Images in Logical-and-Linguistic Artificial Intelligence Systems

Boris A. Kobrinskii

Federal Research Center "Computer Science and Control" of Russian Academy of Science, Vavilova str. 44, kor.2, Moscow, 119333, Russian Federation
kba_05@mail.ru; bak@isa.ru

ABSTRACT

Visual images are holistic, and when verbalized there is a partial loss of semantic content. However, it should be noted the lack of effectiveness of decision support systems only when using images without an effective context, and systems that do not include holistic images. Inclusion of images in the knowledge base of intelligent systems can significantly improve their effectiveness. At the stage of formation of intermediate diagnostic hypotheses, the system will present to the user (physician) a hypothesis specific to the verbal and visual characteristics. At the same time, it is necessary to take into account the need to use fuzzy logic at the stages of the derivation of solutions. The subsequent process will depend on the physician's confidence in the coincidence of the image of the diagnosed patient with the image(s) in the knowledge base of the intellectual system.

Keywords: logic-and-linguistic-and-image system, holistic images, confidence factors, visual image number, fuzzy logic, orphan hereditary diseases.

1   Introduction

An image is a necessary fragment of the whole picture of the world. We can see this in many of the humanities – in archeology, history and art, for example. In medicine, images occupy a special place in the diagnosis of a number of diseases. The value of holistic images is determined by the fact that the verbal equivalent may be too complex or erroneous. This can be explained by the theory of the specialization of the "image" brain system in parallel or simultaneous image processing [1]. I don't understand why this different font thinks it is Times New Roman. I can't change it on my PC.

For example, it is extremely difficult to convey such visual images as "a skull in the form of a shamrock", "rough or 'gargoloid' facial features", or "beaky nose" in a verbal form that is understandable and can be perceived unambiguously. Metaphorical terms can be perceived differently by physicians who have never seen such images, or only in the literature, but otherwise have met them only in a linguistic form describing the appearance of these patients. In every case, there is at least a partial loss of information. This is due to the fact that it is impossible to fully reflect one-to-many relationships in metaphorical representations that contain latent polysemy (for example, 'poultry head' includes features of the shape of the skull, nose, etc.).

The defragmentation of the image with subsequent reintegration is also unacceptable, since individual features (for example, the shape of the forehead, ears, etc.) do not create an integral picture of the individual's face [2]. The verbalized image can be generally difficult or incorrectly perceived without context (for example, a rash on the skin without describing its shape, colour, size or other characteristics). At the same time, visual images help to distinguish diseases similar in their manifestations. As modifiers for the perception of images, certainty (probability) may also appear. Naturally, the full picture of the disease develops in the context of linguistic and image characteristics. Context can contain both specific ethnic features of appearance (eye section, lip thickness, etc.), and pathological manifestations on the part of various systems of the human body (colour and thickness of the skin, the nature of breathing, etc.).

Until now, in medical diagnostic systems using computational methods or approaches based on artificial intelligence, images of patients have not been used or demonstrated after making a diagnostic decision based on the analysis of the linguistic signs. Recently, various systems have been implemented in which only images are offered. However, based on our own experience, analysis of the literature and discussions with medical experts, it seems more promising to build logical-lingual-image diagnostic systems.

2 Visual images and their meaning

Consider an entity or concept as a collection of properties of a visual image. We will compare the image of the entity (for example, a beastlike face), whose properties will be called the eyes, lips, teeth, hair, etc. That is, hypertrophied facial features can in general represent the essence of the phenomenon. Analogy can serve as a mask (anger, despair, etc.; for example, threat by the Medusa jellyfish, or anger by Stromboli in a Pinocchio cartoon). Graphics can serve as a way to highlight the main message in the image, giving the "convexity" characteristic features.

At the same time, the "floating" image in memory is associated with the subject's thinking and this contributes to the physician's doubt. This is despite the fact that the image in the system of mental reasoning has the function of an intuitive representation.

Numerous variations of the representation characterize the intermediate states of the image, which can be complete or represented by separate relevant fragments, as a contour representation. It is possible to compare the individual segments (for example, the shape of the forehead, chin, ears, etc.).

From the point of view of logic, an image or object characterized by a certain content is a symbol or token. For example, an image may be a symbol from ancient times, such as a petroglyph. Nevertheless a symbol is not always an image; for example, Japanese characters are images, but Greek, Latin and Cyrillic are not. Petroglyphs can be considered as an intermediate variant of image-symbolic representations, which in the knowledge base of the intellectual system will correspond to linguistic and image symbols.

