Short Report on Kick Start Grant for “Social Sensing”.

Prof. Patrick G.T. Healey, Queen Mary University of London.
Dr. Hamed Haddadi, Imperial College London.

1 Project Summary:

There are many contexts in which it would be useful to have a better understanding of human interaction ‘in-the-wild’. In particular there is clear evidence that frequency and quality of social interaction are critical factors in determining physical and mental health outcomes (including a substantial impact on mortality (Landis-Holt, 2010). However current methods for assessing social engagement are coarse grained and rely heavily on subjective self-report. This project assessed the feasibility of developing unobtrusive, quantitative methods for capturing the frequency, quality and context of everyday social interactions. The aim was to identify new ways of enabling machines to perceive, recognise and engage with basic patterns of human interaction to enable more effective communication and collaboration.

The approach used is based on results obtained from work on optical motion capture of live conversation that people move in characteristic ways during face-to-face conversation (Battersby and Healey, 2010; Healey, Plant, Howes and Lavelle 2015). In particular, speaker’s hand movements during increase during conversation whereas their addressees move their hands significantly *less* than normal. This leads to the hypothesis that the frequency and degree of engagement interaction might have distinct motion signatures. If correct, this would provide a way to sense patterns of social interaction without requiring explicit self-report or potentially intrusive audio or video recordings. While a great deal of attention has been paid to sensing physical activity using motion sensors it has not been applied to capturing the quality of social activity in this way. For example, the Avon Longitudinal Study of Parents and Children (ALSPAC) and UK Biobank have wrist-worn accelerometer data but do not contain significant information on social interaction and have not been analysed to detect this (Willets et. al. 2018, Mattocks et. al. 2008).

2. Objectives:

The basic objectives of the project were to capture and analyse data from wrist-worn motion sensors and test whether a) whether or not people were interacting and b) how engaged they are in interaction could be determined from the motion data.

The target outputs of the project were:

1. Short conference paper on the findings of the pilot study submitted to ICMI.
2. Grant proposal to EPSRC on Social Sensing for Health and Wellbeing.
3. Modifications to the sensingkit software library for processing movement data.

3. Results:

Rather than collecting a new data set the data for the Social Sensing project was drawn from an existing corpus of unscripted social interactions collected at QMUL (Katevas, K. 2018) using mobile phones and wrist worn accelerometers (E4 Writsbands) captured from the onboard sensors using sensingkit (https://www.sensingkit.org). The accelerometer data had not been previously analysed. The corpus includes video recordings of the interactions without audio capture, sensor data from mobile phones and sensor data from wristbands. All research protocols were approved by the Queen Mary University of London ethical review board (reference: QMREC1705).

Video Coding:

The videos (which did not have audio recordings for privacy reasons) were initially analysed to identify all conversations and to code who was talking and who was listening. Videos were annotated using ELAN (https://tla.mpi.nl/tools/tla-tools/elan/). Participants were coded one at a time by the second author on a binary distinction: Speaking / Not Speaking with Speaking defined as visible mouth movements, ignoring backchannels (isolated utterances of less than 700ms) but allowing short pauses and including laughter. The fixed position of the video cameras combined with free movement of participants resulted in some occlusion of one or both participants in an interaction. Only interactions where both participants faces were visible are included in the analysis. This resulted in data for 14 participants in 7 speaker - listener pairs (6 males). Three participants were annotated twice in conversations with different people.

Following the intial data analysis (see below) a second pass coding was made to capture all walking sequences - typically only a few seconds long- for the participants in the first analysis. This yielded nine short walking sequences.

Accelerometer Data:

The E4 wristbands sample acceleration (ACC) in three dimensions at 32fps. The wristband provided only partial recordings for three participants who are excluded from the analysis.
The raw accelerometer data is converted to G-units by multiplying them with $xg = x^* 8/128$ (a value of $x = 16$ is in practice 1g). The magnitude of acceleration (r_i) is then calculated from the combined 3-axis acceleration data at each time point (i), by taking the square root of the sum of squared x, y and z values, leaving a single timeseries vector for each participant.

