Extracting automatically social signals from psychotherapy sessions
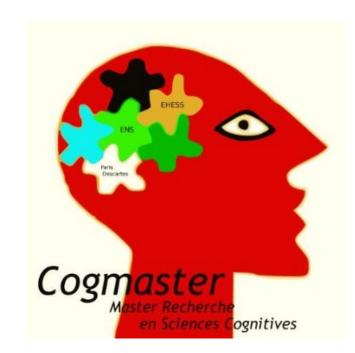
Thomas Gargot[1,2,*], Giovana Varni[2], Jonathan Aigrain[2],
Mohamed Chetouani[2], Michel Spodenkiewicz[1,2], David Cohen[1,2]
1. Hôpital de la Pitié Salpêtrière, 2. Institut des Systèmes Intelligents et de la Robotique, Paris, France
* thomas_gargot@hotmail.com

Introduction and motivation

Psychotherapy is an important part of treatment of mental disorders alone or complementary with pharmacotherapy. Somes techniques are now widely evidence-based and very cost effective (Layard & Clark, 2014).

Most of studies are indirectly based on patient reported outcomes or problematic behaviors that are evaluated before and after the psychotherapy. Unfortunatelly, studies hardly control what is directly happening during psychotherapy especially the interaction between the patient and the therapist that could be a predictor of the psychotherapy efficacy. Consequently, it is difficult to make precise links between theory and practice, control its application and to understand which of their ingredients are the most importants.

It is yet possible to annotate manually videos. However, this task is challenging since it can be either very repetitive (annotation of turn-taking or non-verbal behavior) or very technical (annotation of application of some specific techniques like in motivational interviewing (Moyers et al, 2015).

In the future, an automatic feedback of the psychotherapy could also even help a therapist to guide them to reorganize a treatment.

Project website

Towards an automated research framework

Here we suggest a research framework to extract automatically social signals from psychotherapy videos. We focused on motor synchrony since, it was found to be a predictor of psychotherapy outcome in a first study (Ramseyer, 2011). Varni et al. 2015 developed the SyncPy open source python library to help researchers and practitioners to automatically analyse synchrony. It could be possible to measure synchrony even in familial therapies. Other features could be quite easilly extracted manually like direction of the body, gaze direction or smiles.

Using an interdiciplinary and open science approach, we developped some modules toward this goal that could be freely and easily re-used by other teams in other databases or with other modules.

Future work

The next step would be to design a specific database anticipating some technicals problems (speech segmentation, overlapping of subjects, micro-movements) and define more precisely different conditions of psychotherapy that could be contrasted.

Flowchart of the processing pipeline including: Video collection (245 videos, 252.44 GB) (VLC), Video selection, Conversion video from VOB to AVI (Open CV Python), Frame Extraction (Open CV Python), Participant zone segmentation (Paint Brush), Cut non relevant big movements (Quicktime), Machine learning upper body detection, Mean motion History by minute (R), Background Extraction, Motion History Extraction (Open CV C++), Filtering (R Sliding Interval), Normalization (R log), Synchrony computing with Syncpy (SSI) (Python SyncPy Call S_Estimator), Speech Extraction (Elan), Clinic qualitative assessment, Psychometry collection (Excel), Importation in CSV and Graphic Representation (R). Other automatic features: Motion direction detection, Gaze direction detection, Emotions detection, Smiles detection. Features to extract during psychotherapy: Operationalized psychotherapy, Relationship evaluation, Physiological markers, Speech segmentation.

Legend: Raw data, Final data compared, Manual Step, Automatic step, Improvement suggestions, Laborious work.

No potential conflicts of interests