3 Image series

Images can be static or dynamic. The dynamics of the image can be associated with the different perceptions of people in different epochs and in different countries (for example, the image of a sphinx), or with changes in a human face over a lifetime or during severe illness. These dynamics can be

considered as an image series [3]. By analogy with the topology of the possible [4], the genotype in hereditary diseases can be compared with the image-archetype and a series of phenotypic images. This situation, "genotype-phenotypes", corresponds to the "one-to-many" relationship.

In the medical domain, this is particularly evident in orphaned hereditary metabolic storage diseases. Such visual phenotypic characteristics of pathological signs form in the physician's mind on the basis of observed real objects, mental images stored in their memory, although in time they can be deformed. This corresponds to the opinion of J.T.E. Richardson [5] that structures located in the occipital part of the left hemisphere of the brain play an important role in the formation of images, and the structures of the right hemisphere are associated with the processes of image transformation.

Identification of images by a person occurs by comparison with a prototype in long-term memory. The theory of sensor prototypes [6] assumes a complete correspondence between the object and its internal representation (benchmark). This identification of persons occurs even after a long period of interruption, although during this period there are often significant changes. Therefore, the dynamic image series proposed for viewing can be extremely useful in the process of forming diagnostic hypotheses.

The images in one visual series, although different from each other, are characterized by a commonness with a typical representative, a kind of figurative archetype. The image series considered as a quasi-continuum that is characterized by fuzzy transitions between similar images related to the same disease can also be characterized as a set, and individual age groups can be considered as subsets. A rough set can also be associated with the image series [7, 8]. In this case, two boundary lines are established to describe inaccurate concepts that can allow us to determine the conditional "distance" between the proximity of the new image and one of the image series. Rough sets are a mathematical tool for solving uncertain tasks, which in principle can be applied to the analysis of elements in the image series. In addition, each image can be accompanied by the confidence factors of experts. Respectively, image series of phenotypic (external) manifestations of a disease, representing a wide class of similar but different individual images for different diseases, can serve as a basis for visual diagnostics (prediagnosis).

Inclusion in an image series can also be carried out on the basis of association. Let us consider this with examples from medicine:

- association by similarity – close persons, objects, situations, including separate fragments;

- association on metaphorical reflection, for example, gargoyle-like facial features in mucopolysaccharidosis;

- memory associations (by mental connections) between images or between images and verbal signs (for example, at a certain age, gross facial features and "clawed hands", rough facial features and corneal opacity when suspected of the Hurler or Hunter syndrome, the possibility of combining macrocephaly and external hydrocephalus). In this case, associative signs can be presented in the image series (Figure 1).
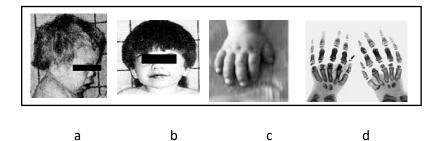
a   b   c   d

Figure 1. Associated series of images of the Hurler and Hunter syndromes: coarse facial features (a, b), (c), radiographic clawed foot (d)

Elements of an image series should be arranged in a certain order. First, the most typical representative of the series (archetype) should be located, and then, in descending order of similarity, the remaining images, ending with the least characteristic for the given image series (in medicine atypical manifestations of the appearance of a specific disease). The arrangement of images in the composition of the series can be carried out in accordance with expert estimates of the similarity of certain attributes. Confidence-building measures are similar to the scale of confidence factors [0,1].

4   Logic-and-linguistic-and-image models

Decision making using images in medicine is determined by the intuitive-image nature of thinking. This is important to bear in mind, since the analysis of a holistic image (for example, rough facial features) is more informative than the examination and mental unification of its fragments. However, combining verbalized signs in intellectual systems of knowledge with imaginative visual series of clinical manifestations that allow an experienced physician to form a primary diagnostic hypothesis for a number of diseases, especially of a hereditary nature, remains unrealized for the time being. Nor does the use of only a holistic image, without appropriate linguistic content, allow us to make the final diagnostic decision.

Logic-argumentative mental processes include signs and images, which indicates the expediency of constructing context-dependent logic-and-image models. It should be borne in mind that the image-argumentation models should assume an alternating order of inclusion of linguistic concepts and visual representations (images).

Verbalized qualitative and visual image knowledge can be represented in the form of frame or ontological models [2]. Visual imagery series included in the frame or ontological model of knowledge representation can be used at the stages of formation and confirmation of hypotheses by their sequential presentation to identify the expert system to the user. This will make it possible to include in the knowledge base poorly verbalized or non-verbalized concepts that correspond to holistic images. Along with medicine, where there are characteristic phenotypic manifestations in certain diseases, this approach can find application in ethnography, archaeology, sculpture, painting and other subject areas in which images are of fundamental importance.

