Lecture Notes in Artificial Intelligence 746

Subseries of Lecture Notes in Computer Science
Edited by J. Siekmann

Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

Andranick S. Tanguiane

Artificial Perception and Music Recognition

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Series Editor

Jörg Siekmann
University of Saarland
German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3
D-66123 Saarbrücken, Germany

Author

Andranick Tanguiane
Fernuniversität Hagen
Postfach 940, D-58048 Hagen, Germany

CR Subject Classification (1991): I.2, I.5, E.2

ISBN 3-540-57394-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-57394-1 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Typesetting: Camera ready by author
Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
45/3140-543210 - Printed on acid-free paper

Foreword

In this book I summarize my studies in music recognition aimed at developing a computer system for automatic notation of performed music. The performance of such a system is supposed to be similar to that of speech recognition systems: acoustical data at the input and music score printing at the output.

In this essay I develop an approach to pattern recognition which is entitled *artificial perception*. It is based on self-organizing input data in order to segregate patterns before their identification by artificial intelligence methods. The performance of the related model is similar to distinguishing objects in abstract painting without their explicit recognition.

In this approach I try to follow nature rather than to invent a new technical device. The model incorporates the *correlativity of perception*, based on two fundamental perception principles, the grouping principle and the simplicity principle, in a very tight interaction.

The grouping principle is understood as the capacity to discover similar configurations of stimuli and to form high-level configurations from them. This is equivalent to describing information in terms of generative elements and their transformations.

The simplicity principle is modeled by finding the least complex representations of data that are possible. The complexity of data is understood in the sense of Kolmogorov, i.e., as the amount of memory storage required for the data representation.

The tight interdependence between these two principles corresponds to finding generative elements and their transformations with regard to the complexity of the total representation of data. This interdependence justifies the term "correlativity", which is more than relativity of perception.

The model of correlative perception is applied to voice separation (chord recognition) and rhythm/tempo tracking.

Chord spectra are described in terms of generative spectra and their transformations. The generative spectrum corresponds to a tone spectrum which is repeated several times in the chord spectrum. The transformations of the generative spectrum are its translations along the $\log_2$-scaled frequency axis. These translations correspond to intervals between the chord tones. Therefore, a chord is understood as an acoustical contour drawn by a tone spectral pattern in the frequency domain.

Time events are also described in terms of generative rhythmic patterns. A series of time events is represented as a repetition of a few rhythmic patterns which are distorted by music elaboration and tempo fluctuations associated with the tempo curve. The interdependence between tempo and rhythm is overcome by minimizing the total complexity of representation, e.g., the total amount of memory needed for storing rhythmic patterns and the tempo curve.

The model also explains the function of interval hearing, certain statements of music theory, and some phenomena in rhythm perception.

Generally speaking, I investigate hierarchical representations of data. In particular, I pose the following questions:

(a) Why a hierarchy?

(b) Which hierarchy? and

(c) How does the hierarchy correspond to the reality?

From the standpoint of the model, the answers to these questions are, respectively:

(a) A hierarchy makes a data representation compact, which is desirable in most cases;

(b) consequently, a better hierarchy is one which requires less memory for the related data representation; and

(c) under certain assumptions such a hierarchy reveals perception patterns and causal relationships in their generation, making the first step towards a semantical description of the data.

One can see that the main distinction of this approach is finding optimal representations of data instead of directly recognizing patterns. In a sense, analysis of patterns is replaced by synthesis of data representations. Since self-organization is used instead of learning, the threshold criteria used in most pattern recognition models are avoided.

The correspondence between music perception and the performance of the model, together with the diversity of its applications, can hardly be regarded as simply a coincidence. It makes an impression that the model really simulates certain perception mechanisms. Probably, the related model can be applied to speech recognition, computer vision, and even simulation of abstract thinking. All of this is a subject for discussion.

This book has been written during my work in Grenoble at the ACROE–LIFIA (Association pour la Création et la Recherche sur les Outils d'Expression–Laboratoire d'Informatique Fondamentale et d'Intelligence Artificielle).

I acknowledge Professor Philippe Jorrand, the director of the LIFIA, and Dr. Claude Cadoz and Dr. Annie Luciani, the heads of the ACROE, for inviting me to Grenoble and the stimulating working conditions; my colleagues, especially Martial Barraco, for their friendly attitude and help in various domains; and Remi Ozoux, a student of ENSIMAG (Ecole Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble), who has fruitfully collaborated on the project and who has programmed the latest version of the algorithm for chord recognition.

Hagen, 14 September 1993 Andranik Tangian (Tanguiane)

Contents

List of Figures

List of Tables