3D cameras in the service of psychiatry and neurology

Daphna Weinshall

Hebrew University

Our research

- **Problem**: most psychiatric and neurological disorders are diagnosed in a descriptive non-quantitative manner.

  There is an urgent need for objective ways to measure such disorders.

- **Goal**: develop automatic tools for the characterization of non-verbal behavior of psychiatric patients (Schizophrenia) and Parkinson patients.

ICRI-CI
Intel Collaborative Research Institute
Computational Intelligence

Non-verbal cues

- *Speech prosody*: schizophrenic patients may have “indifferent tone”, depressed patients may show small variability
- *Facial expressions*: schizophrenic patients may show “flat affect”, with often inappropriate facial expressions
- *Bodily gestures*: various aspects of Parkinson disease are judged based on the patient’s movement

Speech prosody and mental illnesses

- **Goal**: use acoustic properties of one's speech to 'look into the soul' and actually measure psychiatric disorders
- Domain: schizophrenia, acute depression
- Application example: patient monitoring over the phone

joint work with Roi Kliper and Shirley Portuguese

Acoustic characteristics of the speech signal

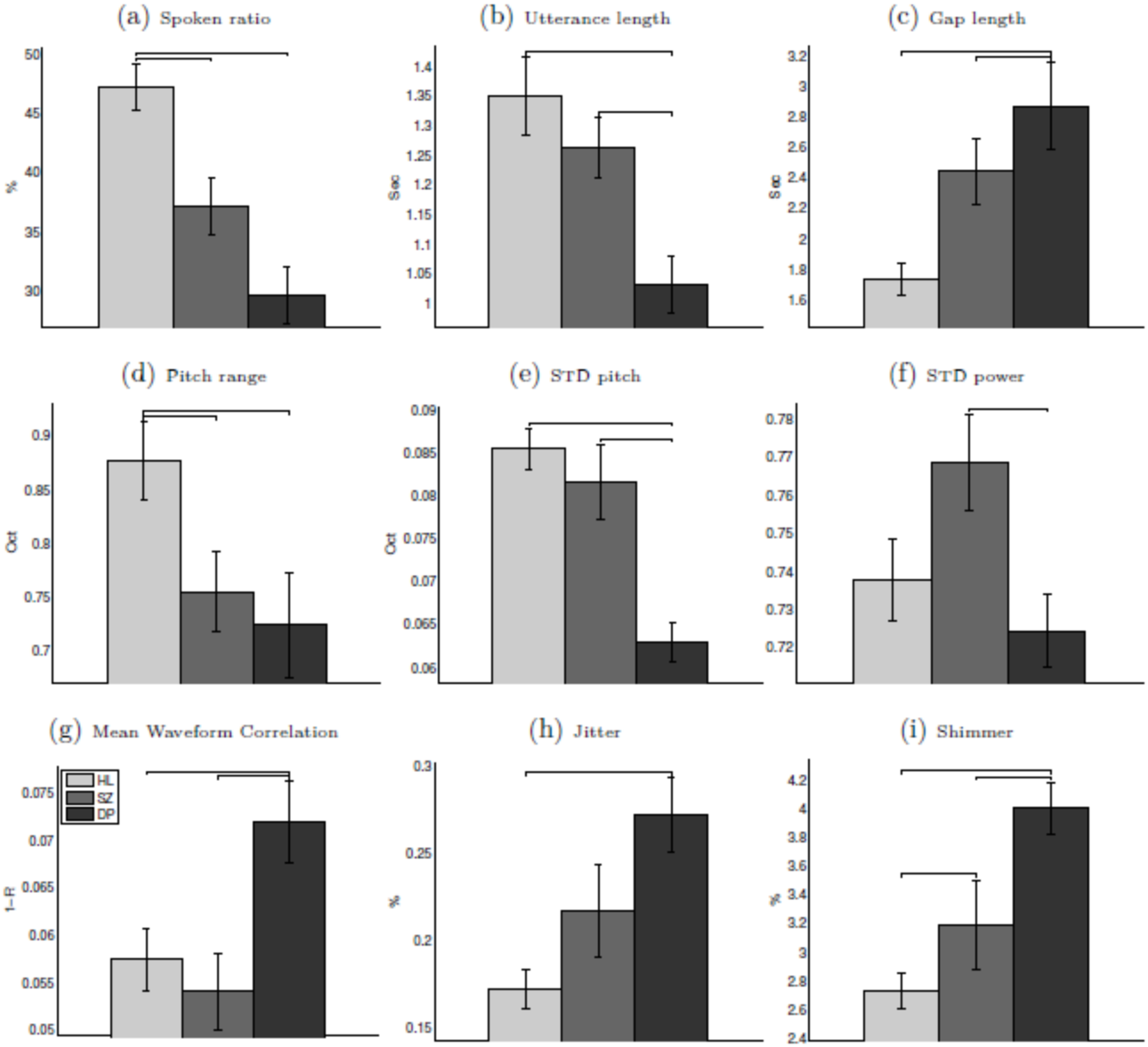
- **Macro scale**: mean utterance duration, mean gap duration, mean spoken ratio
- **Meso scale**: pitch range, pitch standard deviation, power standard deviation
- **Micro scale**: mean waveform correlation, mean jitter, mean shimmer

