Dynamic Sound Fields Clusterization using Neuro-Fuzzy Approach

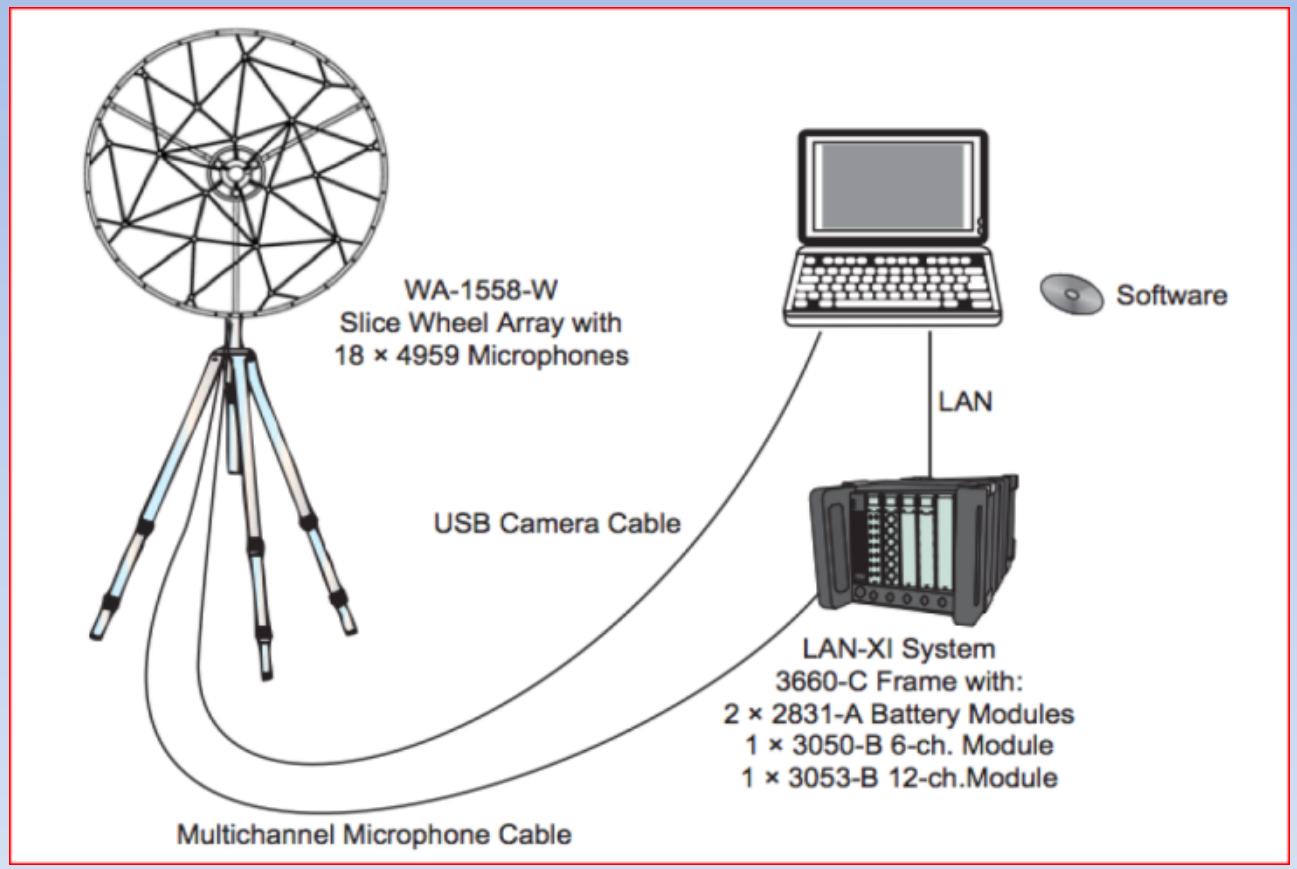
Petia Koprinkova-Hristova and Kiril Alexiev

Institute of Information and Communication Technologies

Presentation schedule

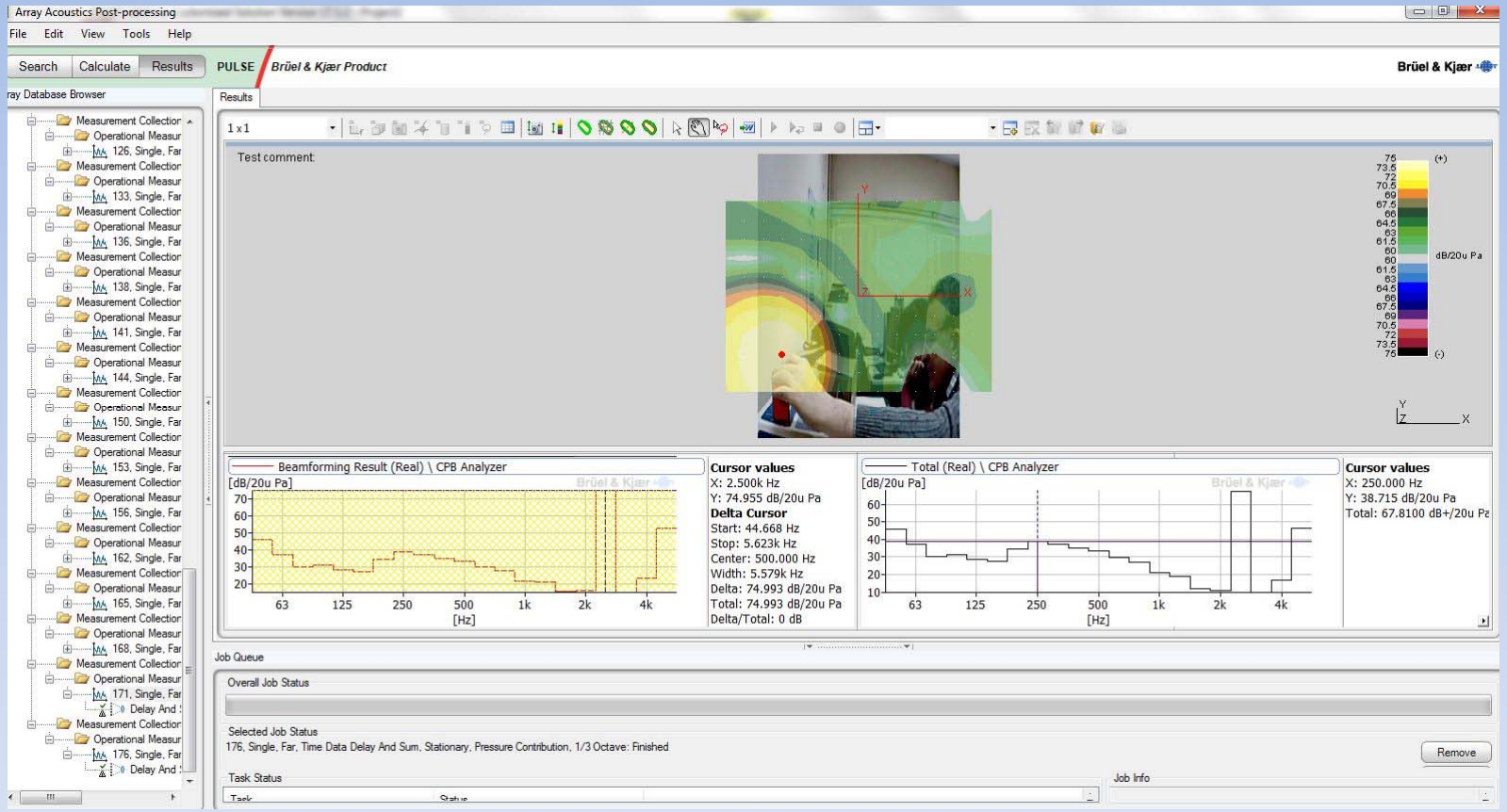
- Experimental set-up
- Proposed algorithm
- Receptive fields for signal preprocessing
- Echo state networks
- IP improvement of reservoir and its exploitation for clustering
- Results and discussion