The calculated magnitude of the acceleration is converted to Mean Amplitude deviation (MAD) to give a measure of the intensity of the associated physical activity. This is a measure of dispersion similar to standard deviation (SD) but less influenced by extreme values. MAD was calculated for each participant's movements using a 5 second rolling window following the norm for physical activity monitoring (Matthews 2012). The resulting data plotted for fourteen participants is shown in Figure 1.

Engagement:

To obtain independent judgements of speaker's levels of engagement 28 clips (2 clips of each speaker) were extracted showing highest and lowest intensity hand movements in an +/- five second window for each speaker calculated from the accelerometer data. The clips were edited to be centered on the speaker only, showing both camera angles with no audio.

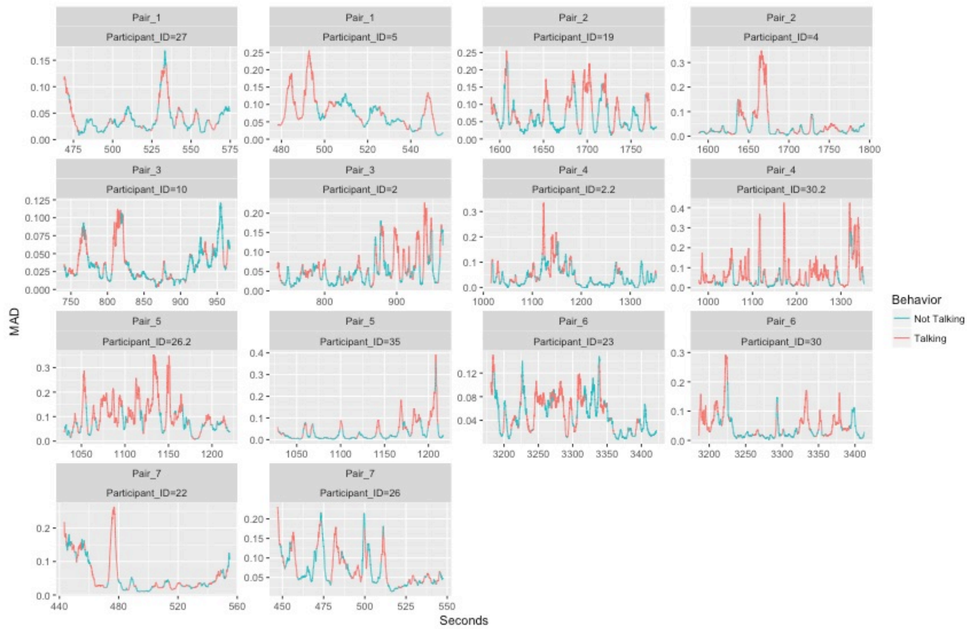

Figure 1: Mean Amplitude Deviation Over Time for Each Participant

Statistical Analysis:

GLMM analysis of the MAD data with participants as a random factor shows a reliable and large difference in the deviation scores for speakers and listeners (GLMM: Chi-sq = 33.06, $p < 0.001$). GLMM analysis of MAD scores for High vs. Low MAD as a predictor of estimated engagement also shows a large effect (Chi-sq = 136.89, $p < 0.001$). The means for each condition are shown in Figure2.

To test the (post hoc) question of whether walking was distinguishable from talking and listening a new set of data was identified by taking shorter, randomly selected sequences from the talking and listening data for the 9 participants who could be coded walking. The raw data is illustrated in Figure 3 and the means in Figure 4. A GLMM analysis with Participants and Behaviour (Walking/Talking/Listening) as random factors showed each behavior was discriminable. Tukey pairwise comparisons: Talking vs. Listening ($Z = 6.074$, $p < 0.001$) Walking vs. Listening ($Z = 4.294$, $p < 0.001$) Walking vs. Talking ($Z = 2.482$, $p < 0.015$).