Data

- 3 groups: schizophrenic patients, people with acute depression, and normal controls

- Experimental conditions:
  - Free interview with psychiatrist
  - Readings of various standard texts

Results


(a) Spoken ratio
%
(b) Utterance length
Sec
(c) Gap length
Sec
(d) Pitch range
Oct
(e) STD pitch
Oct
(f) STD power
(g) Mean Waveform Correlation
1−R
HL
SZ
DP
(h) Jitter
%
(i) Shimmer
%

Visual non-verbal cues

- **Facial expressions and bodily gestures:** very informative about a variety of mental and neurological disorders, but how can we measure them reliably?

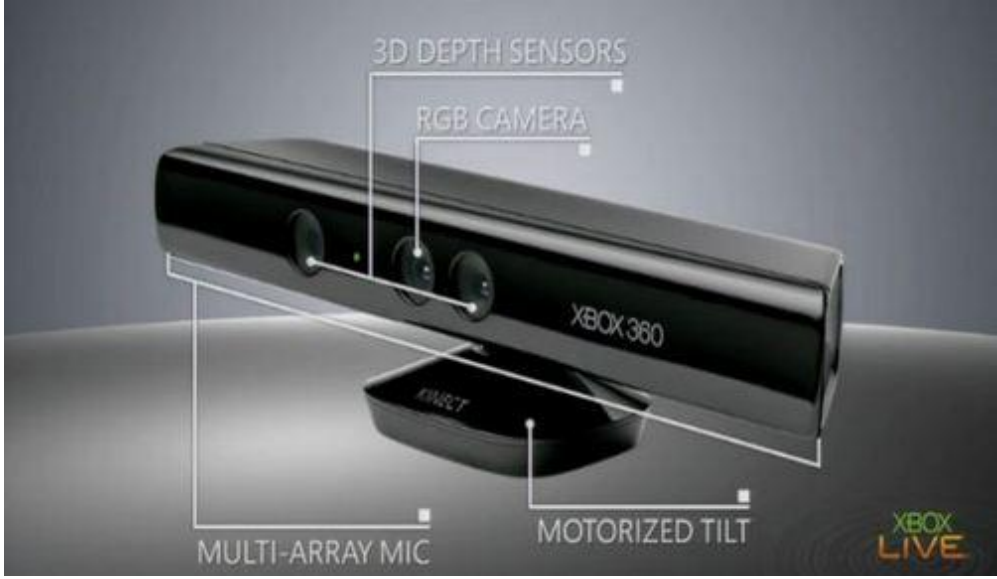
- **Our current research**: extract this information from 3D depth cameras that have become available in recent years

Two posters next door (preliminary results)

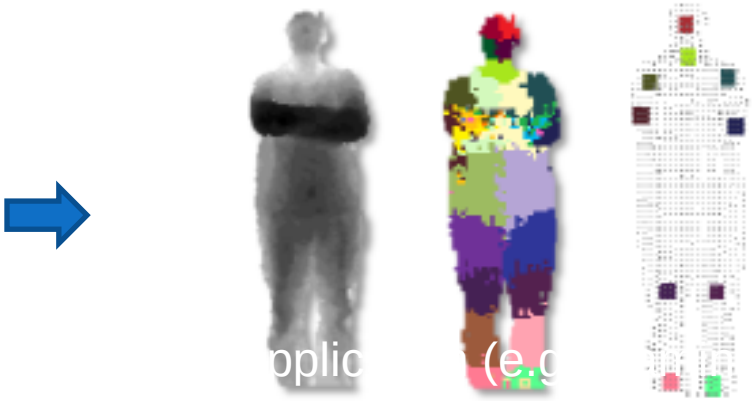
- Schizophrenia:
  - Tali Tron
  - Dr. Avi Peled (Shaar Menashe mental hospital)

