Machine Understanding - a new area of research aimed at building thinking/understanding machines

Zbigniew Les and Magdalena Les

St. Queen Jadwiga Research Institute of Understanding,
Melbourne, Australia
sqjriu@outlook.com

***Abstract*—In this paper machine understanding, that is referring to a new area of research the aim of which is to investigate the possibility of building a machine with the ability to think and understand, is presented. Machine Understanding, the term introduced by the authors to denote understanding by a machine, is the first attempt to establish the scientific method to investigate the complexity of understanding problem, and is based on the results of philosophical investigations and assumptions of the logical positivists. Machine Understanding, defined in the context of both human understanding and existing systems that can be regarded as the simplest understanding systems, is based on the development of the shape understanding system (SUS) and 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. Machine understanding refers to the categorical structure of learned knowledge and one of the most complex problems that is solved within this framework is understanding of visual objects (visual understanding). In this paper only some aspects of visual understanding, as examples of understanding process, are presented. The first stage of visual understanding involves perceptual reasoning that consists of the perceptual categorical reasoning and visual reasoning. The visual reasoning consists of assigned reasoning that assigns the perceived object to one of the shape categories. The assigned reasoning consists of the consecutive stages of reasoning where at each stage of reasoning the specific data are acquired based on the results of the reasoning at previous stages.**

Keywords—machine understanding; visual understanding; visual thinking; perceptual reasoning; assigned reasoning; image understanding,

I. Introduction

This Wiener's book [22] that laid the theoretical foundation for servomechanisms, analog computing, artificial intelligence and neuroscience exploits an old paradigm where trust in mathematical modeling is the basis for development of the scientific approach in explaining natural phenomena and designing complex machines. However, there are very complex problems such as understanding or thinking, where this approach cannot be applied and for this reason the new area of research the aim of which is to investigate the possibility of building a machine with the ability to think and understand was proposed by authors [11], [12], [13]. Machine Understanding is defined in the context of both human understanding and existing systems that can be regarded as the simplest understanding systems. The problems related to human understanding were discussed in a more detail in [11], [12], [13] and in this paper only main points of these problems are presented and most of references can be accessed from our books. Understanding, the result of thinking, involves processes such as learning, problem solving, perception, and reasoning, and requires abilities such as intelligence. There is no unique definition of human understanding and comparison of human understanding and machine understanding is based on the results of philosophical investigations, and not on the results of scientific research. Some problems related to human understanding are topics of research in the area of psychology, linguistics, cognitive science or artificial intelligence, however there are also problems that are not subjected to scientific methodology (empirical research). Human understanding was differently defined during the long period of philosophical inquiries. One view is that perceived object and idea are a key to understand human understanding (Plato, Aristotle, Lock, Berkeley, Leibnitz or Kant (see e.g. [5], [14], [6], [1], [10], [9]). For Plato [19] understanding is grasping of ideas and the idea refers to particular things in the empirical world that are imperfect reflections of that idea. For Aristotle understanding is connected with perception were ideas (concepts) are extracted from perceived data based on the abstraction and generalization. For Locke [14] understanding is grasping of the relations between ideas and for Kant [9] understanding begins by means of objects which affect our senses, produce representations, rouse our powers of understanding into activity (to compare, to connect, to separate) and to convert the raw material of our sensuous impressions into the knowledge of objects (ideas). For Husserl [16] meaning of the object is a key for understanding. Husserl, when still absorbed with an object, pointed to the meaning of the object as its essential cognitive ingredient. He introduced distinction between natural and phenomenological modes of understanding. Natural understanding is based on the perception that constitutes the known reality whereas phenomenological understanding is based on phenomenological reduction that is based on consciousness of any given object that discerns its meaning as an intentional object. For Frege, Wittgenstein and Russell language is a key for understanding and formal language and mathematical modeling were an important components of understanding. For Russell understanding is connected with searching for an ideal language for representing the scientific facts and Wittgenstein [23] developed a comprehensive system of logical atomism as a formal language of science. For analytic philosophy (logical positivism) [16] understanding is based on logical clarification of thoughts by analysis of the logical form of philosophical propositions and using formal logical methods to develop an empiricist account of knowledge. Logical positivists adopted the verification principle according to which every meaningful statement is either analytic or can be verified by experiment, and rejected many traditional problems of philosophy as meaningless. For hermeneutics philosophers (Schleiermacher, Gadamer, Heidegger) interpretation of the text is a key to understanding. Hermeneutics, as the art of understanding the written discourse of another person correctly, was initially applied to the interpretation of scripture and emerged as a theory of human understanding through the work of Schleiermacher and Dilthey. For Schleiermacher understanding of the text is to find the author's intentions whereas for Gadamer [4] the context of interpretation determines a text's meaning and reveals something about the social context in which texts were formed. For philosophers such as Hobbes or Spinoza brain and its functioning is a key for understanding. They believed that humans are deterministic machines with understanding explainable by scientific methods. Modern philosophers (logical behaviorism or functionalism) regarded the problem of understanding as the problem of mind functions. Functionalism identifies mental states with brain states and explains understanding in terms of cognitive theory that tried to explain human understanding by comparing the mind to a sophisticated computer system.