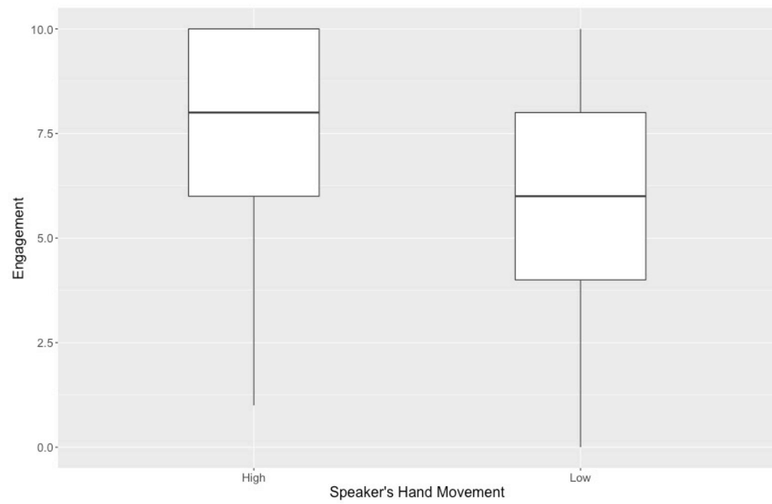

Figure 2: Estimated Engagement of Speakers with High and Low Hand Movement

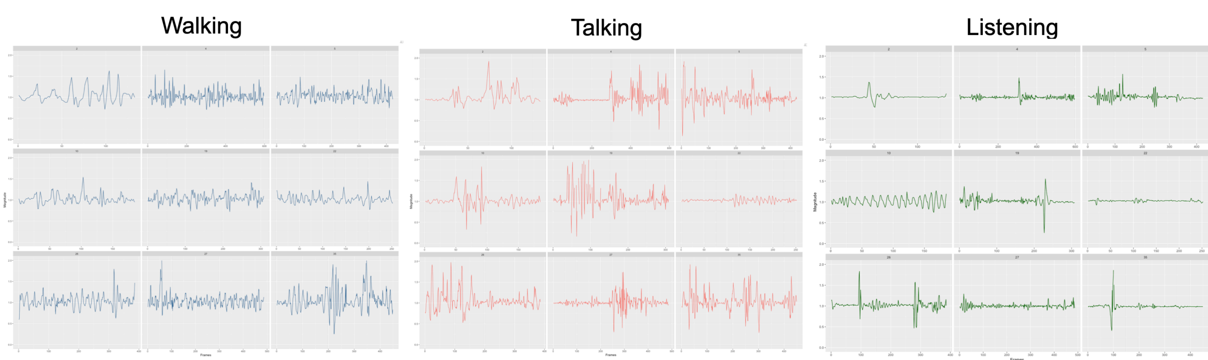

Figure 3: Accelerometer Data (MAD) for Nine Particpants (order preserved across panels) showing short segments of Walking, Talking and Listening.

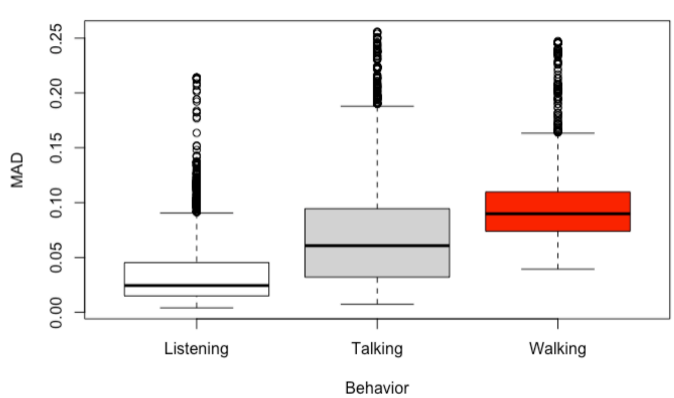

Figure 4: Mean Average Deviation: Listening, Walking, Talking.

Discussion:

The results of these analyses are clearly promising. Despite a very simple approach to the signal processing of the accelerometer data e.g. the analysis ignores differences in the relative contribution of the x,y and z axes and uses only a simple measure of energy in the signal (MAD) - there clear and large differences between speaking and listening and, on a smaller data set, speaking, listening and walking. The small number of participants and lab based, even if unscripted, interactions means there are questions about generalisation to larger or different demographic samples and also whether other daily activities, such as cooking or cleaning, would make interaction harder to distinguish.

