JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. 18 No. 1 (2010), pp. 75–85

The Role of Artificial Intelligence, Knowledge and Wisdom in Automatic Image Understanding

Piotr S. Szczepaniak[1], Ryszard Tadeusiewicz[2]

[1]*Institute of Information Technology*
Technical University of Lodz
Wólczańska 215, 90-924 Łódź, Poland
piotr@ics.p.lodz.pl

[2]*AGH University of Technology*
al. Mickiewicza 30, 30-059 Kraków, Poland
rtad@agh.edu.pl

Abstract. *In the paper, the roles of intelligence, knowledge, learning and wisdom are discussed in the context of image content understanding. The known model of automatic image understanding is extended by the role of learning. References to example implementations are also given.*
Keywords: *artificial intelligence, automatic image understanding, learning, wisdom.*

1. Introduction

Automatic analysis of images has a long history. A lot of methods based on diverse approaches have been described in the literature; they usually operate on a pixel level. The problem becomes more difficult if the question about the image content needs to be answered. In this paper, we address the relation between the concept of automatic image understanding [1] the conditions that need to be fulfilled for the system to be called artificially intelligent [10].

2. Basic concepts

Let us start with the general question what by definition *intelligence*, *artificial intelligence*, *knowledge*, *understanding*, and *wisdom* are. We are interested in explaining what the above named concepts mean for image understanding and creating the basic paradigm for automatic image understanding.
Intelligence: a specific set of mind capabilities which allow the individual to use the acquired knowledge efficiently and to behave appropriately in the presence of new tasks and living conditions [1].

The encyclopedia refers the reader to the term *thinking*, which is defined as any sequence of conscious psychological processes, in other words *the activity of mind* which takes place only when the individual comes across a situation requiring a certain task to be solved (either theoretically or practically) to which he has no ready solution, either instinctive or learned. Thus, it is an active cognitive process within which determination of dependencies between the elements of the reality under examination makes it possible to solve a problem faced by the thinking individual. The nature of the activity of the cerebral cortex is defined as analytical-synthetical. What is specially strongly emphasised is the connection between thinking and speaking (language), as the development of either of these abilities is believed to be impossible in separation from the other.

Kernels of the definition of human intelligence are:
Intelligence is *a human mind capability*; it integrates *instinct and thinking*; it uses *primitive and acquired knowledge*; it allows *solving of new problems*.

Knowledge: awareness or familiarity gained by experience; a theoretical or practical understanding of a subject [2].

Understanding: perceiving the meaning of something; perceiving the significance or explanation or cause of something [2].

Wisdom: experience and knowledge together with the power of applying them critically or practically [2].

If we aim to create something that is artificial, then, according to [2], our aspiration is to imitate the natural original. Here, we aim at the construction of a machine or another system which in some in a sense manifests intelligence and is able to solve problems for which we use our knowledge, understanding and experience. By analogy to the definition of human intelligence, we can argue that:

Artificial intelligence (AI) is a set of system capabilities which allow the system (machine) to use human (expert) primitive and self-acquired knowledge and to solve new tasks efficiently in the presence of new conditions of their performing.

When the system is a machine, then it integrates *a hardware construction with computational procedures*. With another kind of system, e.g. bio-chemical, the integration of *system construction with its specific processes* is imposed.

According to the original definition, one uses the notion of *computational intelligence* if the computational system uses a certain set of methodologies (cf. [3-9]). Note that this original definition is not based on semantics. The synergy of hard-technology and soft-methodologies improves the ability of machines to solve new problems and create the natural need for formulating new concepts.

Russel and Norvig [10] list the following requirements imposed on the "intelligent" computer:

- natural language processing,
- domain knowledge representation,
- reasoning mechanism,
- learning ability,

and additionally

- computer vision, and
- robotic equipment.

The researchers are interested in explaining whether it is possible to state that a machine can demonstrate any kind of intelligence (e.g.[9]). In other words, if human level of intelligence can be achieved in an artificial system. The reference point here is the human intelligence, as proposed by Alan Turing [11] - the *comparative Turing test*. In practice, for a number of tasks, such as object recognition, one reduces this general test to the determination of the percentage of correct results obtained by a machine for the test set of known cases, followed by the false-positive results, and the like.

