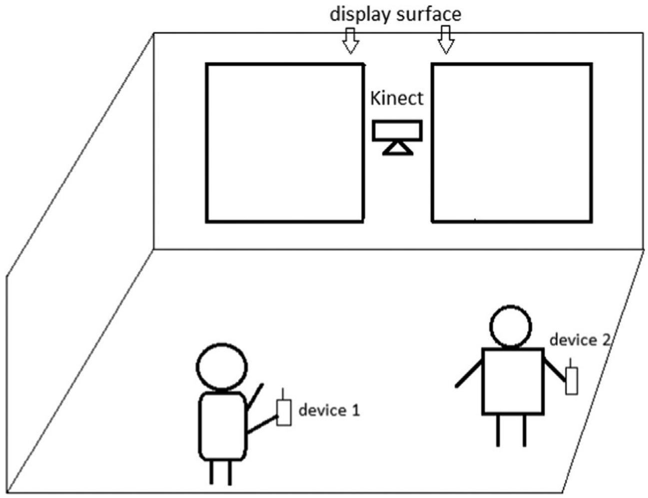

Fig. 1. A scenario of two people using a proxemic interaction system in a room. Proxemic interaction relates the two users to their personal devices by matching the motion sensed by the Kinect with the motion sensed by the devices when they carry the devices and move in the field of Kinect's view. The personalised content will be displayed when the user approaches the surface as the system knows the identity of the user through matching the user with the personal device.

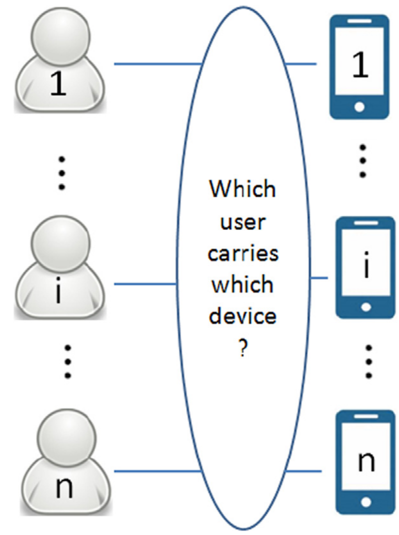

Fig. 2. Illustration of the user–device matching problem.

system to provide a personalised service when a user approaches a display surface through proximity interaction.

Designing technologies that are embedded in people's everyday lives plays an important role in context-aware applications (Bilandzic and Foth, 2012). The process mentioned above may involve a variety of people's everyday movements, including moving with a device in the trouser pocket, the subtle hand movements or walking with a device held in the hand (Barnard et al., 2005). Vogel and Balakrishnan (2004) proposed an interaction framework for ambient displays that support the transition from implicit to explicit interaction by identifying individual users through registered marker sets, and argued for the need of marker-free tracking systems and user identification techniques.

In the above scenario, the system can achieve user matching and identification implicitly. At certain spatial locations above the surface, we can design a spatially aware display application that links the digital books with the spatial locations. An important issue in this proxemic interaction system is the accuracy of position tracking. In order to reduce the joint position uncertainty and improve the interaction performance and experience of the users (Jim and Tom), we proposed a sensor fusion approach to stabilising the hand position in the Kinect space by fusing the Kinect sensor and the mobile inertial sensors in Feng et al. (2015). After user matching, we can apply the acceleration sensed by Jim's device to compensate for the effects of position uncertainty and lags in Jim's skeleton tracking sensed by the conventional Kinect system, giving a smoother, more responsive experience.

1.2. Sensing in proxemic interaction

Sensors provide the position and posture of users in a proxemic interaction system. As the advanced sensors become ubiquitous, many systems are composed of a range of elements which observe the world via a diverse set of sensors. These sensors might work at a range of sampling rates, depending on power constraints. They may measure different derivatives of measurands (e.g. position, velocity, acceleration) in the world and they might have different noise characteristics. If we can fuse information from such systems in an efficient and principled manner, we can potentially improve the capability of the system without adding extra sensing hardware. A concrete example of this is integration of inertial data from mobile devices such as phones or tablets with position sensing from an embedded Microsoft Kinect sensor, but the same principle can be found in many systems. The Kinect is a commercially successful sensor for human body tracking, being low-cost, portable and unobtrusive in a room. If the Kinect can sense multiple people in the room and each has a device in the hand or pocket, can we devise algorithms to rapidly and reliably determine which person carries which device (Fig. 2), even when movements are subtle?

The visual sensing and inertial sensors have complementary properties (Hol et al., 2007). The Kinect senses human posture and location, but the inferred positions are subject to significant uncertainty (Casiez et al., 2012). Inertial sensors have been used for sensing human movement (Luinge, 2002) and can be used to measure the acceleration of body segments. These sensors have different noise characteristics and are subject to disturbances like muscle tremor (Strachan and Murray-Smith, 2009). Solving this efficiently by exploring the complementary properties of these sensors, minimising the extent of the movement needed for matching, would allow a system to rapidly link the Kinect-sensed people with their devices equipped with inertial sensors, and customise services appropriately for them.

1.3. Probabilistic model-based user matching approach

In this work, we focus on using a probabilistic approach to deal with the smoothness of human movement, and apply the inference results for user matching and identification. The inputs are the timing information t and the outputs are the sensor measurements including position and acceleration. There is a relationship between inputs t and outputs y, which will be a time-series when the largest changes in time are limited by the motor ability of the human body, $t \rightarrow y$. We will take a probabilistic approach to correlating the measurements of human movement sensed by multiple sensors, and make use of the likelihood of the model to infer the best estimate of correlation between different sensor measurements, i.e. the position time-series and the acceleration time-series, considering the smoothness of human movement.

For a user, we want to match a specific series of position observations with one of the acceleration time-series. In this paper, we will look at how the user–device combinations can be ranked based on which model they fit. In order to do this, we would prefer to take a flexible probabilistic approach to function modelling. We propose a variation of a Gaussian Process prior