# Studies in Computational Intelligence

Volume 842

**Series Editor**

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

The books of this series are submitted to indexing to Web of Science, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink.

More information about this series at http://www.springer.com/series/7092

Zbigniew Les · Magdalena Les

# Machine Understanding

## Machine Perception and Machine Perception MU

Springer

Prof. Zbigniew Les
The St. Queen Jadwiga Research Institute of Understanding
Toorak, Melbourne, VIC, Australia

Magdalena Les
The St. Queen Jadwiga Research Institute of Understanding
Toorak, Melbourne, VIC, Australia

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-030-24069-1 ISBN 978-3-030-24070-7 (eBook)
https://doi.org/10.1007/978-3-030-24070-7



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*This book is dedicated to our Patron*
*St. Jadwiga Queen of Poland*

# Introduction

This book presents the selected research results in a newly established area of scientific research which we call machine perception MU. Machine perception MU research is carried out within the machine understanding framework. Machine understanding is based on further development of the research that was presented in our previous books titled *Shape Understanding System: the First Steps toward the Visual Thinking Machines*, *Shape Understanding System: Knowledge Implementation and Learning* and *Shape Understanding System: Machine Understanding and Human understanding* [1–3]. This is the fourth book that presents the research results in the area of thinking and understanding, carried out by authors in the newly founded the St. Queen Jadwiga Research Institute of Understanding. Machine understanding is the term introduced by authors to denote understanding by a machine (SUS) and is referring to the new area of research the aim of which is investigating the possibility of building the machine with the ability to understand. SUS, as the machine that is designed to have the ability to think and understand, learns both knowledge and skills in the process of learning called the knowledge implementation. A machine to be able to understand needs, to some extent, mimic human understanding and for this reason machine understanding is based on the assumption that the results of understanding by the machine (SUS) can be evaluated according to the rules applied for evaluation of human understanding. The important part of machine understanding approach is investigation of the different forms of explanation how to solve a problem (text problem) or explanations of the causes and context of an object or phenomenon.

In the first book [1], a brief description of philosophical investigations of topics connected with understanding and thinking, was presented. The shape that is the main perceptual category of thinking process and the important visual feature of the perceived world was briefly described. In Chap. 2, the shape classes that are regarded as the basic perceptual categories were presented. Shape classes are represented by their symbolic names. Each class is related to each other and, based on these classes, there is relatively easy to establish the ‘perceptual similarity’ among perceived objects. In Chap. 3, the description of the reasoning process that leads to assigning the perceived object to one of the shape classes was described. In

Chap. 4, the new hierarchical categorical structure of the different categories of the visual objects was presented. In Chap. 5 examples of visual reasoning processes that can be considered as the special kind of thinking processes were presented. The thinking process is regarded as the continuous computational activity that is triggered by perception of a new object, an 'inner object' or a task given by a user.

In Chap. 1 of the second book [2], some aspects of human learning that are related to the newly introduced concept of the knowledge implementation, were described. In Chap. 2, a short survey of literature on the vast topic concerning learning by a machine, was presented. In Chap. 3, knowledge implementation was defined in the context of both human learning and machine learning. In Chap. 4, the selected issues connected with learning and understanding, were described in the context of the newly introduced concept of knowledge implementation. The relations between understanding and learning were also discussed. In Chap. 5, the shape understanding method was presented. In Chap. 6, categories of the visual objects were described. In Chap. 7, the theoretical framework of the knowledge implementation method was presented. In Chap. 8, the knowledge implementation as a new method of learning knowledge and skills of the different categories of objects was presented. The short description of the shape understanding system was presented in Sect. 8.1. In Sect. 8.2, learning of the new knowledge was presented. Learning of the knowledge of the visual objects can be seen as learning of the specific perceptual skills to acquire data from an image. In Sect. 8.2, learning and understanding of the text that belongs to one of the text categories such as the category of text-query, the category of text-task or the category of dictionary-text, was presented.