Experimental set-up



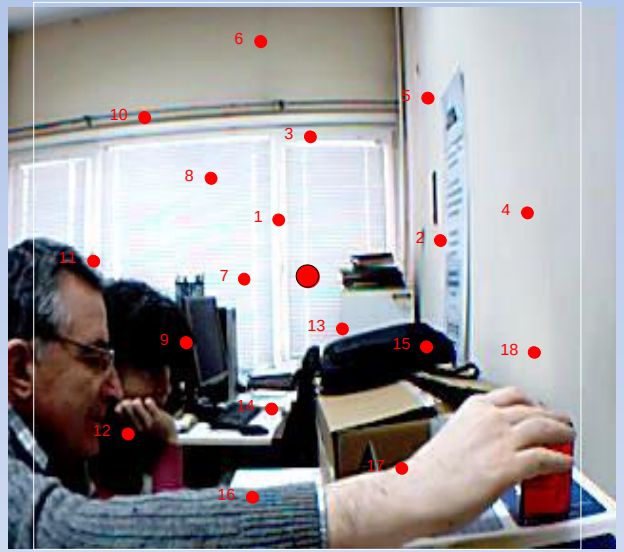
Experimental set-up

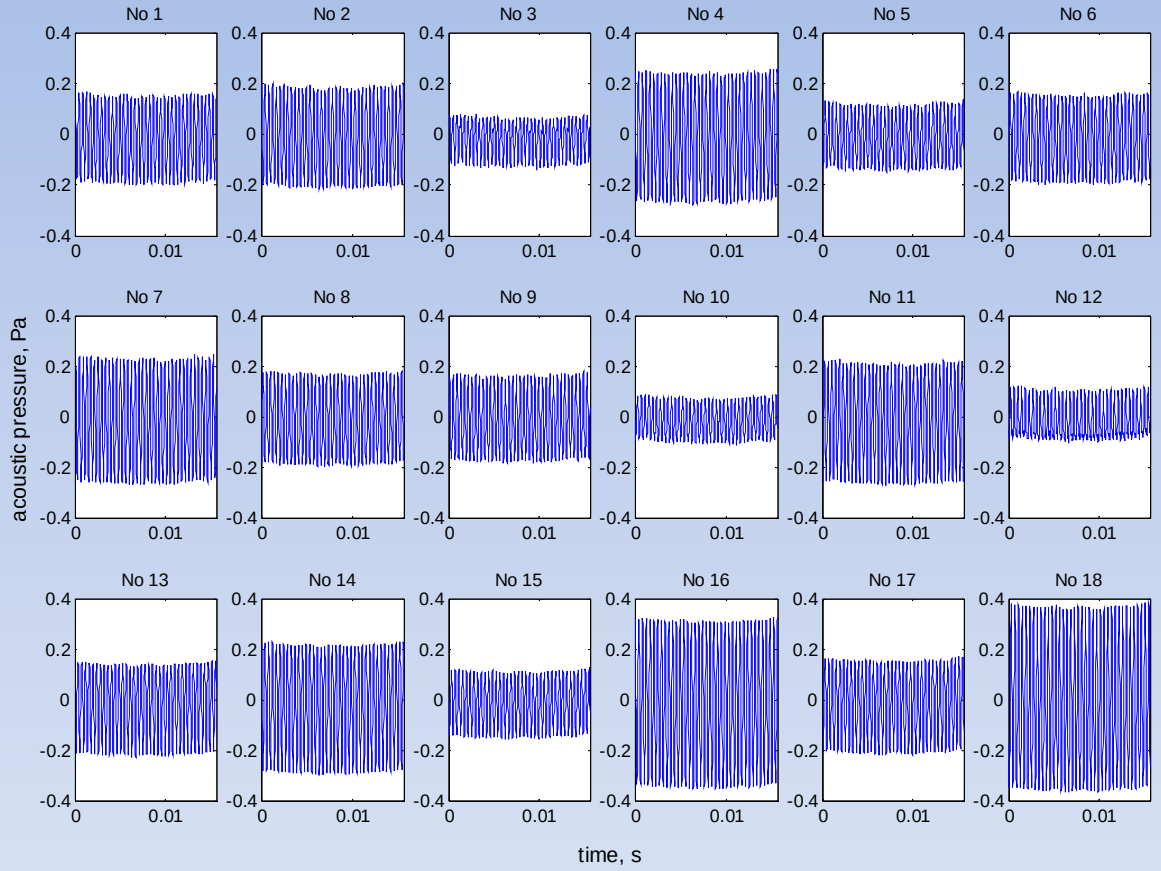
Brüel & Kjær sound analysis system



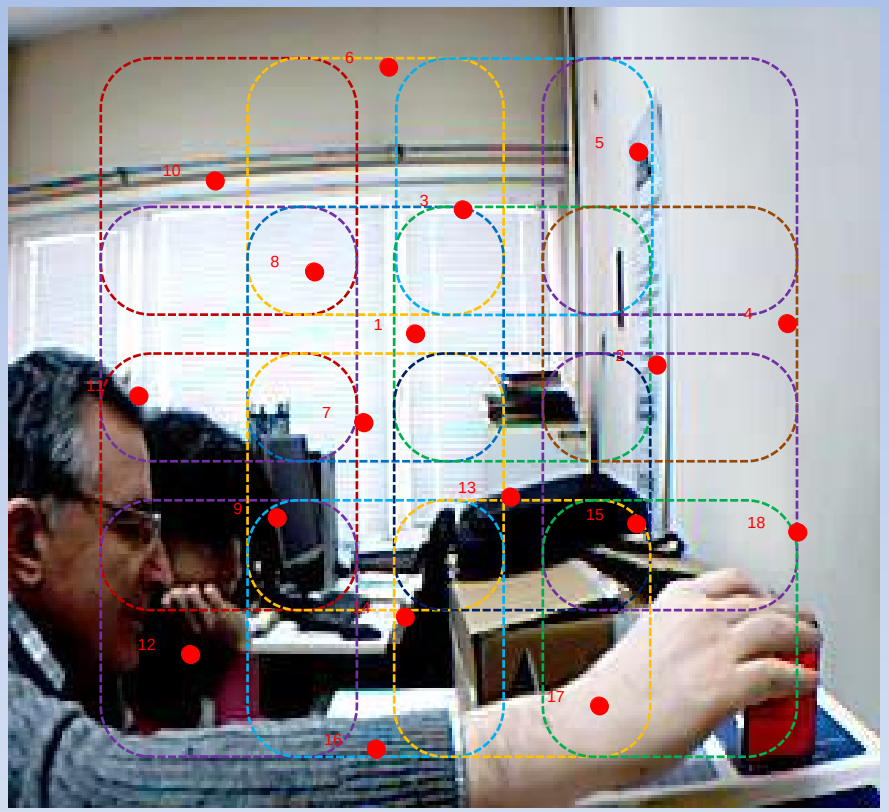
Experimental set-up

our experiment





Proposed algorithm



collect microphone signals

↓

preprocess collected signals using MT receptive fields

features I ↓

preprocess obtained features using ESN

features II ↓

obtain 2D projections

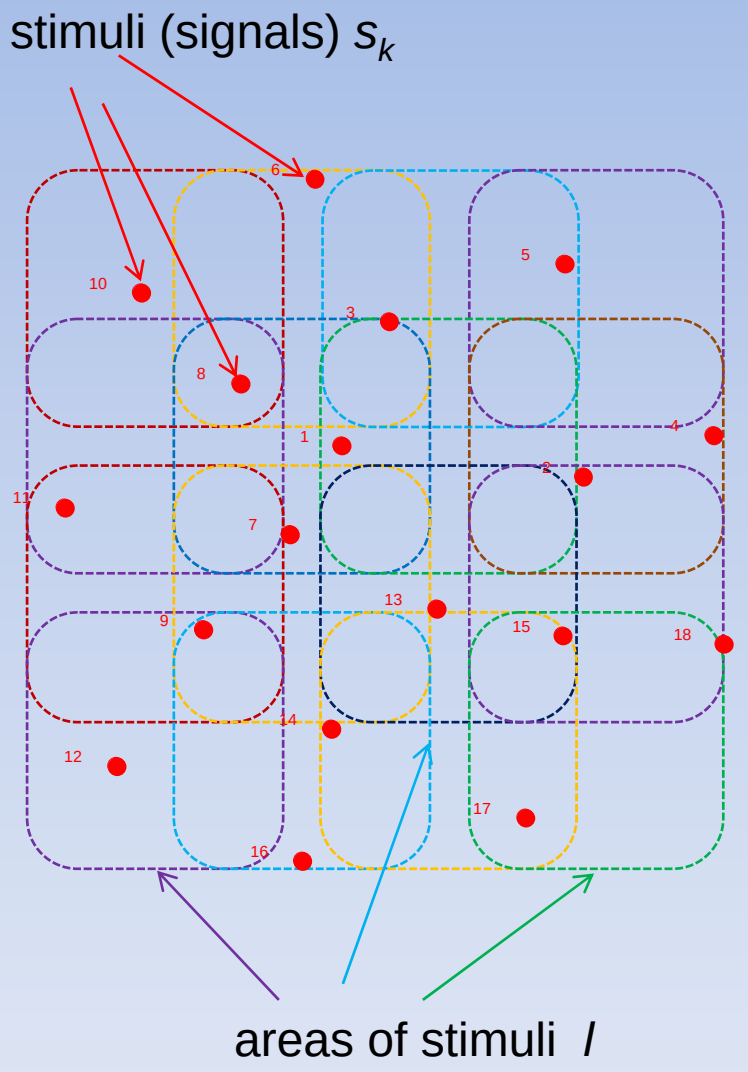
↓

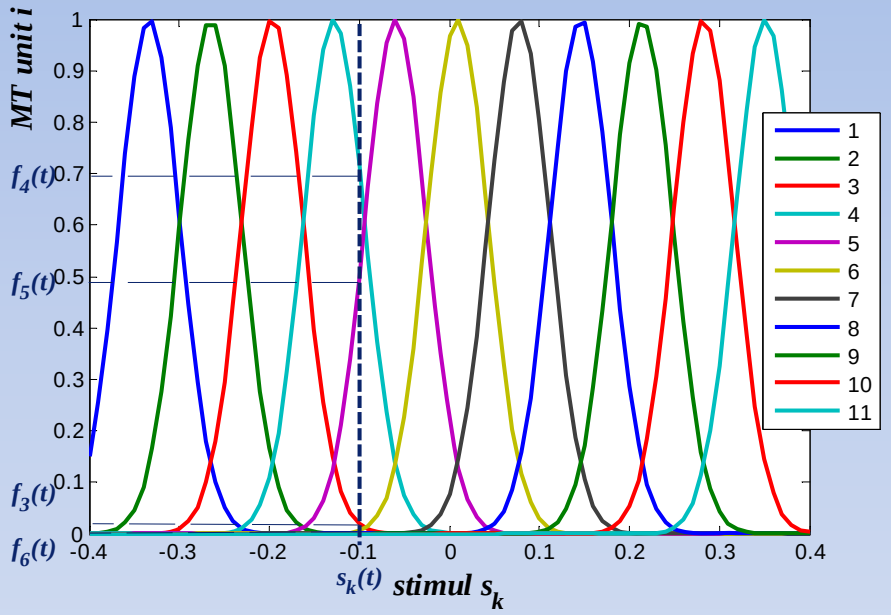
apply subtractive clustering to all projections

↓