Machine understanding is defined in the context of both human understanding and existing systems that can be regarded as the simplest understanding systems. Simple understanding systems are built in the areas of expert systems, image understanding, language understanding, or robotics. Expert systems [7] are computer systems that emulate the decision-making ability of a human expert and are the first computer systems that solve problems that require understanding of the selected fragments of knowledge. The term image understanding [2], [17], [21] refers to a computational information processing approach to image interpretation and knowledge-based interpretations of visual scenes that transform pictorial inputs into commonly understood descriptions or symbols. Image understanding systems, built in order to interpret the perceived object or interpret an image, are based on the research in the area of computer vision and image understanding. Language understanding is an area of research that deals with understanding of a text as the product of the linguistic activity of the mind. The natural language understanding systems [20], [18], [15], [8], [3] usually consist of the subsystems that perform the specific tasks such as lexical analysis, syntactic analysis, semantic analysis, discourse analysis or pragmatic analysis. Neural Networks [24], a set of simple computational units (nodes, neurons) that are highly interconnected, which attempts to model the capabilities of the human brain, can be also regarded as the simplest understanding systems. The most popular neural network in pattern recognition is the feed forward multilayered network, with the back propagation algorithm as the training method.

II. MACHINE UNDERSTANDING

Machine understanding refers to the new area of research the aim of which is to investigate the possibility of building a machine with the ability to think and understand. The term machine understanding, introduced by the authors [13], denotes the process of understanding by the machine SUS (Shape Understanding System). A machine, in order to be able to understand, needs to imitate the way in which humans understand the world and language (text). SUS as the machine that is designed to have an ability to think and understand needs to learn both knowledge and skills. Learning knowledge and skills, which supplies material for thought that leads to understanding, is called knowledge implementation (see [12] for description). Machine understanding stresses the dependence of learning and understanding processes and 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. It is assumed that to understand means to be able to solve a problem and to give the relevant explanations.

Machine understanding refers also to the categorical structure of learned knowledge. The shape categories, presented in [11] and [12], are basis for the intuitive grasping of the sense of perceived objects whereas the basic abstract categories, described in [13], are applied during abstraction in the problem solving when the perceived object is assigned (transformed) into the visual general abstract categories such as the circle category or the rectangle category, and into the basic abstract categories such as the object category or the movement category. Machine understanding refers to different ontological categories of objects: a visual object, a real world object, a sign, a sensory object, or a text object, described in [12], [13]. The sensory object category is a special category derived from the category of visual objects. A sensory object is the object that is named based on a set of measurements that refer to the attributes of the category to which the object is assigned e.g. the category of mineral objects. The text category is referring to any form of the text and is divided into four different specific categories: the text-query category, the text-task category, the dictionary-text category, and the long-text category (see [12], [13] for description). Machine understanding, following the way of scientific understanding, is based on the basic abstract categories such as the set category, the element category, or the belonging category that are defined in the area of set theory based on adopted axioms, as described in [13]. The basic abstract categories are represented as the objects on the SUS normalized perceptual visual field (the rectangle on which all perceived objects are projected). These visual representations that refer to the SUS intuition can be utilized during explanatory process and make it possible to found understanding on the strong intuitive basis.

