Information Fusion 40 (2018) 18–31

Contents lists available at ScienceDirect

Information Fusion

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

Full Length Article

Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach

Eiman Kanjo[a,*], Eman M.G. Younis[b], Nasser Sherkat[a]

[a] Computing and Technology, Nottingham Trent University, United Kingdom
[b] Faculty of Computers and Information Minia University, Egypt

ARTICLE INFO

Article history:
Received 18 January 2017
Revised 13 April 2017
Accepted 28 May 2017
Available online 29 May 2017

Keywords:
Multi sensor data fusion
Regression analysis sensor data
Multivariable regression
Affective computing
Physiological signals
Machine learning

ABSTRACT

Over the past few years, there has been a noticeable advancement in environmental models and information fusion systems taking advantage of the recent developments in sensor and mobile technologies. However, little attention has been paid so far to quantifying the relationship between environment changes and their impact on our bodies in real-life settings.

In this paper, we identify a data driven approach based on direct and continuous sensor data to assess the impact of the surrounding environment and physiological changes and emotion.

We aim at investigating the potential of fusing on-body physiological signals, environmental sensory data and on-line self-report emotion measures in order to achieve the following objectives: (1) model the short term impact of the ambient environment on human body, (2) predict emotions based on-body sensors and environmental data.

To achieve this, we have conducted a real-world study 'in the wild' with on-body and mobile sensors. Data was collected from participants walking around Nottingham city centre, in order to develop analytical and predictive models.

Multiple regression, after allowing for possible confounders, showed a noticeable correlation between noise exposure and heart rate. Similarly, UV and environmental noise have been shown to have a noticeable effect on changes in *ElectroDermal Activity (EDA)*. Air pressure demonstrated the greatest contribution towards the detected changes in body temperature and motion. Also, significant correlation was found between air pressure and *heart rate*.

Finally, decision fusion of the classification results from different modalities is performed. To the best of our knowledge this work presents the first attempt at fusing and modelling data from environmental and physiological sources collected from sensors in a real-world setting.

1. Introduction

Repeated exposures to environmental stressors (such as pollution, noise and crowded areas) cause physical illnesses (e.g., headaches, fatigue, sleeping disorder, and heart diseases) and behavioural issues (e.g., stress, attention deficit, anger, and depression) [1–3].

The effect of these stressors on health has been a focal point in health research. Models have been widely used as indispensable tools to assess effects of environmental factors on human and health. In particular, modelling the level of exposures to environmental pollutants such as [4,5].

A decade-long study of 6.6 million people, published in the Lancet recently, found that one in 10 dementia related deaths in people living within 50 m of a busy road was attributable to fumes and noise. There was a linear decline in deaths the further people lived away from heavy traffic [6].

Additionally, Chen's group [6] noted that because air pollution exposure was estimated at the postal-code level, it may not account accurately for each individual's exposure. The study suggested that more research to understand this link is needed, particularly into the effects of different aspects of traffic, such as air pollutants and noise at a higher granular levels.

In general, epidemiological and statistical analysis are usually studied based on observed environmental data, which have traditionally been obtained from governmental sources or from a number of sporadically distributed sensing nodes. In both cases, the

* Corresponding author.
E-mail addresses: eiman.kanjo@ntu.ac.uk (E. Kanjo), eman.younas@mu.edu.eg (E.M.G. Younis), Nasser.Sherkat@ntu.ac.uk (N. Sherkat).

http://dx.doi.org/10.1016/j.inffus.2017.05.005

performance of these studies is evaluated against relatively few directly measured data points [7].

Conversely, the capabilities and availability of cheaper, more sensitive and sophisticated sensors for gases, particulates, water quality, noise and other environmental measurements have improved and are enabling researchers to collect data in unprecedented spatial, temporal and contextual detail [7,8].

These sensors range from bespoke devices designed for specific applications, to those found on more mainstream personal devices, such as smartphones. In some cases, people may act as environmental sensors by reporting what they see, hear and feel by participating in the citizen science of environmental conditions [9]. By leveraging widely available wearable devices, communication and sensor technologies many new sensor systems are relatively low-cost compared with technologies used in established monitoring stations [10,11].

Advances in data science and fusion techniques are critical to enable researchers to make best use of the vast amounts of additional, heterogeneous sensor data sources.

Despite the popularity of using wearable sensors for emotion recognition, the problem of quantifying the relation between environmental variables and physiological body reactions and emotions has been overlooked. In addition, the relationship between emotions and all the other environmental and body factors have been studied qualitatively.

