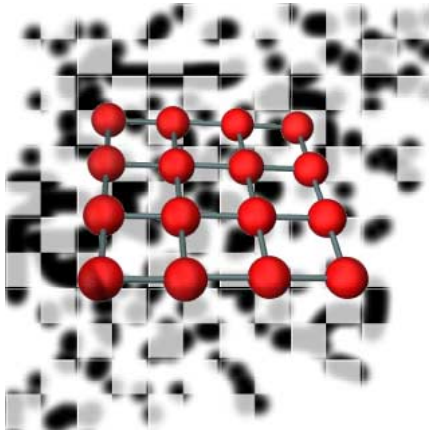


# Dutch Automatic Speech Recognition

## *Using Kohonen Neural Networks*



**Jeroen Schalken**  
June 1998

**TU Delft**

Delft University of Technology  
Faculty of Information Technology and Systems  
Knowledge-Based Systems

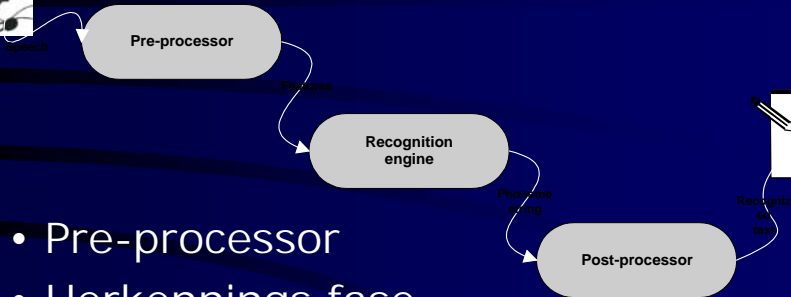
*Graduation Committee:*  
Prof. dr. H. Koppelaar  
Dr. drs. L.J.M. Rothkrantz  
Prof. dr. ir. E.J.H. Kerckhoffs

# Doel

- Self-Organizing Maps (SOM)
  - Phonetic typewriter
- Growing SOM Tree (GSOMT)
- Testbed

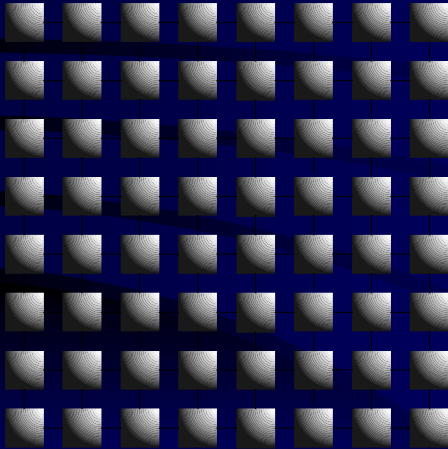


# Spraak herkenning



- Pre-processor
- Herkennings fase
  - phoneme recognition (SOM)
- Post-processor
  - b.v. woord herkenning (HMM, vocabulair)

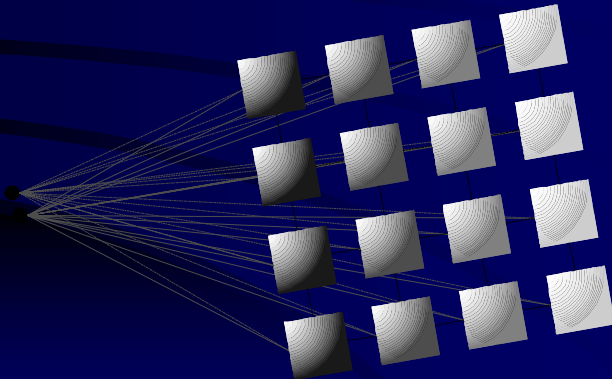
# Self-Organizing Map (SOM)



- Array
- Elke cel is verbonden met zijn burens (neighbors)

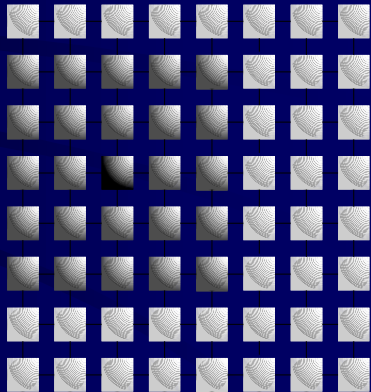
# Self-Organizing Map 2 (SOM)