In the third book [3], the new term machine understanding is introduced, to denote understanding by a machine (SUS) and is referring to the new area of research the aim of which is investigating the possibility of building the machine with the ability to understand. In Chap. 1, a presentation of the point of view of selected thinkers on the topic concerning human understanding and a discussion of some aspects of understanding considered to have implication for material presented in this book is given. In Chap. 2, a short survey of literature, on the vast topic concerning the existing 'understanding' systems, is described. In Chap. 3 machine understanding is defined in the context of both human understanding and existing systems that can be regarded as understanding systems. In Chap. 4 machine understanding that is based on the shape understanding method is presented. In Chap. 6 examples of selected problems used for testing whether these problems can be solved by the machine (SUS) are presented. The special classes of problems that are described in Chap. 6 are problems called text-tasks used for testing the results of learning at school. In Chap. 7 visual understanding, regarded as problem solving that involves naming and recognizing visual objects, is presented. In this chapter examples of learning and understanding of the objects from the leaf category and butterfly category are described. Generalization, specialization, schematization, visual abstraction and imaginary transformations applied during knowledge implementation, which are essential parts of learning and understanding, are also described in this chapter. In Chap. 8 naming of objects that are members of the sign category is presented. Understanding of the objects from the sign category means finding their meaning that is given by assumed conventional meaningful relation, called the coding system. In this chapter understanding of the objects from the selected sign categories such as the musical symbols category, the electronic symbols category or the road signs category is also discussed In Chap. 9 understanding of the objects from the text category is described. Understanding of objects from the text category, regarded as problem solving, means finding the meaning of the text and next interpreting this text in terms of knowledge of the world. In this chapter understanding of the set theory texts is presented as the example of understanding of the text category. In the last chapter understanding of the explanatory text that is generated during explanation of finding of the solution of the command-text-tasks such as '*solve an algebraic equation*' and generating of the explanatory text during explanation of finding of the solution to this command-text-tasks, is presented.

This book presents the selected research results in machine perception MU that is the further development of the machine understanding research concerning perception as the part of the understanding process. Machine perception is term introduced by Nevatia [4] for denoting the research area where researchers investigate the possibility to build the systems that can be endowed with human perceptual ability. Although sharing the same term 'machine perception', research in machine perception MU is based on the very different approach.

Machine perception is based on common believe of researchers from the engineering community that the meaning of the concepts such as the perception is founded on the solid basis of the scientific, empirical and theoretical findings. As it will be shown in Chap. 1, such an expectation is a mere illusion that is the matter of the deep ignorance, rather, than any rational justification. Over the years, many cognitive theories of perception have been proposed, evaluated, revised, and evolved within an impressive body of research. The perception was one of the oldest research interests for both scientists and philosophers and, instead of many efforts and painstaking search for solution to problems of perception, the results are far from being satisfactory. Although scientific theories concerning perception present a valuable steppingstone towards the goal of machine perception to embody this unique human ability within a computational system, there is the urgent need to redefine machine perception taking into account not only human or animal perception but, what is the most important, the machine oriented approach that can lead to some sort of autonomy in defining the basic categories of machine perception. Even among the researchers in human perception there are tendencies to trivialize the meaning of perceptual processes placing them into the corner of the meaningless tasks of the striving for survive. On the other hand, perception is regarded as the inferior of the two cognitive powers because it supposedly lacked the distinctness that comes only from the superior faculty of reasoning, higher cognitive functions of the mind, responsible for creating concepts, accumulating knowledge, connecting, separating, and inferring. Although there is some attempt to promote the view that perception is intelligent, in that it is based on operations similar to those that characterize thought, there is, however, big difference between solving the most complex perceptual problems and solving even simple theoretical problems, such as the mathematical proof of the basic theorems. Machine perception research that is focused on building a machine that will be endowed with human perceptual ability is following the way of thinking of those narrow minded people of science being preoccupied with topics such as object recognition, object detection or solving the navigation problems. However, as empiricist pointed out, the perception is the source of knowledge acquired through the historically evolved process of collecting of the sensory data, the knowledge that is stored in the form of the scientific books and papers, knowledge that leads to understanding of the world. In this context, it is a truism to say that it is understanding that profits from perception, perception that supplies the nutrition in the form of the sensory data. For this reason, the new field of research called machine perception MU, that started with authors' research on machine understanding, at St. Queen Jadwiga Research Institute of Understanding, has emerged. Machine perception MU research, carried out within the machine understanding framework, differs in some important aspects from the classical machine perception approach that is focused on object recognition, object detection or solving the navigation problem. A machine that is endowed with the understanding ability needs to relay on the different perceptual strategies of gathering and transforming data. Not only analysis of problems in machine perception but analysis of any problem must be carried out within some framework that selects the fundamental assumptions and problem definitions on which the research will be founded. For this reason, machine perception MU research that are carried out within machine understanding framework is deeply rooted in this research context that supply fundamental problems definitions.