choose a projection with maximum clusters

Receptive fields for signal preprocessing



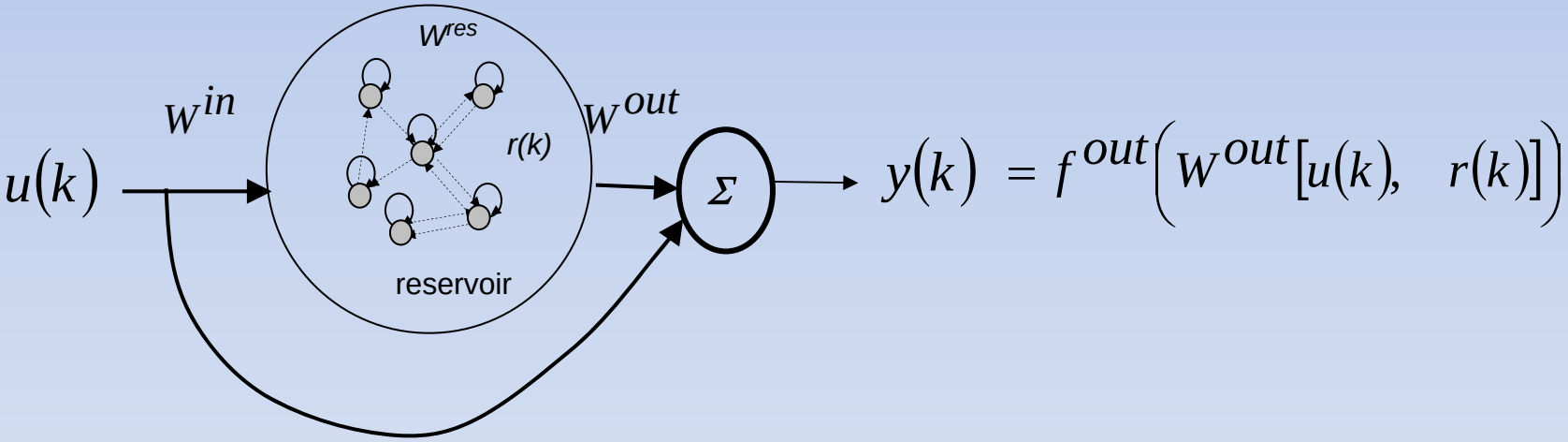


$$f_{il}(t) = \frac{1}{N} \sum_{k=1}^{N} MT_i\left(s_k(t)\right)$$

$$MT_i\left(s_k(t)\right) = \exp\left(\frac{-\left(\mu_i - s_k(t)\right)^2}{2\sigma^2}\right)$$

Echo state networks basics

$$r(k) = f^{res}\left(W^{in}u(k) + W^{res}r(k-1)\right)$$



IP improvement of reservoir

(Schrauwen, B. et al., 2008)

Kullback-Leibler divergence:

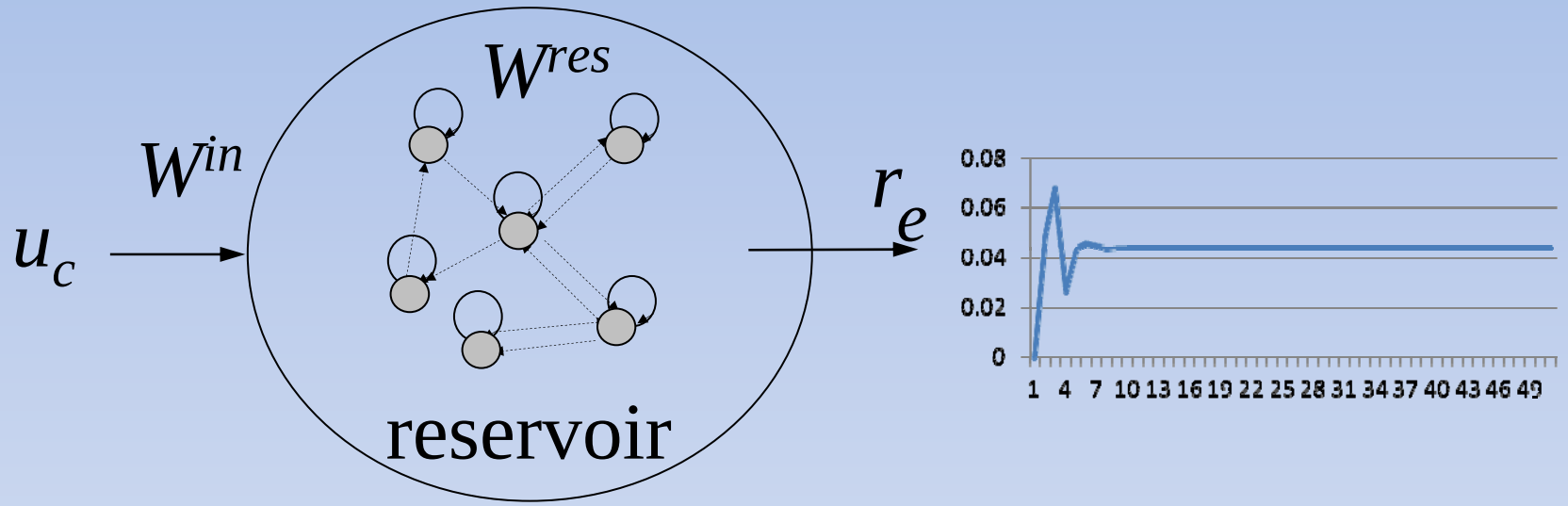
$$D_{KL}(p(r), p_d(r)) = \int p(r) \log\left(\frac{p(r)}{p_d(r)}\right) =$$