- Parkinson:
  - Maria Dyshel
  - Prof. Hagai Bergman (Hadassah & HUJI)

3D Cameras



Get Depth Image

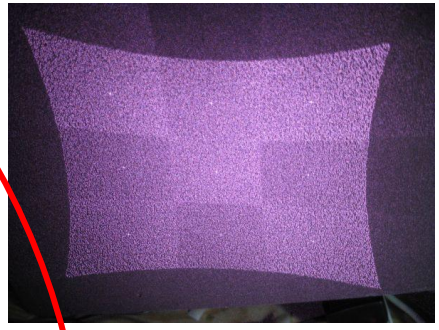


Estimate Body Pose

How does it work: overview

IR Projector

IR Sensor

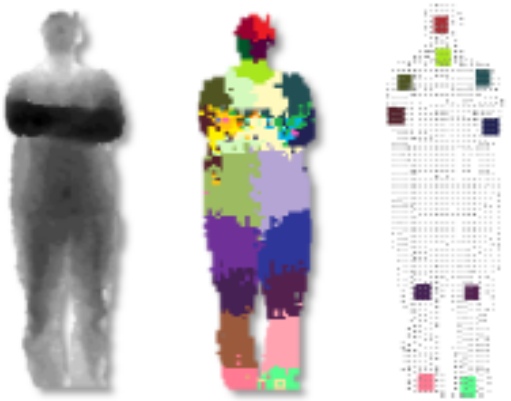