In this work, we define the domain of interest as the *image understanding*. The proposed methodological paradigm includes knowledge representation, reasoning mechanism (cognitive resonance) and learning ability; a language of image content description can also be integrated. Within the presented concept, vision system and robotic equipment can only increase the potential application fields. Consequently, the basis for artificial intelligence is presented.

So far, it has been assumed that *image understanding* is a human specific ability to explain and interpret content of image on the basis of image perception, object recognition, knowledge, experience, reasoning, culture, association, context analysis, etc. Therefore, it is a complicated process.

3. Image understanding

The fundamental requirements of any method of automatic image understanding are:

1. Definition of the domain knowledge representation.
   A qualified professional and well defined field of interest are preferred here. For example, medical images of a certain organ and a restricted class of variations in image content.

2. Consideration of the demands and context.

3. Linguistic representation of the image content using a specialized description language.

4. Reasoning mechanism.

5. Learning ability.

Only if those five demands are fulfilled, can one speak about imitation of the human way in image analysis and in reasoning about the content.

The approach of *automatic understanding of images* described in the book [12], cf. also [17,22], is dedicated to medical applications [13-16]. However, its general idea can be adapted to other fields.

The most important difference between all traditional methods of automatic image processing and the paradigm for image understanding is that there is a feed forward data flow in the traditional methods, while in the new paradigm [12] there are two-directional interactions between signals (features) extracted from the image analysis and expectations resulting from the knowledge of image content as given by experts. When we use the traditional pattern recognition paradigm, all processes of image analysis are based on a feed-forward scheme (one-directional

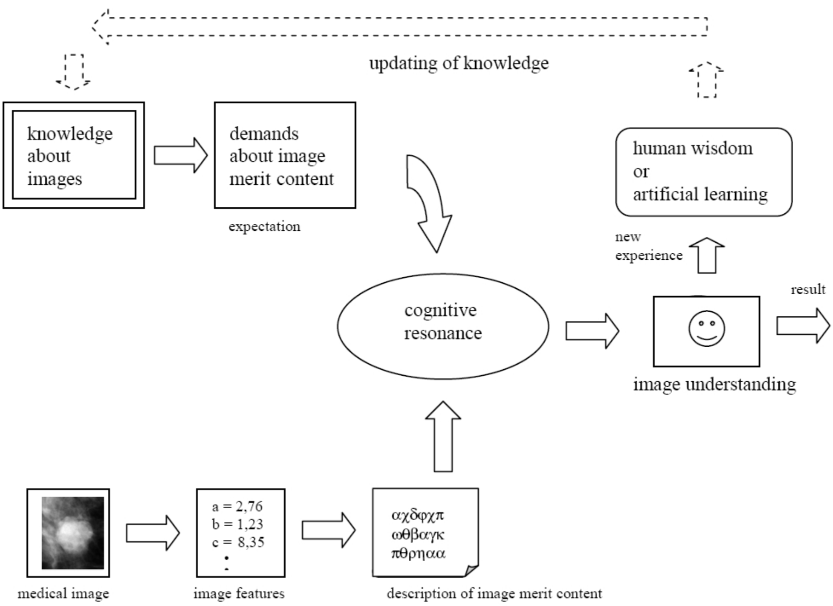

Figure 1. The main paradigm of automatic image understanding

flow of signals). On the contrary, when we apply automatic understanding of the image (1), the total input data stream (all features obtained as a result of an analysis of the image under consideration) must be compared with the stream of *demands* generated by a dedicated *source of knowledge*. The demands are always connected with a specific (selected) hypothesis of the image content semantic interpretation. As a result, we can emphasise that the proposed 'demands' are a kind of postulates describing the desired values of some selected features of the image. The demands result from the knowledge about the image contents in the considered field of application. The selected parameters of the image under consideration must have desired values (determined exactly or fuzzy) when some assumption about semantic interpretation of the image content is to be validated as true. The fact that the parameters of the input image are different can be interpreted as a partial falsification of one of possible hypotheses about the meaning of the image content, however, it still cannot be considered the final solution.