Machine understanding is based on the assumption that the results of understanding by a machine can be evaluated and compared to the results of human understanding. If understanding is defined as the ability to solve problems, then assuming that problems (tasks) are well defined, the understanding (ability to understand) can be tested by testing whether these problems can be solved by the machine (SUS). In this context machine understanding can be regarded as problem solving, however it is assumed that the machine to be able to understand needs also the ability to explain how to solve a problem. The most important part of evaluation of the machine's (SUS) ability to understand is to formulate the problems and to use these problems to test if the machine (SUS) is able to solve those problems. Examples of problems that are solved during machine understanding (problem solving) such as the naming, solving the visual problems (perceptual problems, visual analogy problems or spatial problems), the problems of the signs interpretation, the problems of text interpretation and explanatory problems, are described in [13]. Machine understanding is regarded as the problem solving and explanation can be also regarded as solving the problem of explaining known facts, perceived objects, solved tasks or interpreted texts. The special class of problems used for testing the results of learning at school (text-tasks), is described in [11], [12], [13]. However, in order to test if the text presented to SUS is understood there is a need to formulate the special text-tasks in the form of questions, computing problems or explanatory problems. For example, in order to test the degree of understanding of the mineralogical dictionary text SUS can be asked the questions "what is the name of the mineral that is represented by chemical formula ZnO"? or "explain why malachite is green"?

III. VISUAL UNDERSTANDING IN MACHINE UNDERSTANDING

Machine understanding refers to the categorical structure of learned knowledge and one of the most complex problems that is solved within this framework is understanding of a visual object (visual understanding) that is based on the development of the Shape Understanding System (SUS). In this paper some aspects of visual understanding are presented to show the complex reasoning process. The first stage of visual understanding involves perceptual reasoning that consists of perceptual categorical reasoning and visual reasoning. The visual reasoning consists of the assigned reasoning that assigns the perceived object to one of the shape categories. In contrast to existing approaches in AI, where usually reasoning is independent of the acquired data needed in reasoning process, assigned reasoning consists of the consecutive stages of reasoning where at each stage of reasoning the specific data are acquired based on the results of reasoning at previous stages. The visual reasoning usually assigns the name from one of the visual object categories to the perceived object. The visual object, after naming, is interpreted based on knowledge of ontological visual categories and knowledge of the knowledge scheme. Categories of visual objects are established based on the assumption that a visual object exists and can be perceived by the accessible technical tools (see [11]). The notation of basic categorical knowledge is based on a categorical chain. The categorical chain is a series of categories, derived from the categories of visual objects or categories of body of knowledge, showing the hierarchical dependence of knowledge. The categorical chain derived from the categories of visual objects is given as follows: $\nu_O \supset \pi \supset \sigma \supset \nu \supset \nu .... \supset \{\nu \ ,\nu \ ,....\}$, where the categories are derived from the category of visual objects ν_o. The category at the first level of the categorical chain is called the perceptual category π of a visual object. The category at the second level of the categorical chain is called the structural category σ of the visual object. The ontological category ν begins from the third level of the categorical chain. The symbol $\supset$ denotes moving to the next level of the categorical chain and $\{\nu \ ,\nu \ ,....,\}$ denotes different categories at the same level of the categorical chain. The perceptual categories and structural categories are associated with the visual appearances of objects and are represented by visual knowledge. The structural element category can represent both the visual and sensory objects. Knowledge of the specific category derived from the given visual category such as the symbol category is learned by SUS at the prototype level $... \supset \nu \supset ... \supset \nu$. Ontological visual categories have hierarchical structure and at the bottom of each categorical chain is the prototype category ν . The prototype is defined during learning process at the level for which the training exemplars are available. The prototype is represented by all visual representatives of the specific category and it is assumed that learned visual knowledge is covering the visual domain prototype.

Understanding of a visual object (visual understanding) is a very complex problem and involves perceptual reasoning that consists of perceptual categorical reasoning and visual reasoning. The perceptual reasoning is applied during naming and learning processes. Higher level understanding processes involve the reasoning that is based on the previously learned non-visual knowledge. The visual reasoning is part of the higher level visual reasoning used in the naming process where all learned non-visual knowledge that is connected with the category to which the name is assigned to the object is accessible. During the visual reasoning (naming) the name of one of the learned visual categories is assigned to the perceived object. Naming of the visual or sensory objects is to solve the problem of finding the meaning of the objects. When the object is named its meaning consists of all learned knowledge that is linked to the category to which the named object belongs. For example, understanding signs is to solve the problem of finding the meaning of signs or symbols and to solve this problem the interpretation that is based on the learned coding system is utilized. Similarly, understanding a text is to solve the problem of finding the meaning of the text [12], [13].