Projected psuedo-random light pattern

Stereo Algorithm

Depth Image

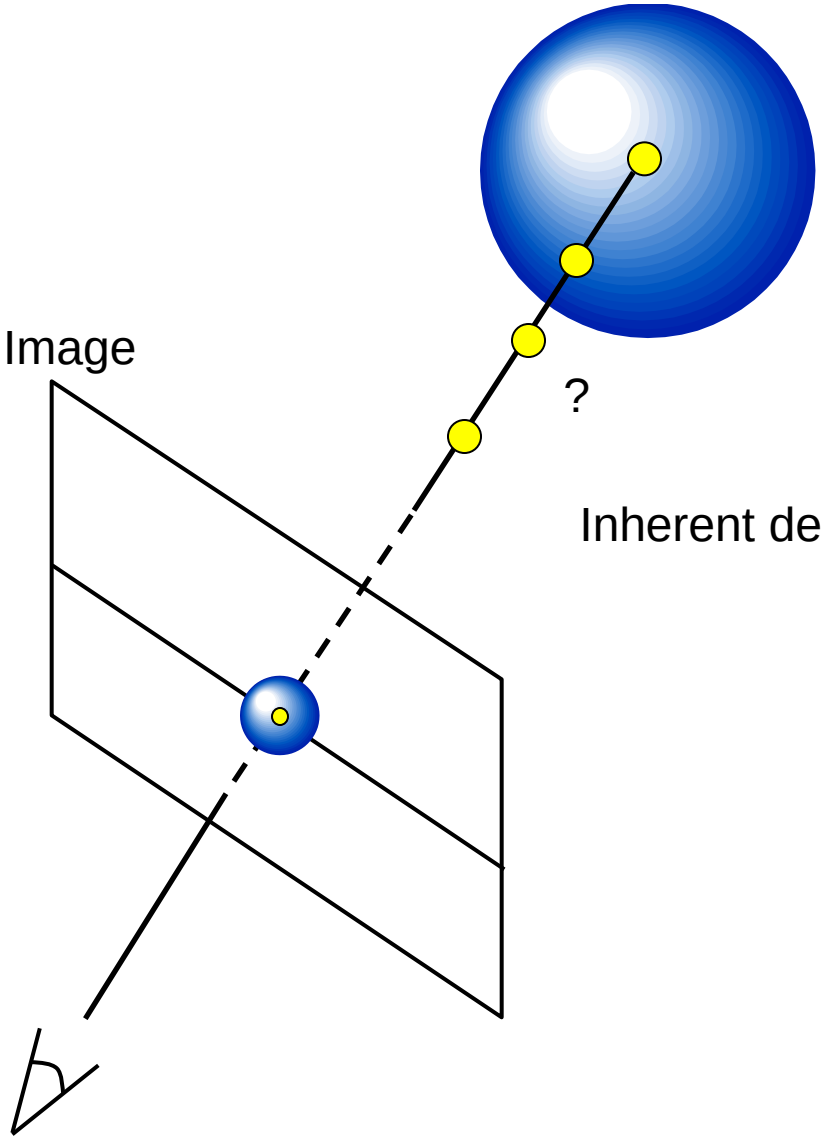
Segmentation, Part Prediction



Body Pose

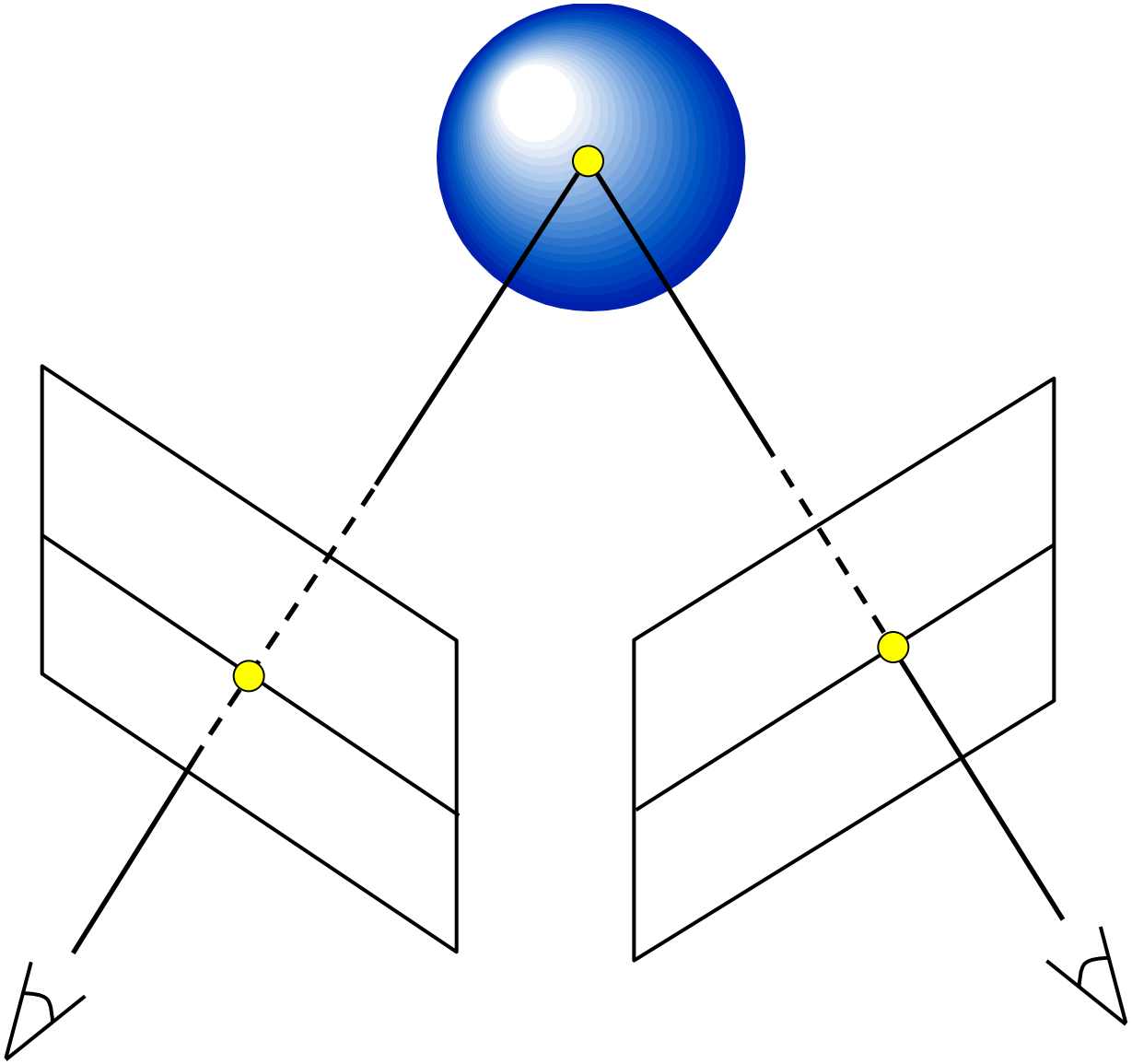
ICRI-CI
Intel Collaborative Research Institute Computational Intelligence