The main drawback of the classical machine perception is that it is focused on the pragmatic side of the perceptual process, without paying enough attention to the cognitive aspect of the perceptual process. However, the pragmatic aspect of the human perceptual gathering and processing data is only the small part of the perceptual process aimed at the securing of the basic existential needs. Most of the human perceptual activities are directed to gather data used to build the knowledge that makes it possible to understand the world. With advent of empiricism the perceived data become the basis for near all scientific knowledge that is used during understanding of the world. For this reason, a machine that is endowed with understanding ability needs to relay on the different perceptual strategies of gathering and transforming data and the research area, part of machine understanding that will investigate these perceptual problems, we will call machine perception MU. Machine perception is attempting to enable man-made machines to perceive their environment by sensory means as humans and animals do. When classical machine perception is focused on gathering sensory information in order to transform it to some form that makes it possible to use it to perform required task, machine perception MU is aimed at building a machine that can think and understand. This new perspective places the research in machine perception MU within the framework of machine understanding research that supplies findings for exploring new possibilities of understanding and thinking abilities of the machine that can be very different from that of human abilities. In our approach the main focus is on perception as the subordinate to understanding and by this directing the research effort not on the sensory devices but on the ability to utilize the transformed data in understanding process.

Machine perception MU is concerned with machine interpretation of any sensory data. In machine perception (robotics) the sensory devices play a big role in supplying the perceptual data, often imitating the biological systems in their perceptual functioning (evolutionary psychology). Biological systems sense their environments by a variety of sources such as sight, sound, touch and smell. The sense of touch is actually many senses, including pressure, heat, cold, tickle, and pain. The sensory system is a part of the nervous system responsible for processing sensory information. It consists of sensory receptors, neural pathways, and parts of the brain involved in sensory perception. Commonly recognized sensory systems are those for vision, hearing, somatic sensation (touch), taste and olfaction (smell). Human being utilizes usually two types of sensory information e.g. seeing and hearing during understanding of the environment, whereas a machine can use more complex perceptual data that are not present in the human perception. Until now, the research in machine perception utilizes sensory data in making decision during object recognition or navigation of the robot. Machine perception MU is concerned with the machine interpretation of any sensory data. With the advent of sensor networks capable of perceiving many different data it opens the new possibilities of exploration of the new methods of building knowledge based on the categories of the sensory objects and sensory concepts. These sensory categories should be capable to grasp the essence of the multisensory data and build the new structure of multisensory knowledge. In machine perception MU the categories of the sensory objects and sensory concepts were introduced to deal with the problem of the data that come from the different sensory channels such as the auditory channel or tactile channel [1–3]. However, in comparison to the human sensory channels the machine can use the different sensory data that are interpreted in the context of the hierarchical structure of the sensory object categories. When machine perception MU is concerned with the machine interpretation of any sensory data, in this book only the visual perception is presented. Reason for this is that the machine perception MU is based on the very different approach than classical machine perception, and in order to present the material concerning the problems solved within machine perception MU research program, there is a need for the very selective arrangement of the presented topics. In this book, the focus is on the visual machine perception and in the following chapters problems concerning our new approach to machine perception will be presented.

For the most practiced discriminations the perception seems to proceed in the specific way. When a percept, even if of a very sophisticated nature, is highly practiced or very important it appears that our minds build up a special-purpose mechanism solely for that purpose. Human and animal brains are structured in a modular way, with different areas processing different kinds of sensory information. Some of these modules take the form of sensory maps, mapping some aspect of the world across part of the brain's surface. These different modules are interconnected and influence each other. For instance, taste is strongly influenced by smell.

Machine perception MU similarly to Marr's approach is regarding perception as a series of progressively more sophisticated inferences, and it is assumed that that there are separate, specialized mechanisms for primitive and sophisticated perception. Machine perception MU is based on the learned knowledge and the well-developed visual problems solving skills. Under these assumptions, to test perceptual ability of a machine is to test if the machine can solve the given perceptual problem. Many visual problems that are solved during visual perception are formulated as the special tasks of the visual intelligence tests. These problems are given in the form of the diagrammatic representations that refer to one of the ontological line-drawing categories.