The perceptual categorical reasoning, the first stage of the perceptual reasoning, is related to the SUS perceptual visual field where an object is assigned to one of the perceptual and structural categories. The perceptual category reflects perceptual properties of the object, determines the visual reasoning process and is divided into a silhouette, a line-drawing, a colour object, or a shaded object. The method of assigning of the object to one of the perceptual categories, based on the histogram, depends on a number of peaks in histogram. The object is assigned to the line drawing or silhouette category if the histogram has one peak, it is assigned to the colour category if histogram has more than two clearly visible peaks and it is assigned to the shaded category if histogram has no clearly visible peaks. An object is assigned to the line drawing category if the object is assigned to the thin class category. In a similar way the object is assigned to one of the structural categories. The structural category refers to the complexity of the visual representations of an object and is divided into the element category, the pattern category, the picture category or the animation category. Assigning to one of the pattern categories is based on computation of a number of objects. An object from the pattern category is the object that is assigned to the silhouette, line drawing or colour category. An object from the shaded perceptual category is usually assigned to the picture category. The element category is the basic structural category that is used during naming of an object.

The shape categories (classes) are derived based on the visual attributes of the visual object that refers to the geometrical and topological properties of the object. The shape categories (classes) are described in [11], [12]. An important part of visual understanding is perceiving of a 3D object and interpreting the object in terms of the 3D geometrical figure or in terms of the real world object. Understanding of the object from the real world ontological category requires usually interpreting it as a 3D object. SUS understands a real world object as the objects extracted from an image in the SUS perceptual visual field. The real world object is usually extracted from the object that is assigned to the picture category. SUS can only differentiate between a real world object and the photograph (picture) of this object by obtaining additional sensory information. There is the assumption of intentionality, that means, SUS knows (assumes) that the photograph is the image obtained by looking at the real world object. The different backgrounds require applying the different segmentation methods to extract the object from the background. Knowing an object (name of object) that we are looking for makes the searching for the object and extracting it more easy task.

As it was described, the first part of the visual understanding involves perceptual reasoning that consists of perceptual categorical reasoning and visual reasoning. The visual reasoning consists of the assigned reasoning that assigns the perceived object to one of the shape categories and is based on the shape understanding method [11]. A member of the shape category is called an archetype. The archetype ω of the class Ω ,$(\omega \in \Omega)$ is an ideal realization of the shape (visual object) in the two-dimensional Euclidean space E^2. An exemplar $e \in E$ of the class Ω is a binary realization of an archetype in the discrete space. The exemplar is one of the regions of a binary image. The binary image is regarded as a set of pixels on the discrete grid (i,j). The visual object o_i , that is perceived by SUS, is transformed by the perceptual transformation Δ: $\Delta(o) \Rightarrow u$ into the phantom u that is the 2D representation (e.g. photograph) of the object o_i . The phantom u is transformed into a set of critical points $\coprod$ by the sensory transformation $\Im$: $\Im(u) \Rightarrow \coprod$ and next into a symbolic description in the form of a string $R\langle\coprod\rangle \Rightarrow \kappa$, and finally into a symbolic name $K(\kappa) \Rightarrow \eta$.