Due to this specific model of inference, the mechanism is called *cognitive resonance*. During the comparison of the features calculated for the input image and the knowledge based demands, we can observe an amplification of some hypotheses (about the meaning of the image content), while other (competitive) hypotheses are weakening. This is similar to the interferential image formed during a mutual activity of two wave sources: at some points in space waves can add to one another, in other points they are in opposite phases and the final result disappears.

Conscious experience enhances the knowledge and is fundamental for reaching wisdom. Nowadays, we do not believe, and consequently we do not expect that any artificial system can demonstrate awareness and, consequently, wisdom. Therefore, the wisdom placed on the Fig. 1 means the human wisdom becoming deeper when the human learns using the system for automatic image understanding. However, we can require from the artificial system a mathematically defined learning ability. Thus, we exchange the *human wisdom* for the system's block *learning module*.

Such a structure of the intelligent system for image understanding corresponds to the *knowledge based perception* which is one of the well known models of the natural human visual perception. The human eye recognises an object if the brain has any template for it. This holds true even when the object is shown in another way, which means that other signals are coming to the visual cortex. This is because the natural perception is mainly a mental cognitive process, based on hypotheses generation and its real-time verification. It is not just the processing of visual signals received by eyes. The verification is performed by permanent comparison of the selected features of an image with expectations taken from earlier visual experience. The presented approach to image understanding is based on the same conceptual basis, the difference being that it is performed by computers.

4. Implementation

The realisation of *automatic understanding of images* needs preprocessing, application of both low-level and high-level methods of image analysis, including the contextual one, as well as the cognitive resonance in which machine concept propositions are verified by confrontation with the hypothesis coming from the formalised human knowledge. A lot of methods of low-level image analysis ~~is~~ are known, some of them operating on different levels of granulation, e.g. two- or multi-layer Kohonen neural network [18,19].

An exemplary implementation approach is presented in [12]. It is based on the base of linguistic description of image structure. Of course languages used for such descriptions are adapted to every particular problem under consideration. There have been languages created for spinal cord images diagnosis [25][31], radiological palm diagnostics [26], foot bones analysis [30], coronary artery images [27], bone fractures analysis [28], [29], etc. After preparing of special type graph-grammar language we must also design translator for this language, engaging special attention to the crucial element of such translator, e.g. parser. As is known parsing process is a very important part in every artificial language translation system, but in cognitive analysis approach, which is considered in this paper, parsing is essential part in this part of scheme given on fig. 4, which is described as "cognitive resonance". During the parsing process the structures of objects located on the image are matched with schematic patterns working as our graph language vocabulary. Every successful matching increases the credibility of some semantic interpretations of the image merit content when other interpretations decrease their credibility. After the parsing process is completed, we obtain a kind of map with most probable semantic interpretations of the whole image.

Certain significant potential stands also behind the active contour approach. It can be used for both contextual analysis [20][21] as well for hypothesis verification [23][24].

5. Summary

Image has a specific properties: size, colour, texture, shape, etc. Those elements can be mathematically formalised and can be used for quantitative description of images. However, they do not allow one to discover the meaning of an image, not even the meaning of objects obtained as a result of segmentation. The context needs to be considered when we aim at understanding the story presented on the image. This is the reason why automatic understanding of images continues to be a challenge for science and technology. Hopefully, the present work is a step towards meeting that challenge.

Acknowledgement

This work has been supported by the Ministry of Science and Higher Education, Republic of Poland, under project number N 519007 32/0978;
decision No.0978/T02/2007/32.

References

[1] *Great Universal Encyclopaedia* (1966), Warsaw, Polish Scientific Publishers (PWN); in Polish.

[2] *The Concise Oxford Dictionary of Current English* (1990), 8th edition, Oxford, Clarendon Press.

[3] Marks II, R. J., Intelligence: Computational versus Artificial. *IEEE Trans. on Neural Networks*, Vol. 4, No. 5, pp. 737-739, 1993.

[4] *Computational Intelligence: Imitating Life.* Żurada, J. M., et al. (Eds.), (1994), Piscataway, IEEE Press.

