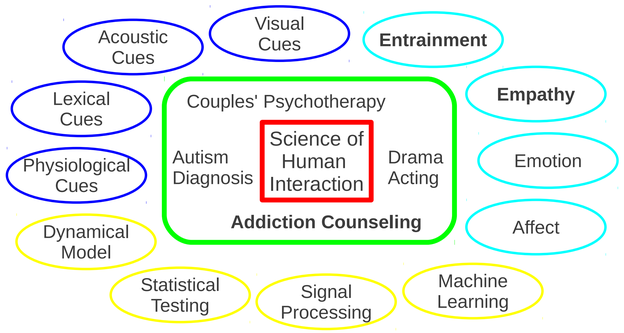
Modeling Therapist Empathy and Vocal Entrainment in Drug Addiction Counseling

Bo Xiao[1], Panayiotis G. Georgiou[1], Zac E. Imel[2], David C. Atkins[3], Shrikanth S. Narayanan[1]

[1]SAIL, Dept. Electrical Engineering, University of Southern California [2]Dept. Educational Psychology, University of Utah
[3]Dept. Psychiatry & Behavioral Sciences, University of Washington, U.S.A.
boxiao@usc.edu, georgiou@sipi.usc.edu, zac.imel@utah.edu, datkins@u.washington.edu, shri@sipi.usc.edu

Behavioral Signal Processing



S. Narayanan and P. Georgiou, "Behavioral signal processing: Deriving human behavioral informatics from speech and language," *Proceeding of IEEE*, 101(5): 1203 - 1233, 2013.

This work is supported by NSF, NIH and DoD.

Computing Entrainment and Empathy

- **Empathy** — feeling for and taking the perspective of others
- Psychological process evident across human and animal
- Rating associated with positive outcome of interactions
- Key performance evaluation index in MI-type psycho-therapy

- **Entrainment** — behaviors of the interactants become similar
- Reflected in multimodal cues, a way to infer empathy level

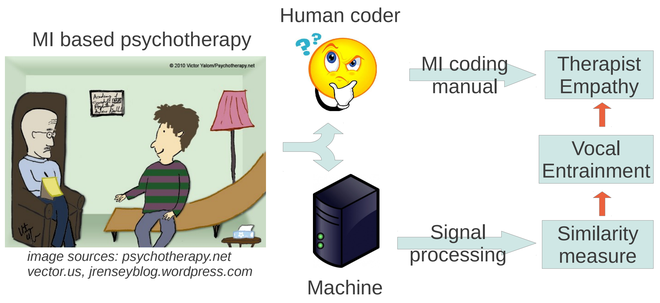
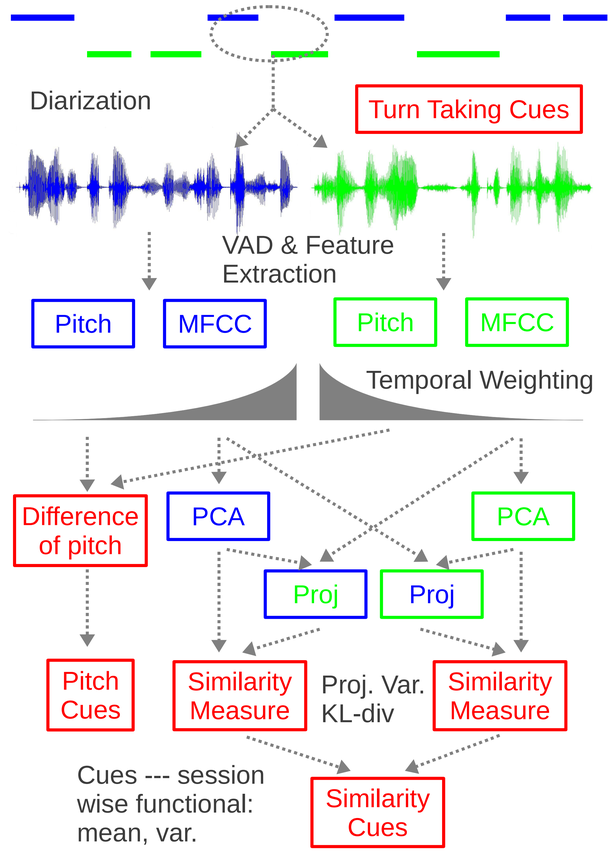


image sources: psychotherapy.net vector.us, jrenseyblog.wordpress.com

Dataset

- Counselor training study of Motivational Interviewing (MI)
- MI: emphasize intrinsic motivation of changing addiction
- Three coders evaluated 144 therapists' empathy score
- Select 72 high & 42 low scored sessions, 20 min each

Acoustic and Turn Taking Features



- Turn Taking cues: ratio of patient segments / speaking time
- Variations of acoustic cues: speaking turn order, PCA space extraction and projection order, number of selected PCA components (J), weighting factor (θ), temporal functionals

Functionals:
Mean (*M*), Variance (*V*)
Similarity measure (*sim*)
Pitch difference (*pit*)

$M_{sim}(PT, PT)$

Turn order, Patient (P) Therapist (T)
T projects to P's PCA space

- **Fusion**: Find PCA space of all features, select the single component that maximizes correlation with empathy score

C.C.Lee *et. al.*, "Computing Vocal Entrainment: A Signal-derived PCA-based Quantification Scheme with Application to Affect Analysis in Married Couple Interactions," *Computer Speech & Language*, 2012.

Experiment

Correlation of features and empathy score

θ^{3000}	1	10^{-1}	10^{-2}	10^{-3}
$M_{sim}(PT, PT)$	-0.24	-0.25	-0.27	-0.29
$V_{sim}(PT, PT)^*$	-0.22	-0.23	-0.21	-0.19
$M_{pit}(PT)$	-0.31	-0.32	-0.31	-0.30
$V_{pit}(PT)$	-0.32	-0.34	-0.35	-0.34
Ratio of patient speaking time				0.27
Ratio of patient speaking segments				0.28

* p<0.05, otherwise p<0.01

Performance of integrated feature

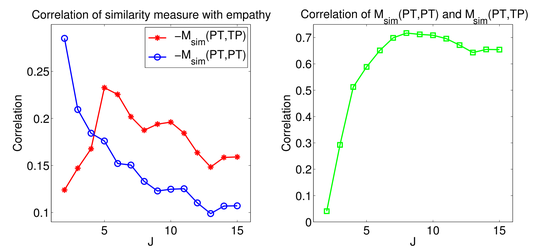
Acc. — logistic reg. of binarized empathy

Training		Testing	
Corr.	Acc.	Corr.	Acc.
0.47 $\pm$ 0.01	0.71 $\pm$ 0.01	0.43	0.70*

$*$ p = 0.02 in binomial test, baseline = 0.61

Discussion

- Projecting T to P and $J = 2$ more effective than others
- The two features are closer in value with larger J



- Different turn orders lead to correlated feature values
- "T following P" correlates with empathy, but not the other

θ^{3000} $(J = 2)$	1	10^{-3}
Corr. $M_{\text{sim}}(PT, PT)$ & $M_{\text{sim}}(TP, PT)$	0.85	0.66
Corr. $M_{\text{pit}}(PT)$ & $M_{\text{pit}}(TP)$	0.71	0.53

Conclusion

- Extend similarity measure for vocal entrainment behavior
- Provide computational model to infer therapist empathy
- In future, more effective features on vocal similarity
- Understand the asymmetry of turn order and PCA projection