- Elke cel krijg de volledige input vector



# Trainen van de SOM

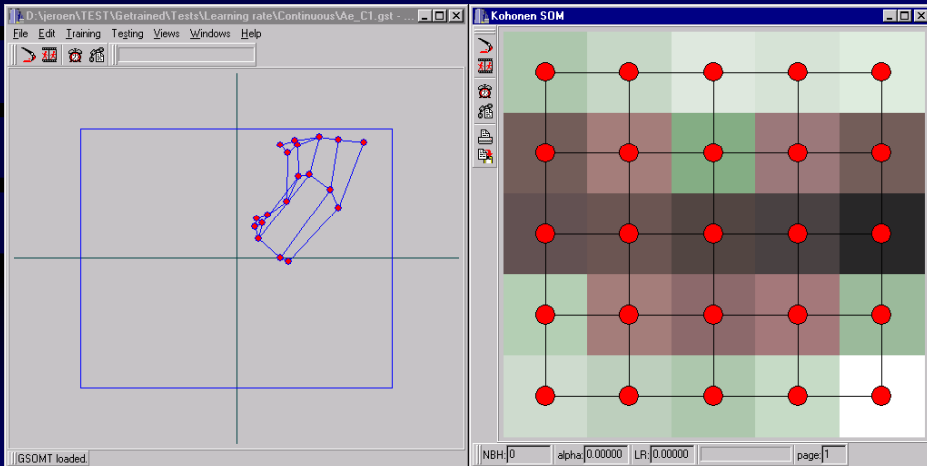
- Winnaar
- Neighborhood
- Updating
- Labeling



# Parameters van de SOM

- Learning rate
  - Over time decreasing
  - Neighborhood decreasing
- Neighborhood function
- Mexican hat function
- Size of the SOM
- Normalization
- Distance measurement