Digression – stereo vision



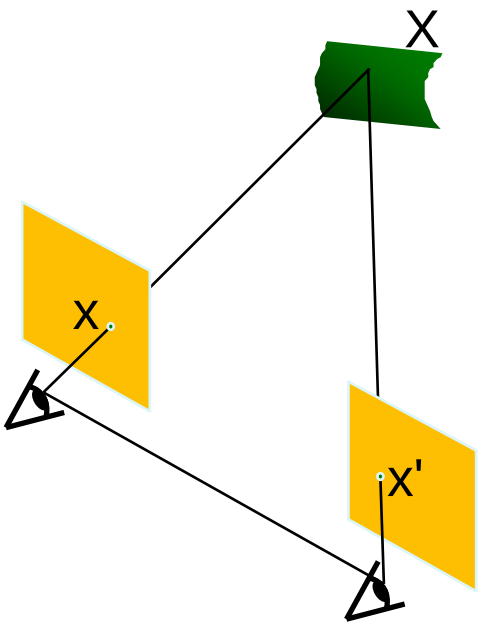
Inherent depth ambiguity in all images

Solution – Stereo Triangulation



Depth from Stereo

- Goal: recover depth by finding image coordinate x’ that corresponds to x
- Sub-Problems
  1. Calibration: How do we recover the relation of the cameras (if not already known)?
  2. Correspondence: How do we search for the matching point x’?



Correspondence Problem

How do we match a point in the first image to a point in the second?

Correspondence is hard !

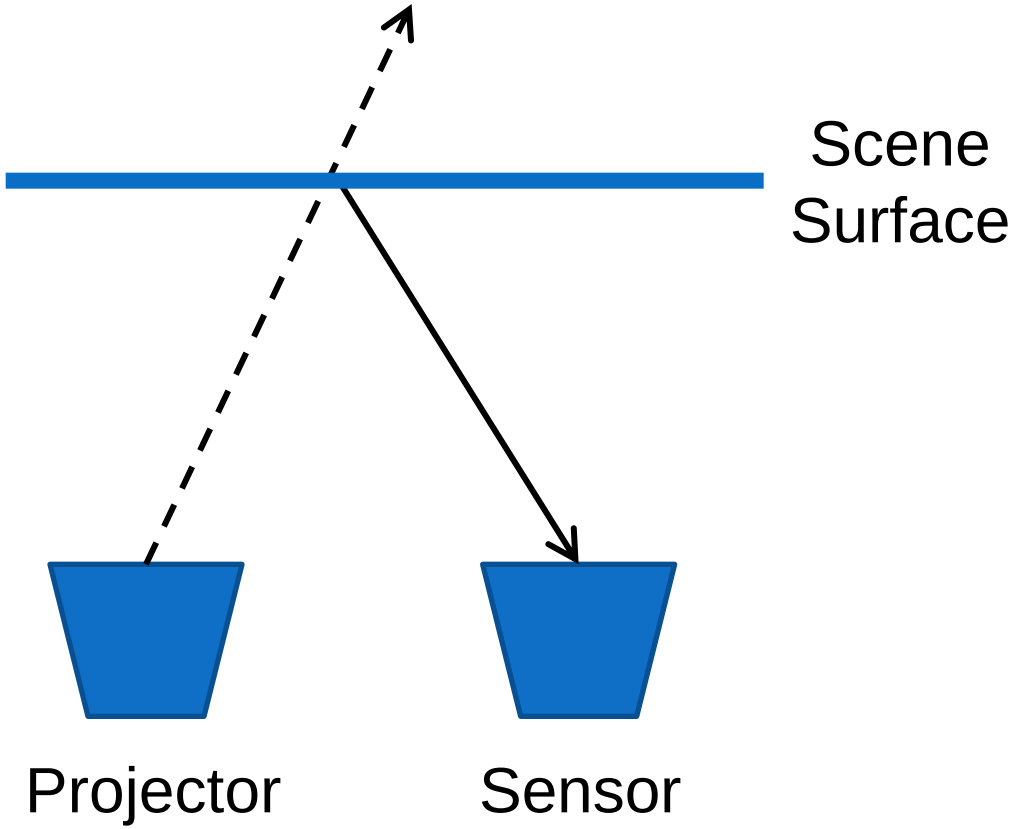
Solution - Structured Light Reconstruction