The organization of the book follows the overview presented in this Introduction. The first chapter of this book presents a short survey of the philosophical inquires and psychological research in the human visual perception. In Chap. 2 the short survey of the literature on research in machine perception is presented. In the second part of this chapter the main points concerning the proposed new machine perception MU approach is outlined. In Chap. 3 the shape classes, introduced by the authors, as the basic visual categories used during the visual thinking and visual reasoning process, are presented. Shape classes are represented by the symbolic names and the description of the class refers to the visual objects such as geometrical figures. In Chap. 4 the 3D object classes that are the natural extension of the shape classes are presented. The 3D object classes, similarly like the 2D shape classes, are established based on general object attributes such as homotopy, convexity, thickness. They play the similar role in learning of the visual concept of the 3D object as the shape classes play in learning of the visual concept of the 2D object. In Chap. 5 picture classes that play an important role in machine perception are described. Perception in machine perception MU is related to the perceptual visual field where the perceived object is regarded as an image. The picture classes were introduced based on the concept of the generic class of pictures (see Les [5]). The picture classes are divided into two categories, the structural and ontological picture classes and the structural picture class refers to the structural organization of the picture plane that in visual art is often called the picture composition. In Chap. 6 the perceptual transformations that play an important role in transforming visual data into symbolic form are explained. The perceptual transformations are regarded as the special geometrical transformations that are applied in the domain of the geometrical figures and are divided into: the problem solving perceptual transformation (PS-perceptual transformation), the interpretational perceptual transformation (IN-perceptual transformation), the class forming perceptual transformation (CF-perceptual transformation), the generating object perceptual transformation (GO- perceptual transformation) and the grouping perceptual transformation (GR- perceptual transformation). In Chap. 7 the perceptual transformations are described in terms of the reasoning process that consists of the consecutive stages of reasoning, and at each stage of reasoning the different processing transformations are applied to transform visual data into the symbolic form. The image is assigned to the structural and ontological picture classes by application of the perceptual transformation that utilizes one or more than one processing transformations that perform the perceptual operations during the reasoning process. In this book processing transformations are presented in the visual form rather than using pseudo mathematical notation that in many computer vision literatures obscures the clarity of the presented material. The reasoning process is also presented in the new form by showing only the stages of application of the perceptual transformation (processing transformation) and visual illustration of the reasoning process. In Chap. 8 the perception as the solving of the perceptual problems is presented. In machine perception MU, the perceptual problems are divided into two categories: the interpretational perceptual problems and the visual problems. The interpretational problem that is solved by application of the IN-perceptual transformation is one of the problems that are connected with seeing and understanding. Visual problems that are given by the sequence or the set of objects, the visual representatives of the problem, and command given in the linguistic form are solved by application of the PS-perceptual transformation. Depending on the level of processing, the perceptual problems are divided into the low-level perceptual problems, the middle-level perceptual problems or the higher-level perceptual problems. The higher level interpretation of the image is based on the notation of the visual concept. In this book, the term 'concept' is used to represent the internal representation of the perceived object, whereas the category is regarded as the model of the object that is part of the hierarchical structure of categories. In Chap. 9 are presented the visual intelligence tests that are used to test the ability of a machine (SUS) to solve the different perceptual problems, at any perceptual level. The following categories of visual intelligence tests were introduced: the category of the visual analogy text-tasks, the category of the 'odd one out' text-tasks, the category of matrix text-tasks, and the category of 'what comes next' text-tasks.

We are aware that this book could be written in a different way where some issues could be explained in more detail or presented in the different way. We would like to explain that this book was written in the 'special' conditions. During the most crucial part of writing of this book, we were distracted, and it was nearly impossible to continue work connected with preparation of this book in Australia. We were forced to leave Australia to be able to continue our research and to finish all works connected with preparation of this manuscript. It was discrimination and racism that we experienced during our living in Melbourne that caused a very big financial damage to us as well as the damage to our health.

## References

1. Les Z, Les M (2008) Shape understanding system. The first steps toward the visual thinking machines. Studies in Computational Intelligence. Springer-Verlag, Berlin
2. Les Z, Les M (2013) Shape understanding system—knowledge implementation and learning. Studies in Computational Intelligence. Springer-Verlag, Berlin

3. Les Z, Les M (2015) Shape understanding system—machine understanding and human understanding. Studies in Computational Intelligence. Springer-Verlag, Berlin
4. Nevatia R (1982) Machine perception. Prentice-Hall, Englewood Cliffs, N.J.
5. Les Z An aesthetic evaluation method based on image understanding approach. In: The First International Conference on Visual Information Systems VISUAL'96, Melbourne, 5–6 February 1996. VUT, pp 317–327

# Contents