Behavioural states of people with PIMD

Matej Cigale
Jožef Stefan Institute,
Jožef Stefan IPS
Jamova cesta 39, 1000 Ljubljana
matej.cigale@ijs.si

Mitja Lutrek
Jožef Stefan Institute,
Jožef Stefan IPS
Jamova cesta 39, 1000 Ljubljana
mitja.lustrek@ijs.si

Abstract

People with profound intellectual and multiple disabilities (PIMD) are an extremely vulnerable group facing great challenges in their daily lives. Due to their disabilities they are sometimes unable to communicate their desires to the outside world. This paper presents two specialised machine learning algorithms developed in Prolog that attempt to classify the behavioural states and communication attempts of people with PIMD based on non-verbal signals (NVS). The advantage of these methods is, that they can incorporate annotated data and expert knowledge from the caregivers.

1 Introduction

People with profound intellectual and multiple disabilities (PIMD) face extreme difficulties in everyday life. They are a heterogeneous group. Severe cognitive, motor and sensory disabilities makes this population reliant on outside care fo most daily tasks, and thus extremely vulnerable. The main problem for interaction with people with PIMD is lack of symbolic communication – they are unable to express their desires in a consistent manner. These individuals would benefit greatly from intelligent systems in their vicinity, they are unable to use these systems due to relative high complexity. The INSENSION project aims to develop a system that will observe behavioural state and non-symbolic communication attempts of people with PIMD and interpret them to people in the vicinity and even automatically control the environment using external services.

The first step of the INSENSION system is to recognise Non-Verbal Signals (NVS) expressed by people with PIMD (e.g., certain gestures [2] and facial expressions [7]) and important features of their environment (e.g., presence of caregivers and objects; temperature, luminosity). Afterwards, these are interpreted as behavioural states (pleasure, neutral, displeasure) and communication attempts (comment, demand, protest). This paper deals with the interpretation of NVS once they are recognised by the machine vision systems. The interpretation of NVS of people with PIMD is a challenging task, since each individual is unique with different abilities and signals, and there are no general signals that could be associated with behavioural states and communication attempts. Because of this, no general-purpose system can be developed and personalised classification methods must be used. Collecting a comprehensive set of annotated data for each individual is not feasible. Additionally, mappings between certain NVS and behavioural states are known to those close to the person, and this expert knowledge should be used in the decision making process. This requires specific methods that can take advantage of machine learning and expert knowledge to reach the decision.

In: A. Editor, B. Coeditor (eds.): Proceedings of the XYZ Workshop, Location, Country, DD-MMM-YYYY, published at http://ceur-ws.org

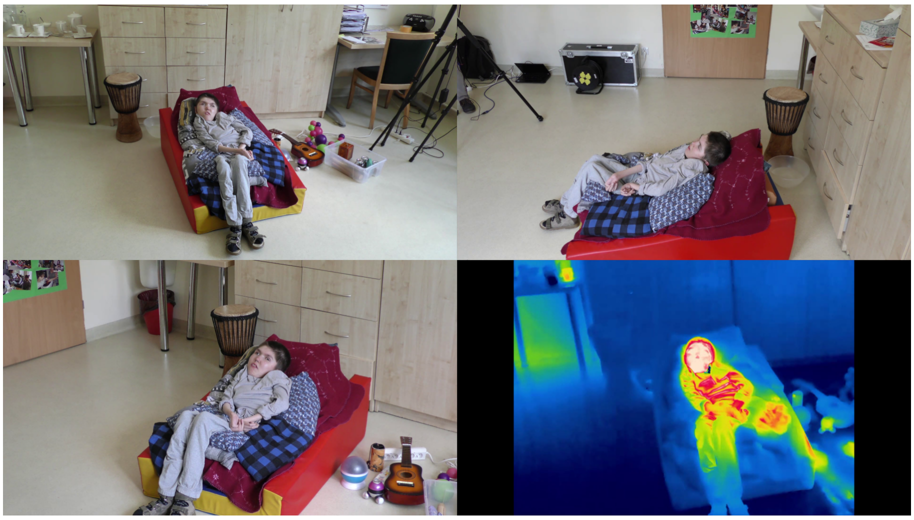


Figure 1: An example of the videos recorded.

A step towards understanding behavioural states closest to what INSENSION will use (from video and audio) was done by Metallinou et al. [4] using USC CreativeIT database consisting of acted-out scenes. Due to the specifics of our target population we cannot assume NVS will be consistent with the rest of the population. Behoora et al. [1] tackled a similar problem, focusing on real-life setting with designers in a team. They used an infrared imaging sensor (i.e., Microsoft Kinect) to extract the body positions, velocity and acceleration of all the joints in the upper part of the body. They used a static table to map the resulting body language poses into feelings. This approach is interesting and applicable to our project but would require creating a new expert system for each individual due to their limitations.