In this paper, we incorporate a sensor-data driven approach to understand the relationship of various environmental measures with wellbeing and emotion. By unobtrusively collecting data from on-Body and environmental sensors we can get better understanding of the association and causality of the environmental bases for human health including psychological changes.

This leads us to investigate the following research questions:

1. How can we model and fuse the relationship between on-body and environmental variables?
2. Can the multi heterogeneous sensors integration improve our understanding of the associations and environmental impact on human health?
3. How can information fusion best make use of the ‘on-body and environmental Sensor Data’ to infer emotion?

Our approach to answer these questions is based on two phase framework in information fusion, which utilizes the new available heterogeneous sensors of multiple modalities as mobile interfaces by studying the relationship between these data sources in spatial-temporal context. Moreover, by studying its relationship with emotion based on decision fusion.

In order to follow our approach, we collected data from forty subjects using on-body sensors ‘in the wild’ around Nottingham city centre environment. The data collected include on-body data such as body movement, heart rate (*HR*), Electrodermal activities (*EDA*) and body temperature and, environmental data including noise level (Env-noise), air pressure and ambient light levels (*UV*), as shown in Fig. 1.

In addition, collected *GPS* data record the user locations while gathering data. The different data channels are collected, cleaned, aggregated and smoothed for different users and user emotions labels are collected using self-report input, based on 5-step SAM Scale for Valence taken in [12].

The selection of sensors and data analysis techniques is optimized from the ground up with the emotion inference application in mind for outdoor environments.

We have adopted an information fusion approach to analyse and model the data since this method offers an effective solution to many of the issues found in analysing data from individual sensors. Information fusion allows integration of independent features and prior knowledge and, provides a better means of identifying specific aspects of the target application domain and improve robustness against interferences of data sources [13].

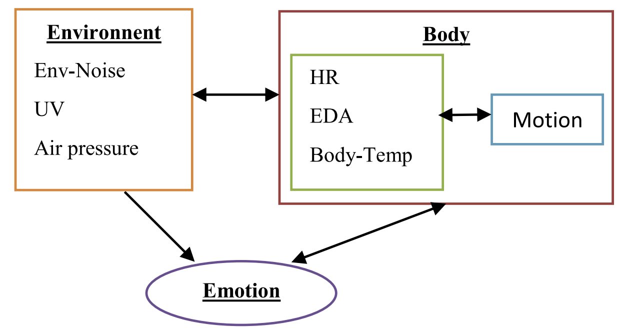

Fig. 1. The relationship between different modalities, the environment, human body, Motion and emotions data.

For examples physiological data, such as heart rate reveal the physical effort of an activity but they may be influenced by external factors such as environmental conditions or social interaction. All of these sources provide only partial information related to the actual individuals’ activity.

In this work we utilise, multi-sensor fusion to demonstrate the feasibility of capturing diverse and multi-model derived features in order to identify relationships, associations and causality and, formalize models describing people’s reaction and emotions.

Our data fusion approach is in three folds: (1) Data fusion by collecting data from multiple sources including *HR, EDA*, body temperature, movement and activity, environmental noise, location, air pressure and *UV*. (2) Feature fusion by examining relationship between our environmental variables and physiological variables based on exploratory statistics and Multivariate Regression modelling, also by looking at the variable importance and variation (3) Decision fusion by combining multiple classifiers from different modalities for emotion prediction.

The rest of the paper is structured as follows. Section 2 discusses related work focusing on previous efforts in quantifying environmental health impact along with a brief review of on-Body sensors and related information fusion techniques. Section 3 covers the methodology including the user study, system architecture of the proposed method, initial data processing and descriptive statistics. Also Section 4 introduces multivariate regression and its math quotation. Section 5 reports the results of the multimodel analysis and emotion prediction based on decision fusion. Followed by discussion and conclusion sections respectively.

2. Related work

2.1. Quantitative assessment of environmental health impacts

Human exposure to environmental pollutants such as noise, air pollution, traffic or even crowded areas can cause severe health problems ranging from headaches and sleep disturbance and heart diseases [1,2].

The relationship between human body and the environmental factors has been extensively studied in social and environmental sciences, psychology and environmental health literature [3,14]. WHO, defines “Environmental Burden of Disease” [15] as one methodology for quantitatively assessing environmental health impacts at the population level in terms of deaths, Disability Adjusted Life Years (DALYs), or occasionally the number of cases. Other indirect measures can be used to estimate health impacts, for example the number of hospital admissions.

According to WHO quantitative assessments of health impacts are based on combining exposure data with exposure-response in-