$$= -H(r) + \frac{1}{2\sigma^2} E\left((r-\mu)^2\right) + \log \frac{1}{\sigma\sqrt{2\pi}} \rightarrow \min$$

If $f^{res} = \tanh$ then $p_d(r) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(r-\mu)^2}{2\sigma^2}\right)$

$$r(k) = f^{res}\left(diag(a)W^{in}u(k) + diag(a)W^{res}r(k-1) + b\right)$$

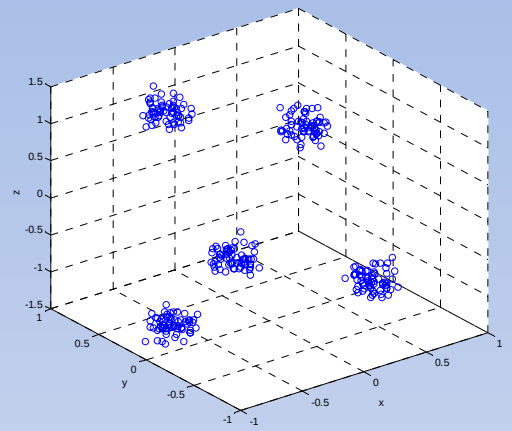
IP Improvement for clustering (Koprinkova-Hristova, P., Tontchev, N., ICANN'2012)

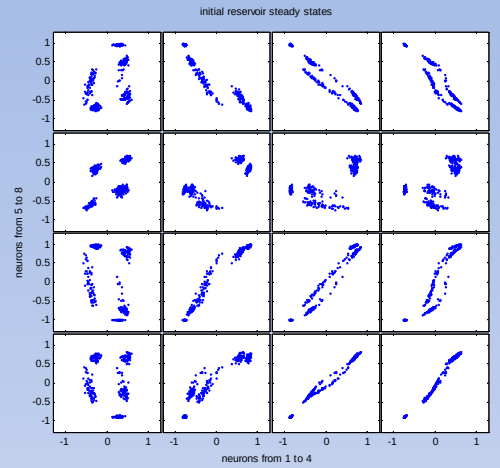


$$r_e = \tanh\left(diag(a) W^{res} r_e + diag(a) W^{in} u_c + b \right), \quad u_c = const.$$

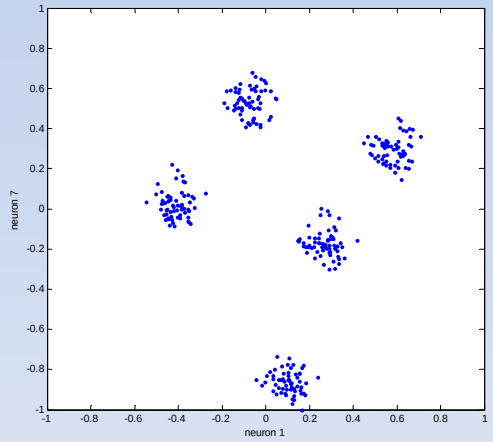
IP improvement for clustering

(Koprinkova-Hristova, P., Tontchev, N., ICANN'2012)

