25 Application of the Self-Organizing Map to the Categorization of Voice and Articulation Disorders, and to Exploration of Emotional Variation of Voice Quality

Lea Leinonen

A series of experiments was carried out to explore the applicability of the self-organizing map to visual imaging of voice and articulation disorders. The visual imaging technique was implemented on a portable PC for demonstrations. For acoustic categorization of voice qualities, a method was developed to select acoustic features with respect to their perceptual significance. The speech samples were assessed by experienced speech pathologists using auditory ratings along 6 dimensions: pathology, roughness, breathiness, strain, asthenia, and pitch.

The clinical studies suggested that visual imaging of speech with the self-organizing map could be used as a feed-back device during therapy: to show the deviation of voice quality from the norm, to show the deviation of phonemes from the norm (misarticulations of children, correction of articulation after cleft palate surgery), and to aid speech training of deaf children. In these applications visual feed-back is superior to auditory feed-back because even hearing subjects with voice or articulation disorders do not usually hear the difference between correct and incorrect performance. At present there is no such visual feed-back device for clinical use. Some training programs to support voice production and correct utterance of phonemes are commercially available for deaf children.

The self-organizing map, or the learning vector quantization, could also be used as diagnostic aid to measure: the degree and the quality of voice and articulation disorder before and after treatment, and deterioration of voice in provocation tests. At present, clinical evaluations are based on auditory ratings. The reliability of auditory rating tests is low because of high intra- and interrater variability. Repeated auditory ratings with several judges are difficult to carry out. For these reasons, the comparison of different surgical or other therapeutic maneouvres is difficult. The lack of reliable measures also restricts the diagnosis of voice disorders without visible anatomical changes, such as those induced by allergens or some inhaled medicines.

For all applications, statistically representative sets of speech data from healthy subjects and subjects with voice and articulation disorders are required. In our clinical studies speech samples were gathered from 200 subjects. This body of data proved to be too small for the selection of acoustic features for comprehensive clinical categorizations.

The self-organizing map was also applied to study emotional variation of voice quality. Spectral energy distributions, modeled by the map, showed differences among speech modes of anger, fear, asthonishment, sadness, scorn, plea, admire, and emotional neutrality.

References

[1] J. Kangas. On the analysis of pattern sequences by self- organizing maps. Doctor's thesis, Laboratory of Computer and Information Science, Helsinki University of Technology, 1994. Among other applications, some studies on pattern recognition of voice disorders were included in this thesis.

[2] L. Liu. Learning vector quantization and simulated annealing in automated feature selection for classification of voice disorders. Master's thesis, Department of Electronics, Helsinki University of Technology, 1995.

[3] H. Rihkanen, L. Leinonen, T. Hiltunen, and J. Kangas. Spectral pattern recognition of improved voice quality. *Journal of Voice*, 8:320-326, 1994.

[4] M.-L. Haapanen, L. Liu, T. Hiltunen, L. Leinonen, and J. Karhunen. Cul-de-sac hypernasality test with pattern recognition of LPC indices. *Folia Phoniatrica & Logopaedica*, 48:35-43, 1996.

[5] L. Leinonen, K. Valkealahti, and H. Rihkanen. Visual imaging of voice quality with the self-organizing map (in Finnish). *Suomen logopedis-foniatrinen aikaukauslehti*, 16:89-96, 1996.

[6] L. Leinonen and H. Poppius. Voice reactions to histamine inhalation in asthma. *Allergy*, 52:27-31, 1997.

[7] L. Leinonen, T. Hiltunen, M.-L. Laakso, H. Rihkanen, and H. Poppius. Categorization of voice disorders with six perceptual dimensions. *Folia Phoniatrica & Logopaedica*, 49:9-20, 1997.

[8] L. Leinonen, T. Hiltunen, I. Linnankoski, M.-L. Laakso. Expression of emotional-motivational connotations with a one-word utterance. *Journal of Acoustical Society of America*, 102:1853-1863, 1997.