The assigned reasoning is the most important part of the perceptual reasoning. The assigned reasoning consists of the consecutive stages of reasoning where a perceived object is at first transformed into a set of critical points $\coprod$ and next into the symbolic name η. In order to fulfill the required task of acquiring the data and processing it in order to obtain a set of descriptors $\Im$, a processing method Φ is used. The processing method applies an image transformation Θ in order to transform the data into one of the data types. The image transformation Θ is the mapping from the one set, called the domain of mapping, into another one called a set of mapping values. The descriptor transformation $\aleph$ is applied to find a set of descriptors ι used to assign the perceived object to one of the possible classes Ω^{η}. A reasoning process that is part of a visual reasoning process is performed passing the consecutive stages of reasoning. During each stage of reasoning a sequence of image transformations is applied in order to find a set of descriptors. The sequence of image transformations $\Theta_{\coprod}^{\lambda} : \coprod -> \coprod$ that are used in reasoning process can be written as: $\lambdabar_{\alpha_1} : \coprod^{\alpha_0} \rightarrow \coprod^{\alpha_1}$ $\lambdabar_{\alpha_2} : \coprod^{\alpha_1} \rightarrow \coprod^{\alpha_2}$,…., $\lambdabar_{\alpha_M} : \coprod^{\alpha_{M-1}} \rightarrow \coprod^{\alpha_M}$ or as a composite given as $\Theta_{\alpha_1} \bullet \Theta_{\alpha_2} \bullet ... \bullet \Theta_{\alpha_M} : \coprod^{\alpha_0} \rightarrow \coprod^{\alpha_M}$, where Θ_{α_1} denotes one of the image transformations and $\bullet$ denotes the sequential operator. Although it was assumed that a visual object is represented by a binary image it is not the cause of a serious limitation to the presented method. The visual object that consists of parts of different colours is assigned into one of the colour classes and during processing stages these parts are interpreted as the new visual objects. The assigned reasoning involves transformation of the description of an examined object s when passing stages $\varsigma_0 \rightarrow \varsigma_1 .... \rightarrow \varsigma_N$, where ς_0 is the beginning stage, ς_N is the final stage of the reasoning process and $\rightarrow$ denotes the move to the next stage of reasoning. At each stage of the reasoning ζ_i the following operations are performed: the processing transformation transforms the set of critical points $\lambdabar_2 : \coprod^1 \rightarrow \coprod^2$, the descriptor transformation computes descriptors $\iota_2 = \aleph_2(\coprod^2)$, an examined object s is assigned to one of the possible classes $[\iota_2 < T_2] \Rightarrow s \succ \Omega[\varsigma_2]$. During reasoning process, a perceived object is first transformed into a set of critical points $\coprod$ and next into the symbolic name η . The symbolic name η is extracted from a symbolic description κ . The symbolic description κ is an intermediate form that has many additional specific data about the perceived phantom. The symbolic description κ_k is used to reason about the specific categories to which the object can belong. For example, the object O1  is transformed into a symbolic description in the form of the following string"[A3][[|L3|AE|]|S79||B100,99,99||A60,61,60||G248||@2691|]]{[|L3|O|]|S52||B58,100,57||A29,30,120||G76||@395|]}{[|L3|O|]|S52||B57,100,58||A30,29,120||G76||@396|]}{[|L3|O|]|S53||B100,58,57||A29,120,30||G76||@417|]}". Next, the symbolic description is transformed into the symbolic name given as the string A3_L3_AE_L3_O_L3_O_L3_O.

Machine understanding is strictly connected with learning of new knowledge. SUS ability to understand depends on the effectiveness of learning process and learning of new knowledge depends on the SUS ability to understand. SUS learning is called the knowledge implementation (see [12]) and is concerned with two main aspects of human learning: learning of the visual knowledge in the context of the categorical structure of the learned categories of visual objects and learning of the knowledge that is connected with understanding of the content of the text. Process of learning consists of acquiring of the new knowledge and learning of the new skills. During learning of the visual knowledge

(knowledge implementation) the generalization, the specification, the schematization and the visual abstraction is important part of the learning of the visual knowledge. Understanding the visual objects from one of the ontological categories requires learning of the visual concepts of this category. The ontological category v_i is given by its name n^i and is represented by a set of visual objects called the visual representatives of the category $v_i(o) = \{o_1, o_2, ..., o_n\}$. Visual knowledge of the category v_i is learned as a visual concept represented as a set of symbolic names $\varphi_c = \{\eta_1, \eta_2, ..., \eta_n\}$. It is assumed that a set $v_i(o)$ represents all visual aspects of the category v_i.

During learning of the knowledge of visual objects, at first, the representative sample of objects from the category $\mathbf{u} \in \mathbf{v}$ is selected, then for each object the symbolic name η_i is obtained and finally the visual concept of this category as a set of symbolic names $\varphi_c^j(\mathbf{v}) = \{\eta_1, \eta_2, ..., \eta_n\}$ is learned. For selected category $\mathbf{v}$ the visual concept is obtained in the following stages of the learning

For all $u_i \in \mathbf{u},\ i = 1, ..., n$, $\mathbf{u} \in \mathbf{v}$ do:

1. Transform a phantom u_i into its digital representation using a perceptual transformation

$$\Im(u_i) = o_i - \text{Ш}^i .$$

For each Ш^k perform reasoning:

2. Assume $\zeta_j \equiv \zeta_0 \equiv \Omega[\zeta_0]$, $\text{Ш}_i^k \equiv \text{Ш}_i^0$.

At the j-th stage $\zeta_j \equiv \Omega[\zeta_j]$ assume that an examined object o_i is assigned to the class $\Omega[\zeta_j]$.