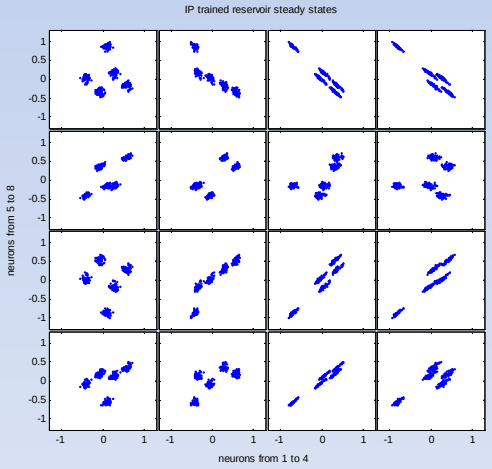




before IP training

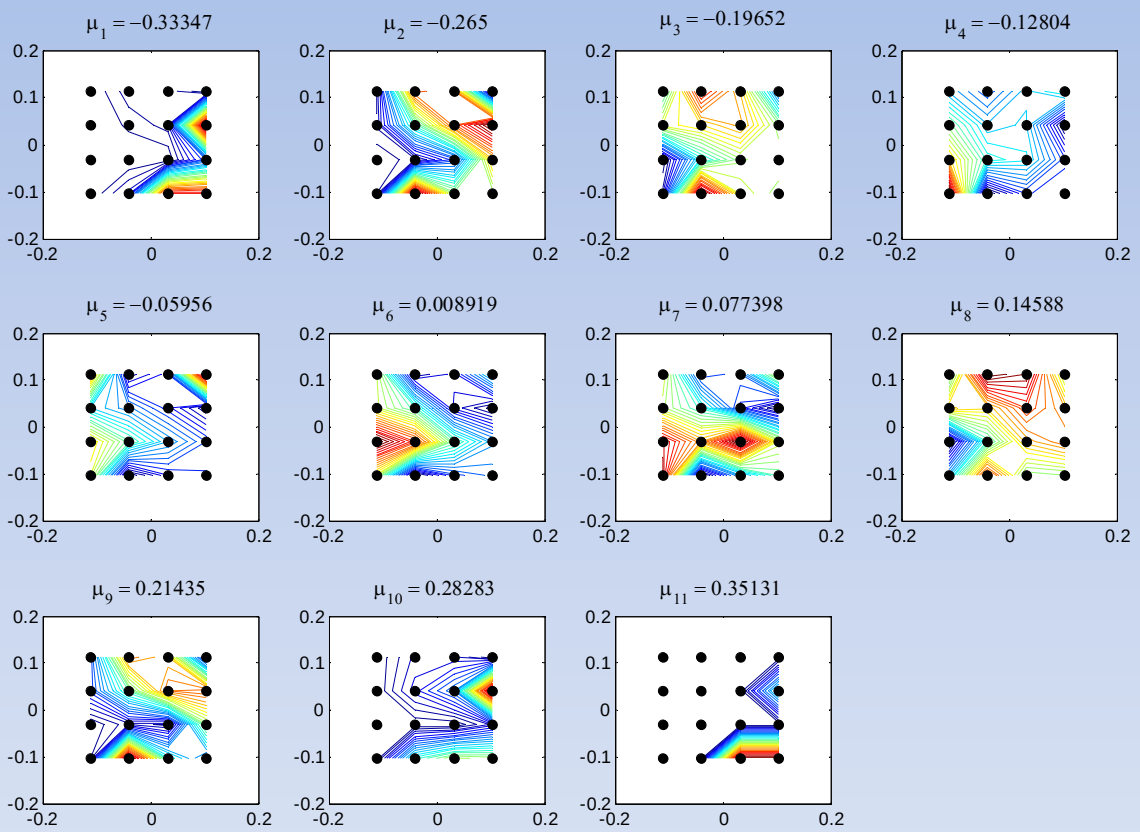


chosen projection

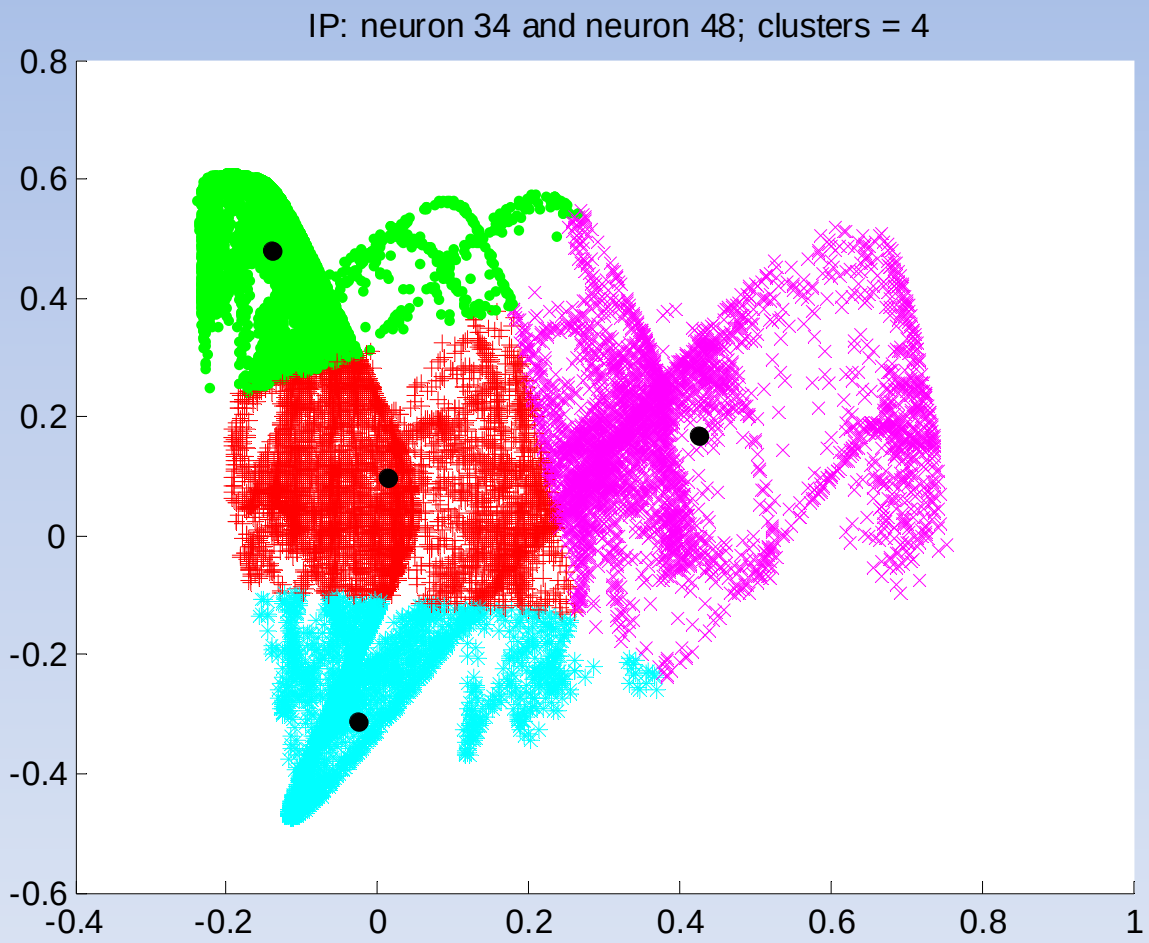


after IP training

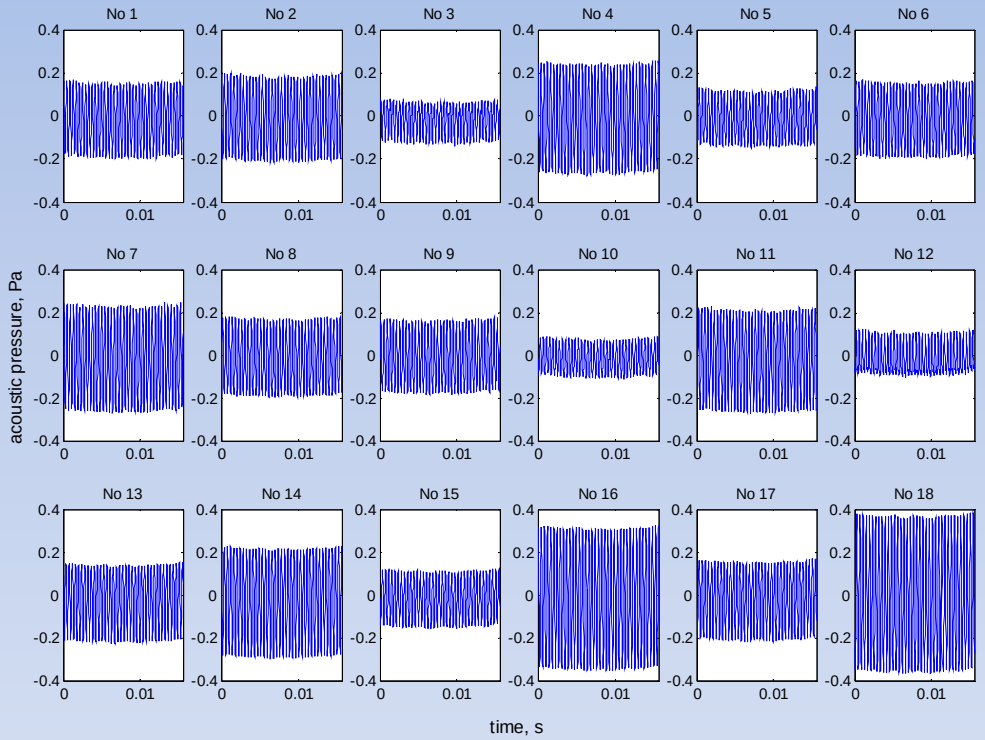
Results
features I



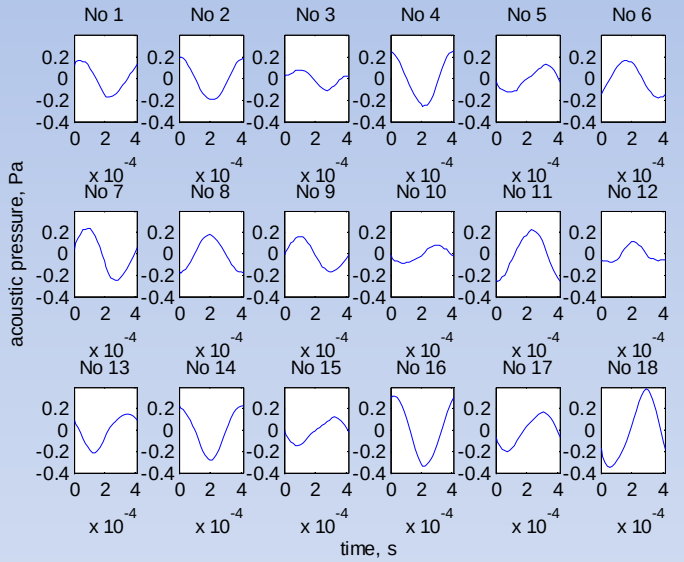