# Testbed (GSOMT)





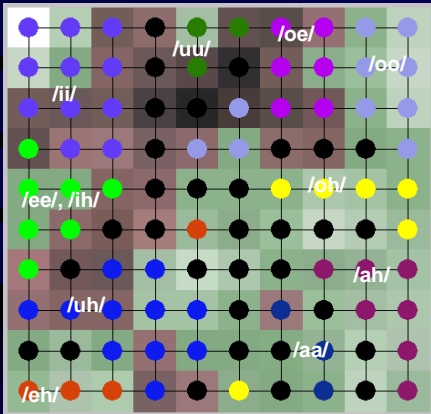
# Demonstratie SOM

- Vorming van een topologisch correcte Map
- Vormen van clusters

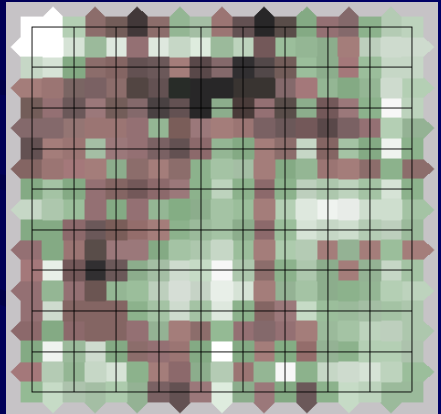
# Phonetic typewriter

- Fins en Japans
- Nederlands
  - Post-processor belangrijk

# Nederlandse klinker-map



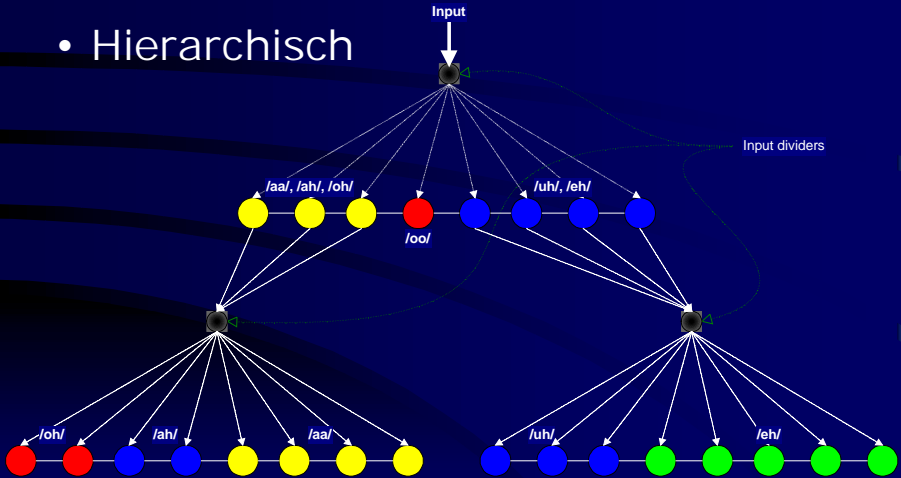
The distribution of the clusters



Combined distances of the weight space

# Growing SOM Tree

- Hierarchisch



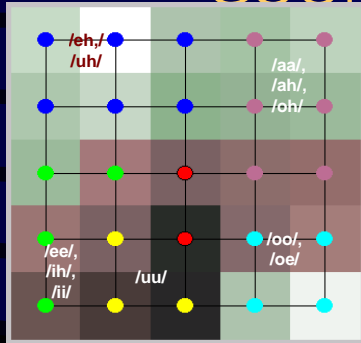
# GSOMT Algoritme

- Train eerste level
- Labelen
- Splitsen
- Toepassen algoritme op volgende levels

# Growing SOM Tree

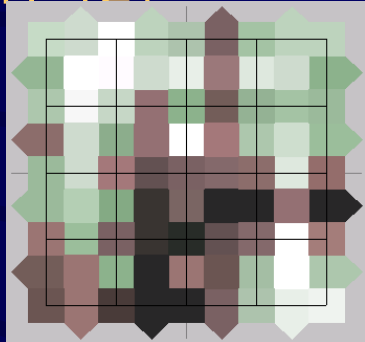
- Voordelen
  - sneller trainen
  - beter onderscheid
- Nadelen
  - later aanpassen van eerste level niet meer mogelijk

# Nederlandse klinker- map, GSOMT (2)

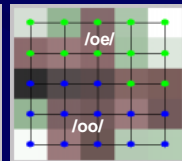
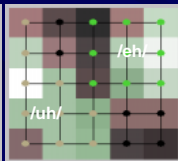
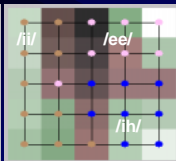
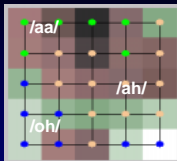


Layer 1

The distribution of the clusters



Combined distances of the weight space



Layer 2

The distribution of the clusters

# Andere phonemen

(medeklinkers)

- onderscheidbaar in groepen (niet individueel)
- onderscheidbaar van de klinkers



# Voordelen SOM

- Trainings data hoeft niet te worden gesegmenteerd
- Tijdens herkenning is het algoritme zeer snel

# Voordelen GSOMT vergeleken met SOM

- Beter onderscheid tussen verschillende clusters
- Sneller tijdens trainen en tijdens herkennen

# Resultaten

- Testbed
  - onderzoek parameters
  - testen nieuwe architecturen
- Klinker-map
  - volledige map

# Letterfrequenties per 10.000

a	b	c	d	e	f	g	h	i	j	k	l	m
633	119	123	512	1586	123	282	232	467	25	187	310	188

n	o	p	q	r	s	t	u	v	w	x	ij/y	z
858	482	70	1	542	351	556	159	223	130	3	115	116

- Klinkers: 3442
- Mede-klinkers: 4951
- Overige leestekens en spaties

# Conclusies

- Herkenner met kleine woordenschat (gebaseerd alleen op klinkers)
- Ter ondersteuning spraak herkenner (bv. hybride met HMM)

# Toekomstig onderzoek

- Nieuwe architecturen om tijdinformatie te bevatten
- Nieuwe manieren voor pre-processing (bv ASSOM, preprocessing in het netwerk)