[5] Bezdek, J. C., On the Relationship between Neural Networks, Pattern Recognition and Intelligence. *The International Journal of Approximate Reasoning*, Vol.6, pp. 85-107, 1992.

[6] Bezdek, J. C. *What Is Computational Intelligence ?* In: Reference [4] of this paper, pp. 1-12.

[7] Eberhart, R. C., Computational Intelligence: A Snapshot. In: *Computational Intelligence: A Dynamic System Perspective*. Palaniswami, M., et al. (Eds.), New York, IEEE Press, 1995.

[8] Pedrycz, W., *Computational Intelligence. An Introduction.* Boca Raton, New York, CRC Press, 1998.

[9] Kasabov, N., Evolving Intelligence in Human and Machines: Integrative Evolving Connectionist Systems Approach. *IEEE Computational Intelligence Magazine*, Vol. **3**, no 3, pp. 23-37, 2008.

[10] Russel, S. J., Norvig, P. (1995, 2002): *Artificial Intelligence. A Modern Approach*. Englewood Cliffs, Prentice Hall.

[11] Turing, A., Computing Machinery and Intelligence. In: *Computers and Thought*, Feigenbaum, E. A., and Feldman, J. (Eds.), New York, McGraw-Hill, 1963.

[12] Tadeusiewicz, R., Ogiela, M. R., *Medical Image Understanding Technology.* Series: Studies in Fuzziness and Soft Computing, Vol. 156, Springer-Verlag, Berlin, 2004.

[13] Ogiela, L., Tadeusiewicz, R., Ogiela, M. R., *Cognitive Approach to Visual Data Interpretation in Medical Information and Recognition Systems*, In: Nanning Zheng, Xiaoyi Jiang, Xuguand Lan (Eds.): *Advances in Machine Vision, Image Processing, and Pattern Analysis.* Lecture Notes in Computer Science - LNCS 4153, Springer-Verlag, Berlin, Heidelberg, New York, pp. 244-250, 2006.

[14] Ogiela, L., Tadeusiewicz, R., Ogiela, M. R., *Cognitive Computing in Intelligent Medical Pattern Recognition Systems*, In: De-Shuang Huang, Kang Li, Irwin, G. W. (Eds.): *Intelligent Control and Automation.* Lecture Notes in Control and Information Sciences, Vol. 344, Springer-Verlag, Berlin, Heidelberg, New York, pp. 851-856, 2006.

[15] Ogiela, M. R., Tadeusiewicz, R., Ogiela, L., *Graph Image Language Techniques Supporting Radiological, Hand Image Interpretations*, Computer Vision and Image Understanding, Elsevier, Vol. 103/2, pp. 112-120, 2006.

[16] Ogiela, M. R., Tadeusiewicz, R., Ogiela, L., *Image Languages in Intelligent Radiological Palm Diagnostics*, Pattern Recognition, Vol. 39/11, pp. 2157-2165, 2006.

[17] Tadeusiewicz, R., Ogiela, M. R., *Why Automatic Understanding ?* In: Beliczynski, B. et al. (Eds.): ICANNGA'2007, Part II, Lecture Notes in Computer Science - LNCS 4432, Springer-Verlag, Berlin, Heidelberg, New York, pp. 477 - 491, 2007.

[18] Tomczyk, A., Szczepaniak, P. S., Lis, B., *Generalized Multi-layer Kohonen Network and Its Application to Texture Recognition.* In: L.Rutkowski, J.Siekmann, R.Tadeusiewicz, L. A.Zadeh (Eds.): *Artificial Intelligence and Soft Computing* ICAISC 2004. Lecture Notes in Artificial Intelligence - LNAI 3070, Springer-Verlag, Berlin, Heidelberg, New York; pp. 760-767, 2004.

[19] Lis, B., Szczepaniak, P. S., Tomczyk, A., *Multi-layer Kohonen Network and Texture Recognition.* In: P. Grzegorzewski, M. Krawczak, S. Zadrożny (Eds.): *Soft Computing Tools, Techniques and Applications.* Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2004.