Results
features II & clusters



Results & discussion

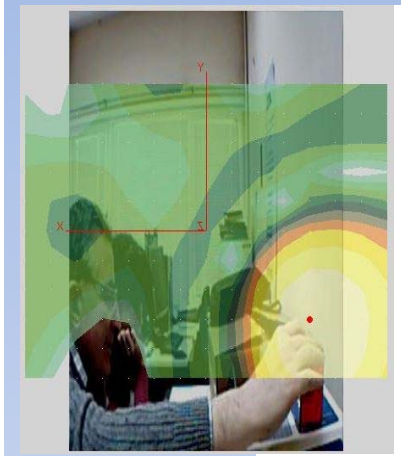


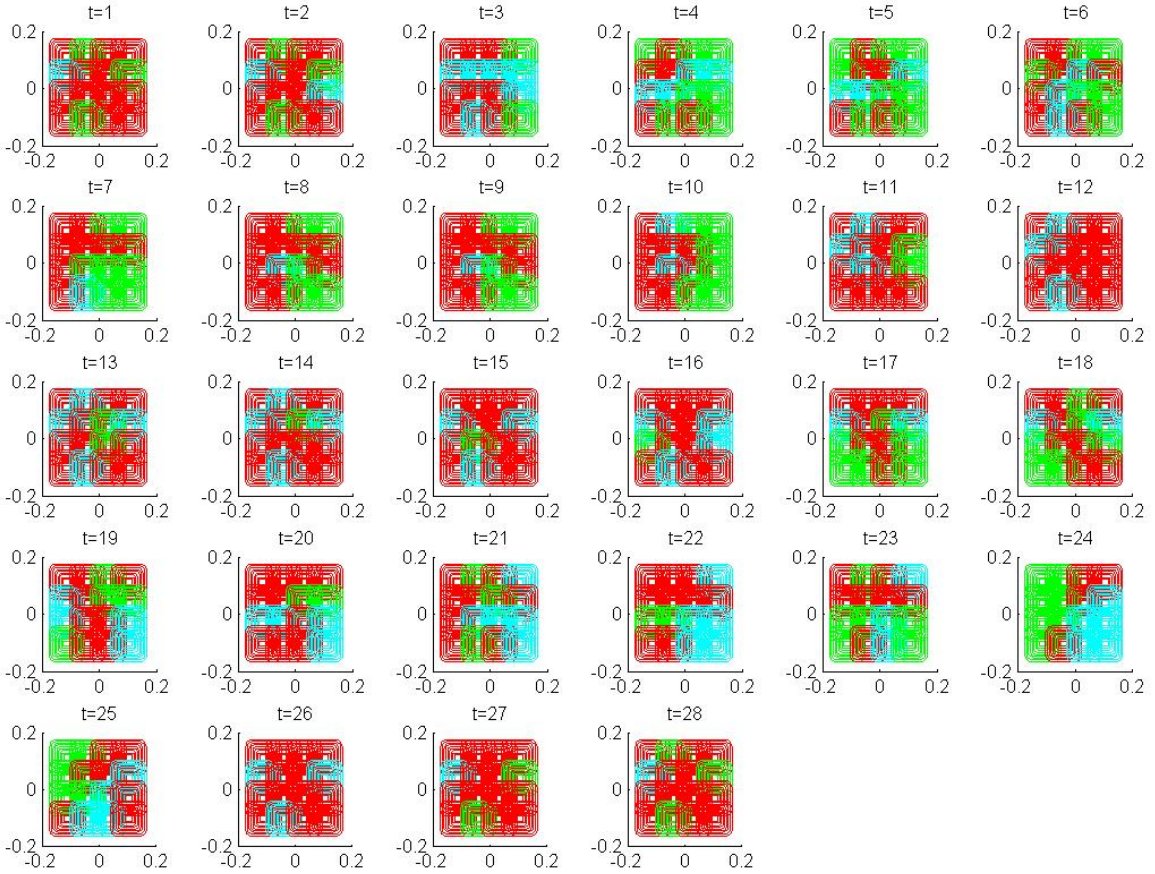
Microphone signals for all period of measurements.



Microphone signals for the first period of approximately 28 time steps.