- Basic idea: project a known light pattern on the scene
- Obtain an effective stereo pair from the recorded image and the projected pattern

- Avoid problems due to correspondence
- Avoid problems due to surface appearance
- much more accurate

Depth from Projector-Sensor

Only one image: How is it possible to get depth?



Application 1: schizophrenia
with Talia Tron and Dr. Avi Peled

- Non-verbal observations such as facial expressions, voice and bodily expressions are essential for the diagnosis of mental disorders
- Our goal:
  - **quantitatively** characterize **non-verbal** behavior in psychiatric patients
  - develop **automatic** tools for describing and analyzing clinically relevant measures of this illness

Extracted features

We use Faceshift™, which returns 4 output types

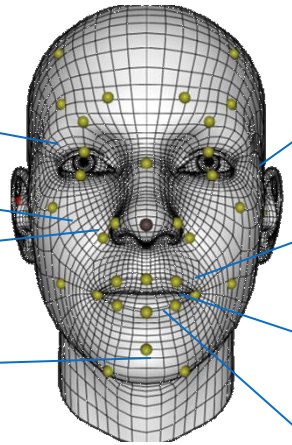
- Intensity level of 49 AUs:

BROWS (up, down)

CHEEK (squint, puff)

NOSE (sneer)

CHIN (raise)



EYES (blink, squint, up, down, in, out)

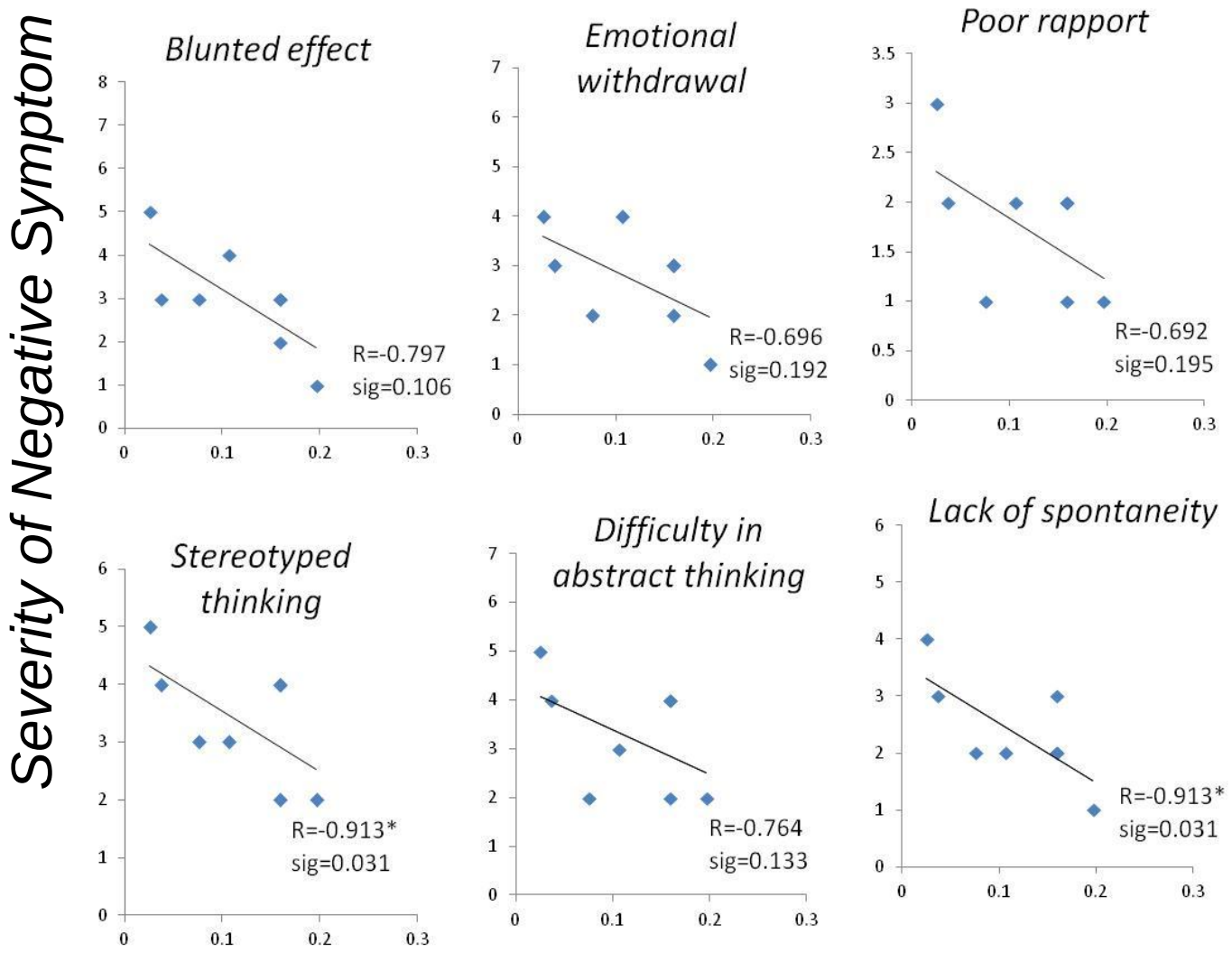
LIPS (stretch, close, open, up, down, funnel, pucker)