[20] Tomczyk, A., *Active Hypercontours and Contextual Classification.* Proceedings of the 5th International Conference on Intelligent Systems Design and

Applications - ISDA'2005. Wroclaw, Poland, IEEE Computer Society Press, pp. 256-261, 2005.

[21] Tomczyk, A., Szczepaniak, P. S., *On the Relationship between Active Contours and Contextual Classification.* In: M.Kurzyński, et al. (Eds.): *Computer Recognition Systems.* Proceedings of the 4th Int. Conference on Computer Recognition Systems - CORES'05. Springer, Berlin, Heidelberg, New York, pp. 303-310, 2005.

[22] Tadeusiewicz, R., Szczepaniak, P. S., *Basic Concepts of Knowledge-Based Image Understanding.* In: N. T. Nguyen, G. S. Jo, R. J. Howlett, L. C. Jain (Eds.) *Agent and Multi-Agent Systems: Technologies and Applications.* LNAI 4953, Springer-Verlag, Berlin, Heidelberg, pp. 42-52, 2008.

[23] Szczepaniak, P. S., *Contribution of Hypercontours to Multiagent Automatic Image Analysis.* In: N. T. Nguyen, G. S. Jo, R. J. Howlett, L. C. Jain (Eds.) *Agent and Multi-Agent Systems: Technologies and Applications.* LNAI 4953, Springer-Verlag, Berlin, Heidelberg, pp. 647-653, 2008.

[24] Szczepaniak, P. S., *Verification of Hypothesis about Image Content using Active Contour Approach.* IEEE Proc. 7th Int. Conference "Computer Information Systems and Industrial Management Applications". Ostrava, Czech Republic. IEEE Computer Society, Los Alamitos, Washington, Tokyo, pp. 207-209, 2008.

[25] Ogiela, L., Tadeusiewicz, R., Ogiela, M. R.: *Cognitive techniques in medical information systems.* Computers in Biology and Medicine, Vol. 38, No. 4., 2008, pp. 501-507.

[26] Ogiela, M. R., Tadeusiewicz, R., Ogiela, L.: *Image Languages in Intelligent Radiological Palm Diagnostics*, Pattern Recognition, Vol. 39/11. 2006, pp. 2157-2165.

[27] Ogiela, M. R., Tadeusiewicz, R.: *Syntactic reasoning and pattern recognition for analysis of coronary artery images*, International Journal of Artificial Intelligence in Medicine (Elsevier), Vol. 26, No. 1-2, 2002, pp. 145-159.

[28] Tadeusiewicz, R., Ogiela, M. R.: *Medical Pattern Understanding Based on Cognitive Linguistic Formalisms and Computational Intelligence Methods*, In: Jun Wang (ed.): 2008 IEEE World Congress on Computational Intelligence

WCCI, IEEE Catalog Number: CFP08IJS-CDR, ISBN: 978-1-4244-1821-3, ISSN: 1098-7584, IEEE Piscataway 2008, pp. 1729-1733.

[29] Tadeusiewicz, R., Ogiela, L.: *Selected Cognitive Categorization Systems*, Chapter in book: Rutkowski, L. et al. (Eds.): Artificial Intelligence and Soft Computing, ICAISC 2008, Lecture Notes on Artificial Intelligence, Vol. 5097, Springer-Verlag Berlin Heidelberg 2008, pp. 1127-1136.

[30] Ogiela, L., Tadeusiewicz, R., Ogiela, M. R.: *Cognitive Approach to Medical Pattern Recognition, Structure Modelling and Image Understanding*, In Yenghong Peng and Yufeng Zhang (Eds.): Proceedings of International Conference on BioMedical Engineering and Informatics BMEI 2008, CPS-IEEE, (ISBN 978-0-7695-3118-2, DOI 10.1109/BMEI 2008.118) Los Alamitos, California, Washington, Tokyo, 2008, pp. 33-37.

[31] Ogiela, L., Tadeusiewicz, R., Ogiela, M. R.: *Cognitive Categorizing in UBIAS Intelligent Medical Information Systems*, in Sordo, M., Vaidya, S., Jain, L. C. (Eds.): Advanced Computational Intelligence Paradigms in Healthcare 3, Studies in Computational Intelligence 107, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 75-94.