Apply the processing transformation: $\lambda_k : \text{Ш}_i^k \to \text{Ш}_i^{k+1}$.

Apply the descriptor transformation: $\iota_h = \aleph_h(\text{Ш}_i^{k+1})$.

Apply the rule: $[\iota_h > T_h] \Rightarrow o_i \triangleright \Omega[\zeta_{j+1}]$.

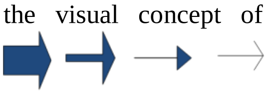
If $\zeta_j \equiv \Omega[\zeta_j]$ is the final stage, assume $\eta_i \equiv \Omega[\zeta_j]$.

If i<n, i=i+1 goto 1 else END.

else

j=j+1, goto 2.

As a result of applying this algorithm the visual concept $\varphi(v_i) = \{\eta_1, \eta_2, ..., \eta_n\}$ is obtained. An example – learning of the visual concept of members of the arrow category:



the result - a visual concept $\varphi_{Arr} = \{Q^2_{L^5}(2L^3), C^2(L^4, L^3), \Xi(\Theta^2, L^3), \Theta^3\}$.

During understanding of an object u, the perceived object u is transformed into the symbolic name η and next a learned set of symbolic names is searched to find the symbolic name of the category that was learned previously. Understanding process can be represented as:

for i=0 to K if $\eta \in \varphi^i$ then $n^i \triangleright N^i$, where $n^i \in N^i$ is the name of the i-th category and a set $N = \{n^1, ..., n^M\}$ is a set of all names of categories to which the object can belong, and M is a number of names in the set N. After naming, all non-visual knowledge, that was previously learned for the category v_i, is now accessible and can be used in the thinking/understanding process.

Understanding of an object is performed at two levels, the intermediate level and ontological level. At the intermediate level of understanding the object is described in terms of the shape classes. The description of the object at intermediate level refers to the symbolic name. For example, for the object , the symbolic name (in SUS notation) "A3_L3_AE_L3_O_L3_O_L3_O" consists of two parts. The first part "A3" gives a general description of the class that means that the object is the acyclic object with three holes. The second part "L3_AE_L3_O_L3_O_L3_O" gives a specific description of the object. The final description of the object, at the intermediate level of understanding, is given in the form of the linguistic description: "acyclic object with three holes". At the ontological level, the object is assigned to one of the ontological categories in naming process. Naming not only attaches the name to the perceived object but also "connects" the object with all knowledge that is relevant to the name of the object. Many names from different categories can be attached to the same object and naming can be given at many different ontological levels. In order to assign the object to the specific ontological category, information included in a symbolic description is used to obtain the additional data needed in the reasoning process. For example, an object  can be interpreted as a symbol "eye of dragon" when additional relation "*all three holes are equal*" is established. In the case of the object O1  the size of holes is given in the string form as |S52|, |S52|, |S53, as part of the symbolic description. The object O1 can be also interpreted as a mathematical object (solid pyramid) or as a real world object (a model of a pyramid).

In this paper only the perceptual reasoning that is the key process of visual understanding is presented. Visual understanding, however, is very complex process that involves solving visual problems such as solving visual intelligence test or visual diagnosis. The important part of the problem solving is to find a suitable form of the problem representation. The visual representation, as one of the forms of the problem representation, can be used as the problem itself (e.g. naming), as the schematic representation of the problem (e.g. solving task with electric circuits), as the imagery transformation (e.g. solving task planning robot action) or as the explanatory process (e.g. explaining a solution). Visual understanding, regarded as a problem solving process, can be described by a sequence of sub-processes and expressed as follows: $\Delta(v) \Rightarrow u$ -> $\Im(u) \Rightarrow \text{Ш}$ -> $R\langle \text{Ш} \rangle \Rightarrow \eta$ -> $T\langle \eta \rangle \Rightarrow \text{Ш}$ … $R\langle \text{Ш} \rangle \Rightarrow \eta$ -> [$\eta \in \varphi^\alpha] \Rightarrow a \triangleright \vartheta^i$, where at first 'the problem transformation'

$\Delta(v) \Rightarrow u$ transforms a given member of the problem category into the visual form (phantom), next the sequence of transformations $R\langle \amalg \rangle \Rightarrow \eta$ -> $T\langle \eta \rangle \Rightarrow \amalg$ … $R\langle \amalg \rangle \Rightarrow \eta$ transforms the internal representation given as a set of critical points into the symbolic names (image transformations), and at the end, the solution is obtained by applying the visual inference. Examples of the visual problem solving are given in [11], [12]. Non-visual problem solving such as solving the educational tasks or interpreting the text is presented in [12], [13].