MOUTH (left, right, frown, smile, dimple)

JAW (forward, left, right, open)

- Eye gaze and position
- 3D head coordinates
- 3D position of facial markers

Correlation between measurements and symptoms



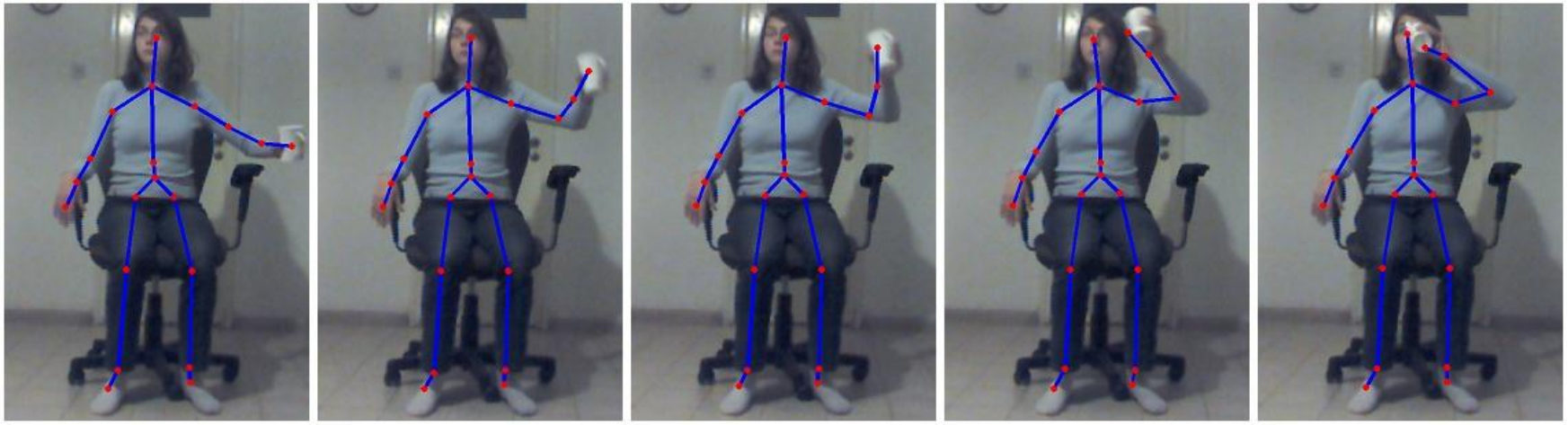
Application 2: Parkinson
with Maria Dyshel & Prof. Hagai Bergman

- Involuntary muscle movement, or dyskinesia, is a debilitating complication of chronic levodopa therapy of Parkinson's disease

- **Our goal:** develop an objective and automatic procedure to assess the severity of dyskinesia using relatively cheap and common devices