Follow-ups:

During the project we contacted Aiden Doherty (Big Data Institure Oxford) who has a dataset from daily life that includes accelerometer data. We collaborated on a preliminary analysis which suggested that it ought to be possible to classify social activity in daily life from accelerometer data however there was insufficient time or resources to tune the machine learning approach.

A key follow on issue is whether people would find such movement monitoring either acceptable or useful. This was taken up with a PPI group (SUGAR group, City University, 25th of April 2019) who were asked to comment on what forms of feedback they might like. This was a very productive session which led to several ideas for visual representations, feedback and goal setting. A clear issue for all was that this information about social or physical activity levels should only be given as advice not instruction.

A second session looking at possible applications was conducted with the Newham Practice Members Council (a group of GPs) on 24th May, 2019 exploring the potential value of data sharing and mapping. The main issue raised was burden on GP's and the difficulty of actioning social activity data. The potential value for social prescribing was explored.

Follow-on Grant Applications:

Healey (PI) QMUL / Ove Arup Partners "Sensing Social Ecologies" Bid to Alan Turing Institute Urban Analytics Calls. 1 Apr 2020 - 30 Sep 2020. (£50k FEC) Unsuccessful.

Healey (PI) QMUL / UCL / CITY / Britsol. "Social Health" Bid to EPSRC Healthcare Techologies Call. Outline Stage. 1 Sep 2020 - 31 Aug 2025. (£7.3m FEC) Unsuccessful.

Conference Presentations:

Healey, P.G.T. Theodorou, L. and Haddadi, H. (2019) "Social Health: Mapping the quality of social interactions in the wild" Invited Talk, Human-Like Computing Machine Intelligence Workshop (MI21-HLC), 30th June – 3rd July Cumberland Lodge, Windsor, UK.

Healey P.G.T., Theodorou, L. and Haddadi, H. (2019) "The Dynamics of Hand Movements in Dialogue" 29th Meeting of the Society for Text and Discourse July 9th - July 11th, 2019 New York City, United States.

Publications:

Healey P.G.T. (forthcoming) “Human-Like Communication” in *Human-Like Machine Intelligence* edited by Stephen Muggleton and Nick Chater. Oxford University Press.

Healey, P.G.T., Theodorou, L., H¨ansel, K., Cavallaro, A. Tokarchuk, L., Haddadi, H. and Katevas, K. (in prep) “Hand Movements Signal Social Engagement”. For submission to Nature Human Behaviour.

4. Staffing:

Lida Theodoru was employed on the project from 1 Dec 2018 - 31 Jul 2019 as a 50% FTE research assistant (Grade 4, spine point 27). During this period she was awarded her PhD and now holds a post-doctoral research assistant job at City University.

5. References:

Battersby, S., & Healey, P. G. T. (2010). Head and hand movements in the orchestration of dialogue. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32, No. 32).

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. PLoS Med, 7(7), e1000316.

Healey, P. G. T., Plant, N. J., Howes, C., & Lavelle, M. (2015, March). When words fail: collaborative gestures during clarification dialogues. In *2015 AAAI Spring Symposium Series*.

Katevas, K., Haddadi, H., Tokarchuk, L., & Clegg, R. G. (2015). Walking in Sync: Two is Company, Three's a Crowd. In Proceedings of the 2nd workshop on Workshop on Physical Analytics (pp. 25-29). ACM

Katevas, K. (2018). Analysing Crowd Behaviours using Mobile Sensing (Doctoral dissertation, Queen Mary University of London).

Mattocks, C., Ness, A., Leary, S., Tilling, K., Blair, S. N., Shield, J., Deere, K., Saunders, J., Kirkby, J., Smith, G.D. and Wells, J. (2008). Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision. Journal of Physical Activity and Health, 5(s1), S98-S111.

Willetts, M., Hollowell, S., Aslett, L., Holmes, C., & Doherty, A. (2018). Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. Scientific reports, 8(1), 1-10.