When it comes to enriching the communication with the context (data from IoT devices, objects and people in the vicinity) of the interaction there are several approaches that produce interesting results. Probabilistic Event Calculus [5] proposes a method that deals with uncertainty in the detected environment to classify the event that is occurring. They mostly deal with the movement of people in the system, but the method is flexible enough and could be adapted to our needs. Another field dealing with context of events is Case-Based Reasoning [3]. Here knowledge is represented as a set of cases - events that happened and the solutions that were used to solve the problem. Events that are detected are conformed into the closest case that is stored in the database and the solution of the problem is used. Both of these approaches can be integrated with our proposed solutions.

2 Behaviour state recognition

Our project currently works with five people with PIMD. Audio-Visual data was collected and took the form of multiple-angle recordings with normal and heat-vision cameras, an example of which can be seen in Figure 1. Expert knowledge was collected from their caregivers in the form of an extensive questionnaire. This data was then incorporated into the behavioural state recognition to improve decisions. Parts of these videos were annotated using the ELAN [6] software. Annotators were asked to annotate pre-defined facial expressions, gestures, vocalisations, presence of caregivers, and to note any special cases including ambient factors, such as music or light, that might play a role in the behavioural state of the subjects in each tree-second window. This is the data used to train our behavioural state classifiers. NVS recognisers will be developed to recognise NVS and context directly from video using computer vision. In addition to NVS, behavioural states of *pleasure* and *displeasure* were annotated, while *neutral or undefined* state was assumed to be any state that was not specifically marked. This simplification was used since recognising more subtle behavioural states of people with PIMD is extremely hard [8]. Nevertheless we feel that people tasked with annotation were familiar enough with their subjects so that they could render an accurate picture of their behavioural state.

Each detected signal can have a meaning, but in people with PIMD that is not guaranteed. The NVS can have no meaning or the same NVS is used to convey multiple dissimilar meanings, ie. the person with PIMD could clap to signify happiness, but also to signify sadness. These signals do not necessary follow social conventions.

The Unique Non Verbal Signals Method makes its decision based on the assumption that there exists a

```prolog
decide_state(NVS_Set, 'Pleasure') :-
    pleasure_marker(Pleasure),
    member(Pleasure, NVS_Set).
decide_state(NVS_Set, 'Displeasure') :-
    displeasure_marker(Displeasure),
    member(Displeasure, NVS_Set).

pleasure_marker(NVS) :-
    assessment('Pleasure', NVS).
pleasure_marker(NVS) :-
    window(Interval, NVS, _),
    window(Interval, 'Pleasure', _),
    not(displeasure_marker(NVS)),
    not(neutral_marker(NVS)).
```

(a) Querying the behavioural state.

```prolog
behaviour_state(NVS_Set, Decision, P_Cut, D_Cut) :-
    calculate_valence(NVS_Set, Valence),
        (Valence > P_Cut ->
            (Decision = 'Pleasure');
        (Valence < D_Cut ->
            (Decision = 'Displeasure');
            (Decision = 'Neutral'))).
```

(b) The valence method.