IV. CONCLUSION

In this paper some aspects of visual understanding, as an important part of machine understanding that is referring to the new area of research the aim of which is to investigate the possibility of building a machine with the ability to think and understand, is presented. The result of our research shows that there is possibility to build the machine with the ability to think and understand based on the framework described in this paper and our books. Presented visual understanding, that involves the perceptual categorical reasoning and visual reasoning, makes it possible to combine the non-visual and the visual knowledge to perform complex higher level reasoning process. Some aspects of visual understanding were also presented to show the complexity of the reasoning process. In contrast to existing approaches in AI, where reasoning is usually independent on the acquired data needed in reasoning process, the assigned reasoning consists of the consecutive stages of reasoning where at each stage of reasoning the specific data are acquired based on the result of the reasoning of previous stages. As it was indicated in this paper, machine understanding not only investigates the possibility of building a machine with the ability to think and understand but also makes it possible to study the selected aspects of understanding and provides the suitable model of understanding that can be approached using scientific methods. However, it is important to stress that machine understanding can only to some extend approximate human understanding and requires very good programming skills in C++ and knowledge of algorithms from the deferent domains such as numerical methods, computational geometry, graph theory, image processing, or signal processing.

REFERENCES

[1] Berkeley G (1996) Principles of Human Knowledge and Three Dialogues. The World's Classics. Oxford University Press, Oxford
[2] Binford T (1982) Survey of Model-Based Image Analysis Systems. *International Journal of Robotics Research* 1 (1):18-64
[3] Dale R, Moisl, H., Somers H (2000) Handbook of Natural Language Processing.
[4] Gadamar H (1989) Truth and Method. Crossroad, New York
[5] Gilson E (1991) The Spirit of Mediaeval Philosophy. University of Notre Dame Press, London
[6] Hume D (1975) Enquires Concerning Human Understanding and Concerning the Principles of Morals. Oxford University Press, Oxford
[7] Jackson P (1998) Introduction To Expert Systems 3edn. Addison Wesley,
[8] Jurafsky D, Martin JH (2008) Speech and Language Processing. 2 edn. Prentice Hall,
[9] Kant I (1996) Critique of Pure Reason. The Everyman Library. Everyman, London
[10] Leibnitz GW (1996) New Essay on Human Understanding. Cambridge Texts in the History of Philosophy. Cambridge University Press, Glasgow
[11] Les Z, Les M (2008) Shape Understanding System. The first steps toward the visual thinking machines., vol 86. Studies in Computational Intelligence. Springer-Verlag, Berlin
[12] Les Z, Les M (2013) Shape Understanding System - Knowledge Implementation and Learning, vol 425. Studies in Computational Intelligence. Springer-Verlag, Berlin
[13] Les Z, Les M (2015) Shape Understanding System: Machine Understanding and Human Understanding, vol 588. Studies in Computational Intelligence. Springer-Verlag, Berlin
[14] Locke J (1961) An Essay Concerning Human Understanding. Dent, London
[15] Manning C, and Schütze, H.; (1999) Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge, MA
[16] Martinich AP (2001) A Companion to Analytic Philosophy. In. Blackwell, Oxford,
[17] Matsuyama T, and Hwang, V. (1990) SIGMA: a Knowledge-based Aerial Image Understanding System. Plenum Press, New York
[18] Mitkov R (2002) Oxford Handbook of Computational Liguistics. Oxford University Press, Oxford
[19] Plato (1993) Republic (trans: Waterfield R). Oxford University Press, Oxford
[20] Radford A (1988) Transformational Grammar: An Introduction. Cambridge University Press, Cambridge
[21] Ullman S, Richards W (1989) Image Understanding. Ablex Publishing Corporation, Norwood
[22] Wiener N (1948) Cybernetics: or Control and Communication in the Animal and the Machine. MA: Technology Press, Boston
[23] Wittgenstein L (1961) Tractatus Logico-Philosophicus (trans: Pears D, McGuinness BF). Routledge & Kegan Paul, London
[24] Zurada J (1992) Introduction to Artificial Neural Systems. West Publishing Company, St. Paul, Minnesota