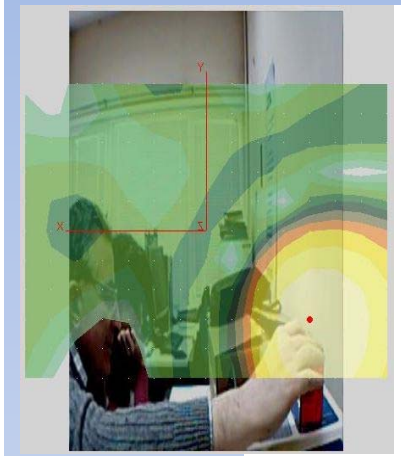
Results & discussions

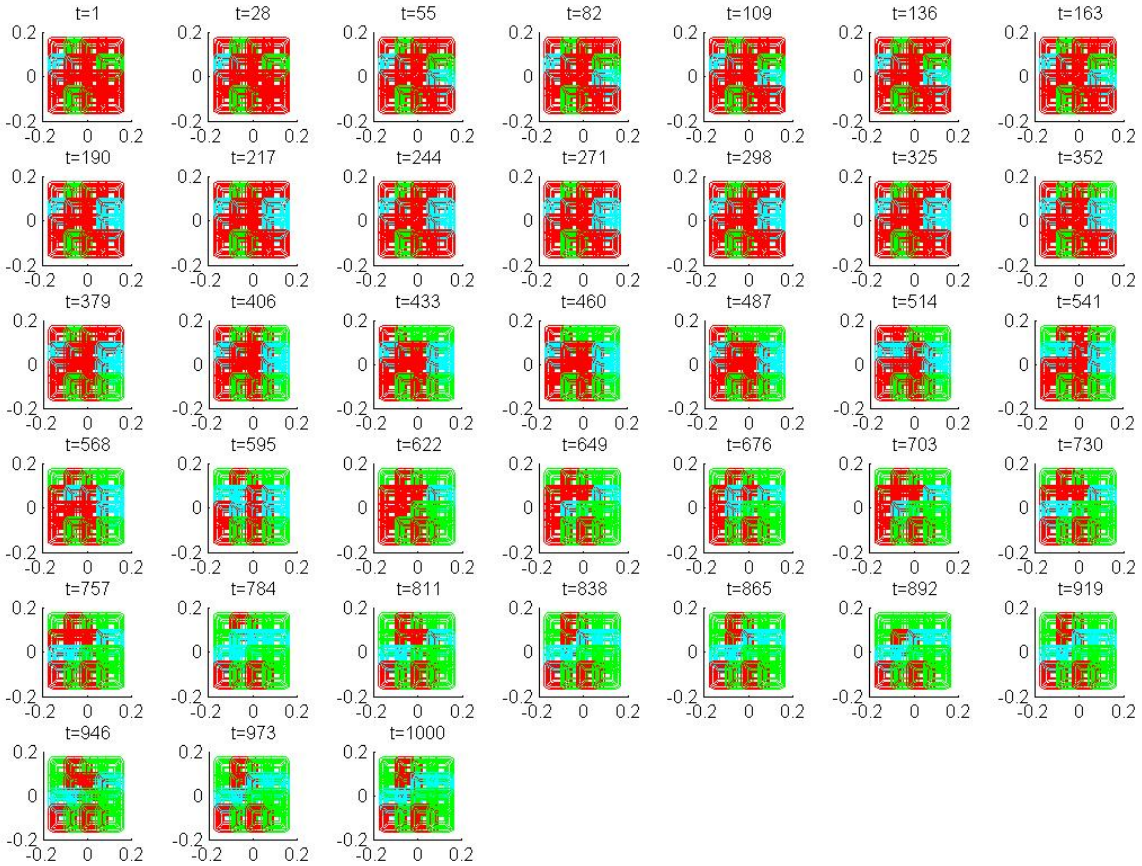




Clusters obtained with 10 neurons for the first period of measurements

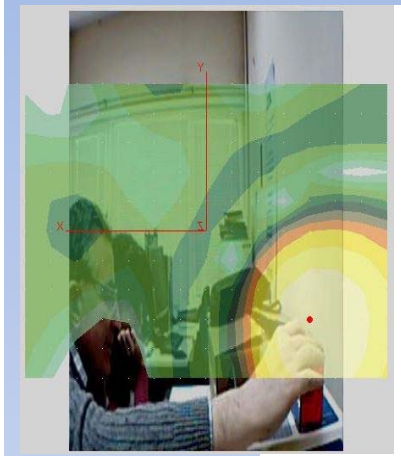
Results & discussion

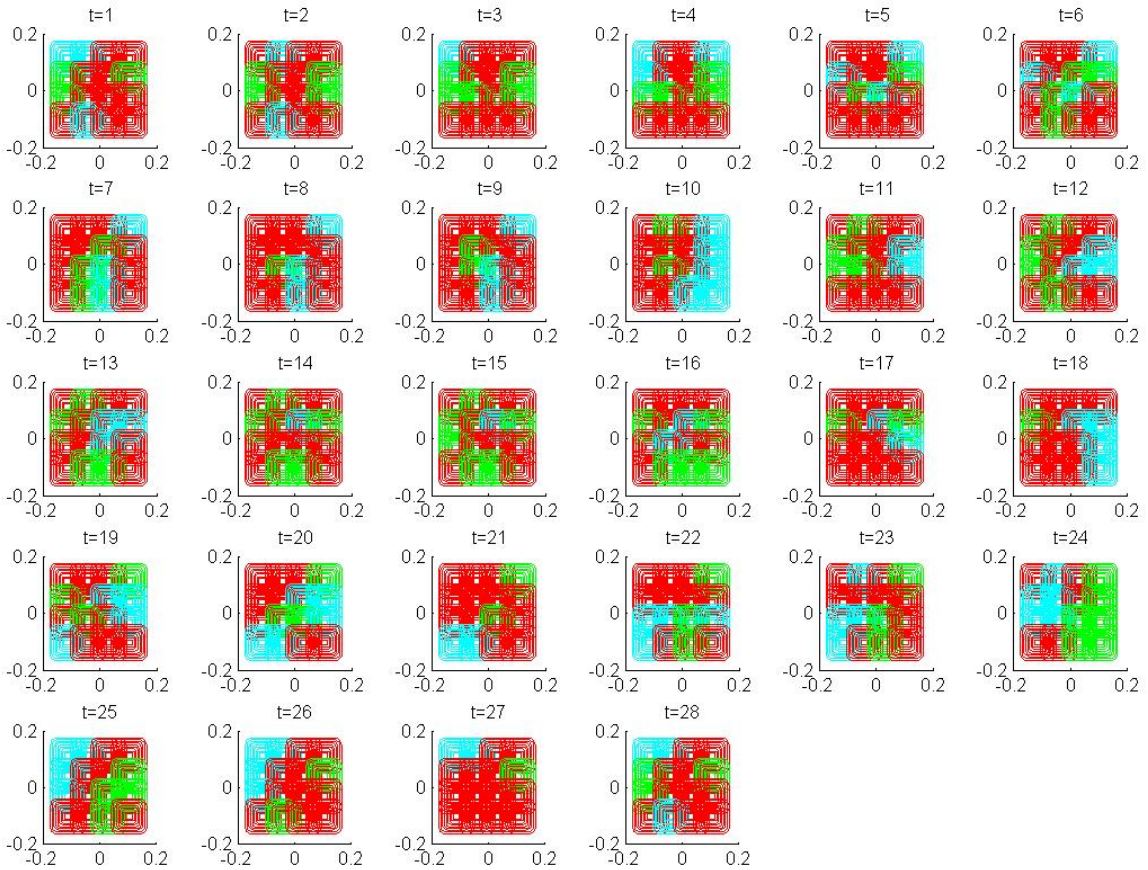




Clusters obtained with 10 neurons for all the time with step 0.412 ms

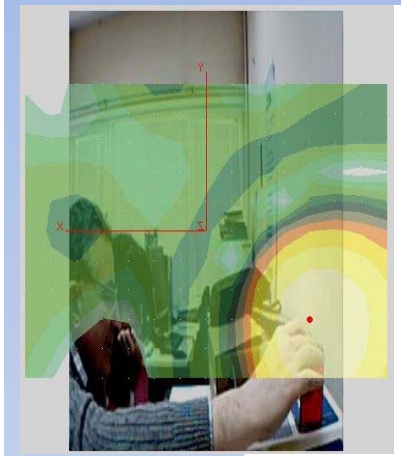
Results & discussions

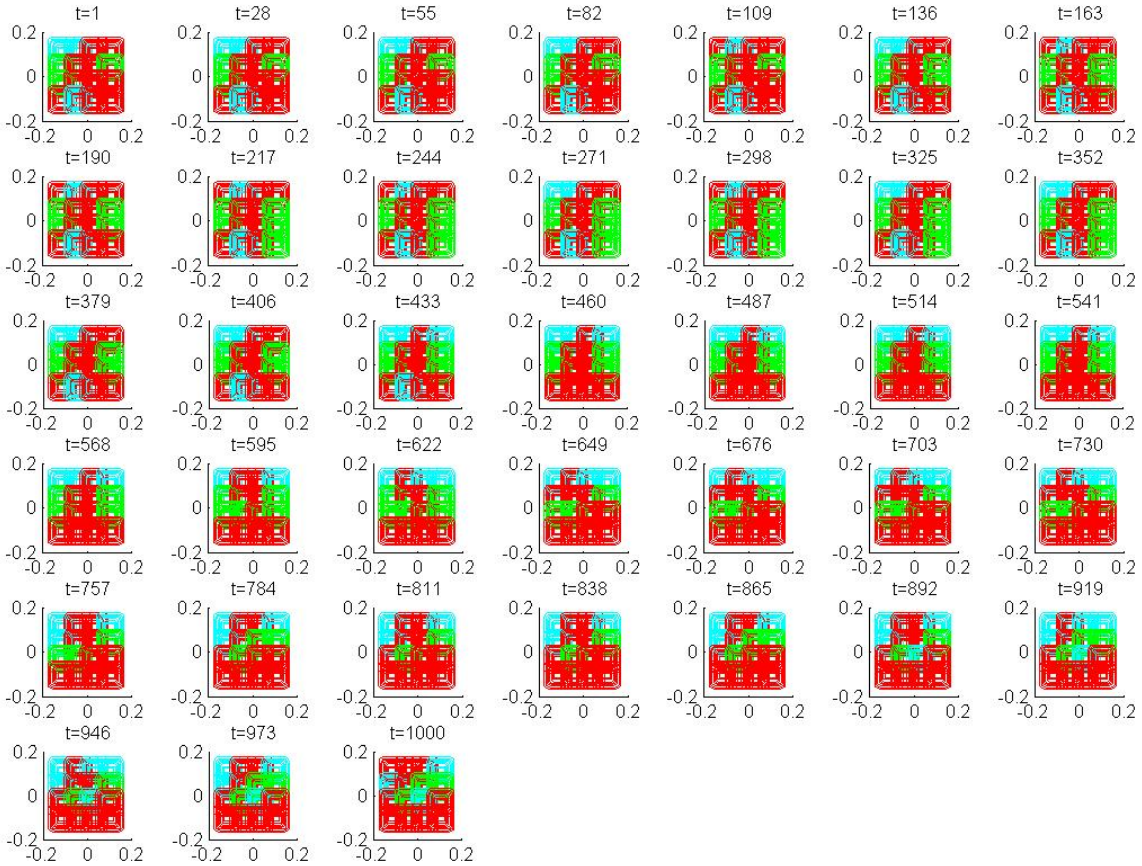




Clusters obtained with 30 neurons for the first period of measurements

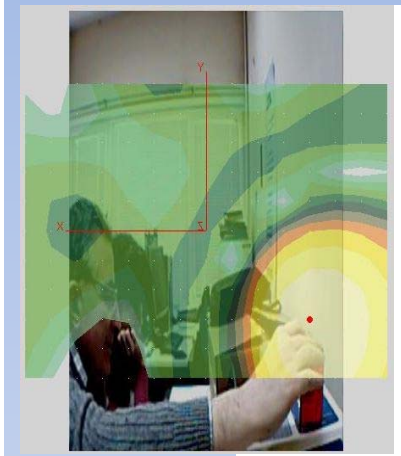