Extracted features

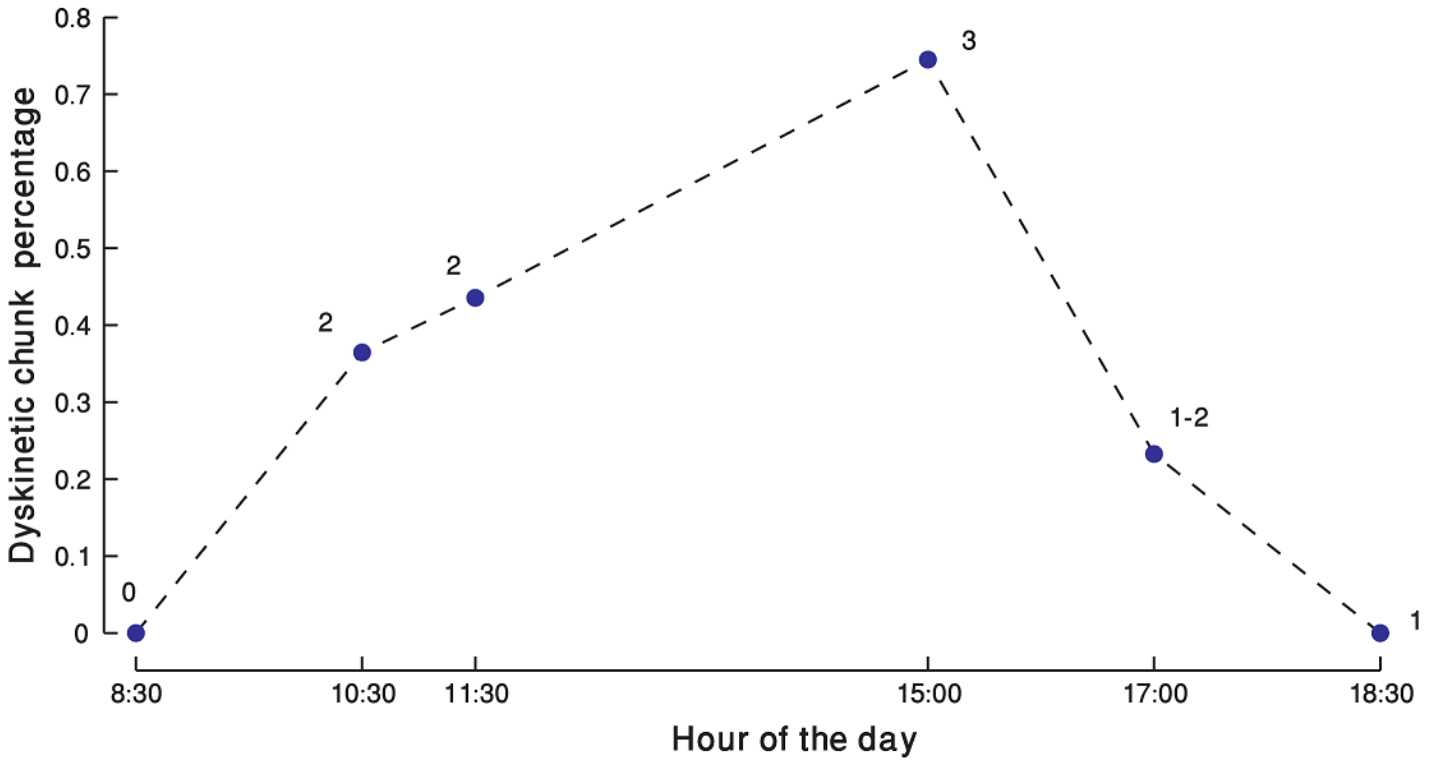
- Mircrosoft Kinect™ and its skeleton tracking interface is used to record and analyze the patients' movements:
- 20 human joints are tracked



Scoring Algorithm

- Recording is segmented into chunks of 1-3 seconds long
- Each chunk is represented by the averages motion of 10 joints
- We train a classifier to classify each chunk as dyskinetic (or not)
- Each video clip is characterized by the percent of dyskinetic chunks

Monitoring the state of a single patient



Technical challenges

- Improved camera technology (better 3D resolution, better depth range)
- Improved software for motion tracking and 3D feature extraction (e.g., better facial Action Units, accurate feet motion)
- Big data acquisition and analysis