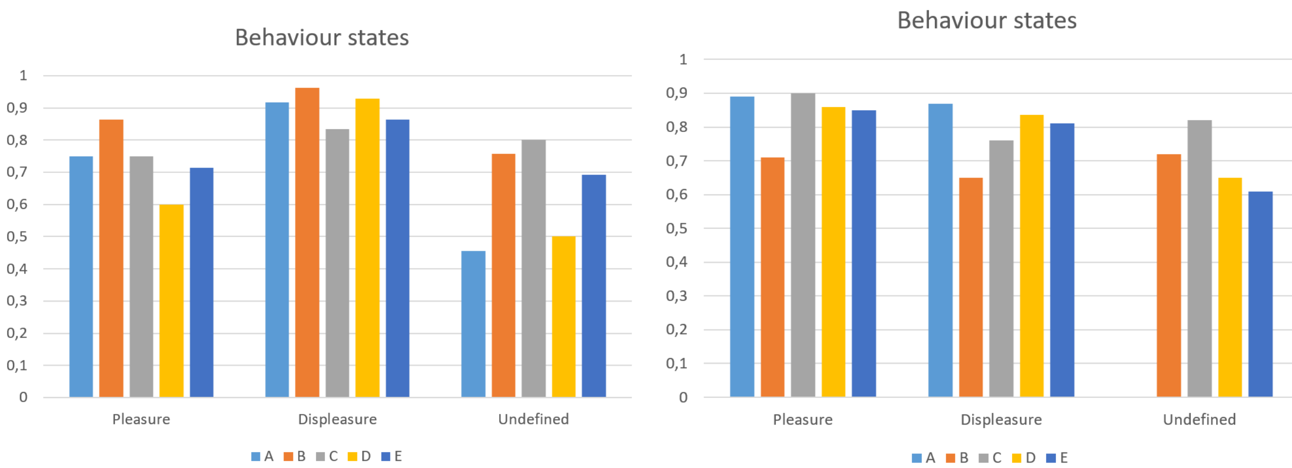
Figure 2: The relevant Prolog terms

set of NVS that signify - for instance - pleasure, but will never be used to signify any other behavioural state. In order for us to detect pleasure, we must from a set of all detectable NVS associated with pleasure, remove all NVS, that are associated with displeasure or neutral state. Additionally if experts decide that a NVS means pleasure we take this into account. Similar logic follows for displeasure and neutral state. Deciding on the behavioural state based on the set of NVS is simple (Listing 2a). We check if there are any NVS that are specific to pleasure, either from the expert knowledge or from the annotated examples. The term *assessment('Pleasure', NVS)* will return a NVS that was annotated as denoting pleasure by the experts. The second part of *pleasure_marker* term will check if there is a NVS in the annotations associated with pleasure and not with displeasure or neutral state. Similar logic is used to calculate displeasure_marker and neutral_marker. The term *window(Interval, NVS, Annotation)* unifies for any three-second window that is annotated with a given NVS and Annotation. If either of these rules holds true we identify the behavioural state as pleasure. This system works well on the available dataset. The results of this can be seen in Figure 3a. This approach is expected to become less viable with more data. Due to uniqueness of the people with PIMD a model is trained for each individual.

The Valence Method treats the significance of a NVS as an indicator of behavioural state on a continuous interval. We assume that each NVS has a certain correlation with emotional valence. In our case valence is a number that is correlated with the three behavioural states *(displeasure, neutral, pleasure)*, a simplified case of mapping feelings to the Arousal-Valence space. Valence is a value in [-1, 1] interval where displeasure is associated with negative and pleasure with positive numbers. Listing 2b contains the code for valence calculation. If there is little or no correlation between pleasure and the NVS it should gravitate towards negative values. Inverse must be true for displeasure. The *correlation_strength(NVS1, NVS2, Num_correlations)* returns the number of all annotated windows that contain NVS1 at the same time as NVS2. If we want to classify the behavioural state based on a NVS we add the valence of all the NVS that are expressed. We determine the behavioural state based on the value of valence. The *calculate_valence* is a recursive function that sums the valence of a set of NVS, and returns 0 for an empty set. The *P_Cut* and *D_Cut* variables determine the intervals of pleasure, displeasure or neutral behavioural state. We use Constraint Logic Programming to determine the optimal values for these values. At its core this is a minimisation problem where we try to find the thresholds for the intervals that produce the smallest classification error. The rationale for the system is as follows. We take all the windows in the annotations we have and attempt to find values where we cut the valence dimension so that our classification error is the smallest possible. This method performs worse then the somewhat naive Unique non Verbal Signals method, as seen in Figure 3b. Person A has very high miss classification of neutral state, owing to a small example size for this state. The Valence method seems to perform better for subjects with more annotations since more data leads to better valence score estimation. The Unique non Verbal Signals method with infinite data converges toward expert knowledge while The Valence method diverges from it.

3 Conclusions

In this paper we presented two machine learning algorithms, specialised for learning behavioural states of people with PIMD. The advantage over the more common algorithms is the ability to incorporate prior knowledge from



(a) Unique Non Verbal Signals method.

(b) The valence method

Figure 3: Model classification accuracy. The letters identify the trial participants.

the assessments. This is important as people with PIMD exhibit little to no standardised expressions due to mental or physical disabilities. The work at this time assumes that the NVS are robustly detected. While the Valence method is easily adapted to the non deterministic nature of the recognisers the Unique Non Verbal Signals method will require additional handling of probability logic to cope with the expected data.

4 Acknowledgements

This work was supported by the European Union's Horizon 2020 research and innovation program [grant agreement No 780819 (INSENSION)

References

[1] Behoora, I., Tucker, C.S.: Machine learning classification of design team members' body language patterns for real time emotional state detection. Design Studies **39**, 100–127 (jul 2015)

[2] Cao, Z., Simon, T., Wei, S.E., Sheikh, Y.: Realtime multi-person 2D pose estimation using part affinity fields. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 pp. 1302–1310 (2017)

[3] Lupiani, E., Juarez, J.M., Palma, J., Marin, R.: Knowledge-Base d Systems Monitoring elderly people at home with temporal Case-Based Reasoning. Knowledge-Based Systems **134**, 116–134 (2017)

[4] Metallinou, A., Katsamanis, A., Narayanan, S.: Tracking continuous emotional trends of participants during affective dyadic interactions using body language and speech information. IMAVIS **31**(2), 137–152 (2013)

[5] Skarlatidis, A., Paliouras, G., Artikis, A.: Probabilistic Event Calculus for Event Recognition (September 2015) (2014)

[6] Sloetjes, H., Wittenburg, P.: Annotation by category - ELAN and ISO DCR. Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08) pp. 816–820 (2008)

[7] Sun, X., Lv, M., Quan, C., Ren, F.: Improved facial expression recognition method based on ROI deep convolutional neutral network. 2017 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017 **2018-Janua**, 256–261 (2018)

[8] Vos, P., De Cock, P., Petry, K., Van Den Noortgate, W., Maes, B.: What makes them feel like they do? investigating the subjective well-being in people with severe and profound disabilities. Research in developmental disabilities **31**(6), 1623–1632 (2010)