Results & discussion

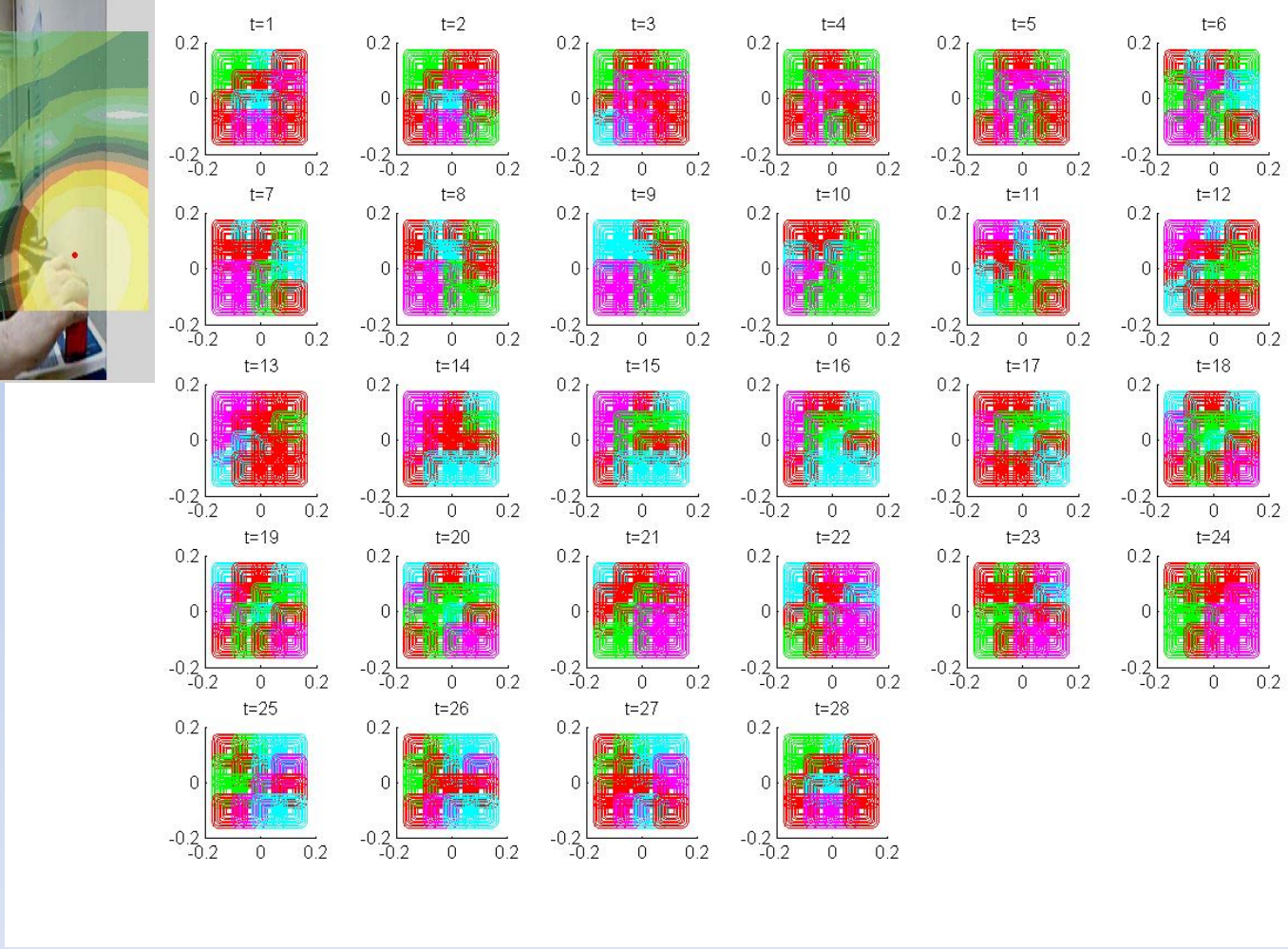




Clusters obtained with 30 neurons for all the time with step 0.412 ms

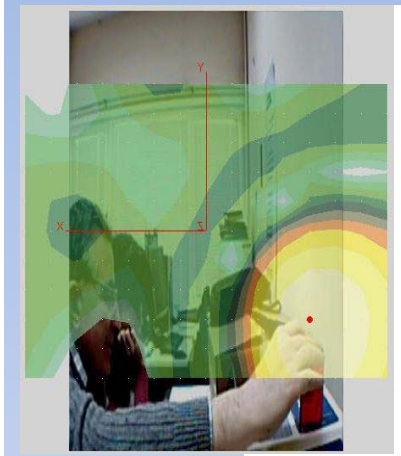
Results & discussions

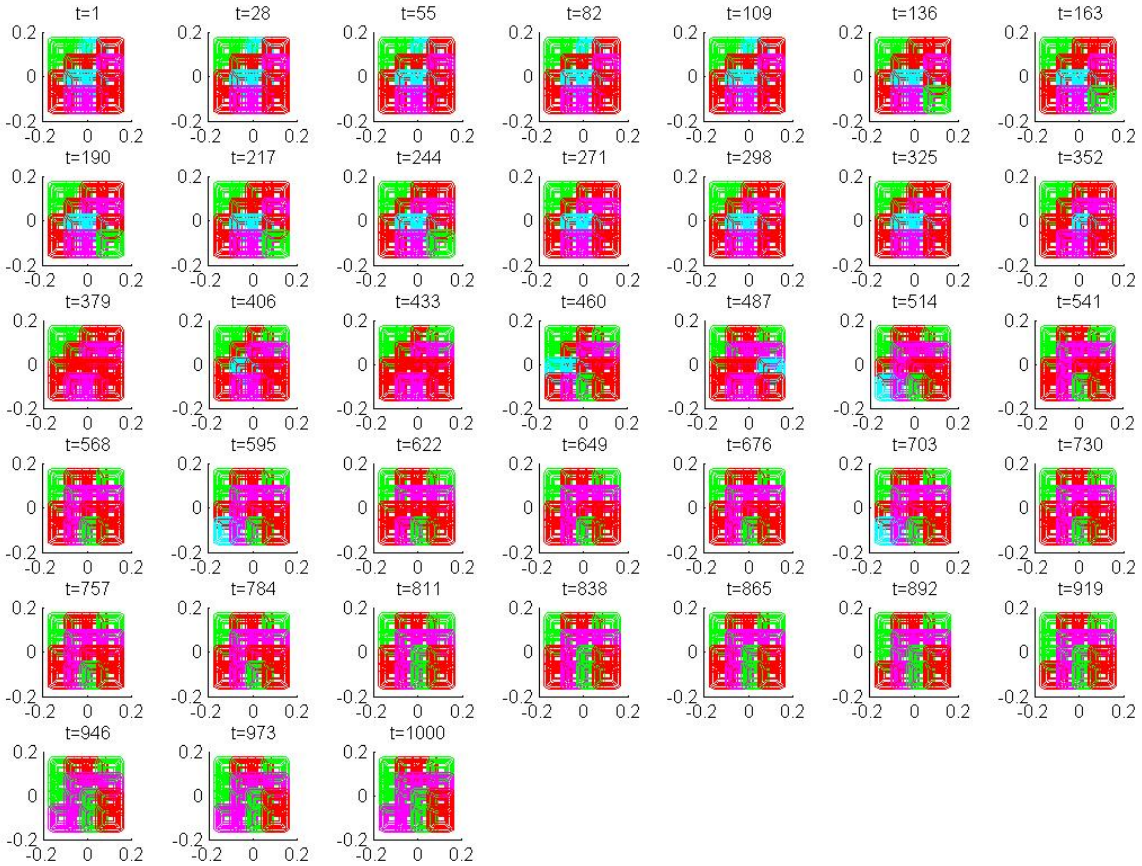




Clusters obtained with 50 neurons for the first period of measurements

Results & discussion





Clusters obtained with 50 neurons for all the with step 0.412 ms

Acknowledgments

The research work reported in the paper was partly supported by the Project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions)

Thank you for attention!

Questions?