Chapter 9

Conclusions

1. We develop an artificial perception approach to pattern recognition. The problem is to recognize structure by identity and structure by similarity. After the structure has been recognized and its elements have been segregated, their recognition is much easier. The model is based on a so called principle of correlativity of perception which is a joint use of grouping and simplicity principles known in psychology. The idea of the correlativity principle is finding a multilevel grouping of stimuli with least total complexity, where the complexity is understood in the sense of Kolmogorov, as the amount of memory store needed for the data representation.

2. The mathematical model considered is based on self-organizing data by their representation in terms of generative elements and their transformations, while subordinating the total representation to the criterion of least complex representation. This is realized by correlation analysis of data under their transformations. In order to perform a directional search for the data transformations which provide high correlation of similar messages, the method of variable resolution is proposed.

3. The model is applied to the problem of chord recognition. A power spectrum of a chord is represented as generated by a spectrum of a tone which is translated along the $\log_2$-scaled frequency axis. We prove the optimality of such a representation of spectral data, corresponding to the causality in the data generation. For this purpose a special mathematical machinery is developed, the deconvolution of a spectrum into irreducible spectra, similarly to the factorization of integers into primes.

4. A simplified model better suitable for practical purposes is substantiated. Instead of power spectra we consider Boolean spectra which are more stable with respect to pitch transpositions of tones. We formulate a necessary condition for generative tone patterns which justifies the use of correlation analysis for discovering generative tone patterns.

5. The model is tested on a series of computer experiments. Since the recognition of intervals is the main procedure in our model of chord recognition, the performance of the algorithm is firstly investigated for the recognition of chords by intervals. In our experiments we use synthesized chords taken from J.S.Bach's four-part polyphony. In particular, we examine different types of recognition mistakes, efficiency and stability of the recognition procedure, and decision making approach based on confronting several hypotheses about the chord obtained from analysis of different types of intervals. Under the worst conditions the reliability of recognition is about 98%.

6. Secondly, we investigate a structural approach to chord recognition where the interval structure of a chord is recognized simultaneously by multi-autocorrelation analysis of chord spectrum and finding optimal spectral representations. This approach is tested on recognizing four-part and five-part synthesized chords. It is shown that the recognition capacity of the model is close to the limits of human perception and resembles its properties: The chord recognizability is better for voices with fewer partials and for more accurate spectral representation, corresponding to better sharpness of music hearing. The harmonies are recognized best, next goes the chord structure, and then chord tones and their pitch.

7. The model is applied to the problem of rhythm recognition. A series of time events is represented in terms of transformations of generative rhythmic patterns, while their temporal distortions being associated with the tempo curve. In order to substantiate a directional search for generative rhythmic patterns, a kind of rhythmic grammar is developed. We formulate segmentation rules based on timing accentuation with which rhythmic syllables are determined. Then we define elaboration, sum, and junction of rhythmic syllables which are used in the structural representation of a sequence of time events. The performance of the algorithm is traced with an example of recognizing the structure of a given rhythm and determining its time.

8. Some applications of the model to psychoacoustics are outlined. In particular, we show that logarithmic scale in pitch perception and insensitivity of the ear to the phase of the signal are necessary for the perception of musical tones as indecomposable sound objects. Besides, we explain the nature of interval hearing as the appearance of correlative perception in audio domain. We show that interval hearing is based on recognition of audio structure by similarity and estimation of distance between similar components. In other words, interval hearing results from the capacity to separate and track simultaneous acoustical processes.

9. The model implies a refinement of definitions of interval, chord, and melodic line. An interval is defined to be the translation distance between similar spectral patterns. A chord is understood as a contour drawn by a spectral pattern of tone in the frequency domain. A melodic line is understood as a trajectory drawn by a spectral pattern of tone in time. Since spectral similarity is the only condition for linking tones in intervals, chords, and melodic lines, their traditional definitions in terms of pitch are generalized to voices with no pitch salience.

10. The model implies a refinement of definitions of rhythm, tempo, and time. The enumerated concepts are associated with different levels of the representation of time events in the model of correlative perception. They are defined in interaction, corresponding to the interaction of the levels of the model. The ambiguity in their determination (rhythm is defined with respect to tempo, and tempo is defined with respect to rhythm) is overcome by optimizing the total complexity of the data representation.

11. The model explains some statements of music theory as conditions aimed at simplifying music perception adequate to the score. From the standpoint of the model, some empirical rules of counterpoint, harmony, and orchestration are rationalized. In particular, we analyze the reasons why certain rules of counterpoint are incompatible with that of orchestration.

12. Finally, we discuss further applications of the model of correlative perception to computer vision, speech recognition and modeling of abstract thinking.