5   Principles of constructing logic-and-linguistic-and-image systems

The architecture of the hybrid linguistic-and-image intellectual system is similar to the traditional one. At the same time, the consequence of the appearance of an image component will be the presence in the knowledge base of a block of images rigidly connected with a block of verbal knowledge [2].

A specific frame will represent a certain image (archetype, parent frame or a typical representative of an image series), and slots will be represented by separate images (elements of an image series). In this range, holistic and fragmentary images which are specific characteristics (markers) of a specific disease can be included. Context in frames can explain (accompany) visual images, but images in a series can be the context in relation to linguistic arguments as logic premises.

When the user enters a sign, the "daemon" in the frame system will activate the "linguistic" frame and the associated frame with images. At this stage, in accordance with the preliminary hypothesis, the corresponding image series will be displayed to the user of the intellectual system to visually identify any similarity with the diagnosed patient. This will speed up the diagnostic process, since erroneous system decisions will be rejected immediately if the images do not match.

As an example, consider the diagnostic fragment for Hunter syndrome, whose knowledge base project includes linguistic and image frames. In the linguistic slot are two signs (thickened skin, noisy breathing) as conditions for the hypothesis. They represent a context and are accompanied by an image series represented by three photos of the heads of patients (Figure 2). Visual images presented to users will be judged by the external similarity with the appearance of the new patient. It is possible to focus on the essential features of the visual image, its fragments. The scale of visual images for assessment at the stages of diagnostic hypotheses promotion ranges from "it seems unlikely", through "it is possible" to "one can not exclude the similarity". When the similarity of the diagnosed patient to one of the members of the imaginary range is revealed, the diagnostic hypothesis will be practically confirmed. In case of doubt, the physician will be able to enter information about other symptoms. If the appearance does not coincide, the search will continue through connection to other image series.

And {thickened skin expressed}

And {noisy breathing expressed}

Figure 2. Fragment of linguistic and image frames for Hunter syndrome (3 fotos)

To compare the new object with the elements of the image series, the method of subjective scaling in the quantized-continuous form can be used, with an estimate of the degree of similarity of visual images on a categorical scale. Objective knowledge in the form of pairwise proximity measures between objects of the image series is a "semantic space" of the subject domain [9]. Alternative approaches to logical inference based on similarity relationships can be based either on a strictly probabilistic approach [10],

or on the apparatus of fuzzy logic and theory of possibilities [11]. In case-based reasoning systems, use-case libraries can also include visual images or image series, including those reflecting the dynamics of temporal changes.

Yang et al. [12] showed that similar images from the inquiry library can provide a deeper interpretation of the newly observed radiological image in mammography. For this, the choice of a metric that preserves both visual and semantic similarity is important. The authors evaluated the algorithm using interactive search-assisted decision support and medical images obtained with ImageCLEF.

6   Intellectual analysis in the classification of images

The use of image series can allow us to set and solve the problems of searching for characteristic unifying features among images of one type (for example, faces of patients with the identical disease or belonging to one group of diseases) or a person's face in the course of development of the disease, i.e. time-series or sequencing of images that change over time. In principle, it is possible to analyze images using previously obtained expert estimates of each individual image.

Accordingly, denotative spaces of image series can be formed on the basis of similarity of certain characteristics into classes called metadenotates (for example, according to the principle of forming subclasses and classes of diseases). In some cases, metadenotates can pool groups of closely related diseases with similar or "masked" images (for example, the Hurler syndrome and pseudo-Hurler polydystrophy or, in modern classification, mucolipidosis II).

7   Discussion

The mental image represents an amalgam of known facts with knowledge in the given area. It is possible to speak about associative-rational activity on the basis of memories of similar situations. In this respect, image representations can be of great importance.

In the earlier systems for the diagnosis of hereditary diseases, including intellectual systems [13, 14, et al.], the verbalization of images led to a partial loss of information or even to misinterpretation. Displaying the photos on-screen after the formation of the diagnostic hypothesis did not allow for correction of earlier mistaken decisions. Nor does the transition to systems that support only image representations without the necessary context [15, 16, 17, et al.] solve the problem. Accordingly, the paradigm of the image-contextual knowledge base was formed, as presented above.

The complex of programs which allows the presentation of visual images to the user-physician at the stage of hypothesis promotion can be regarded as an expert partner in a human-machine system.

8   Conclusion

Traditional expert-type intellectual systems are focused on extracting and processing verbalized knowledge. Clinical decision support systems can be physician assistants in the diagnostic process, including the Fuzzy Arden syntax [18]. However, individual concepts need to be included in the whole, since even their defragmentation results in the loss of information. The inclusion of formal specifications in the form of visual images of the corresponding subject area into classical logical-and-linguistic models can fundamentally solve the problem of improving the efficiency of diagnostic solutions. However, the creation of diagnostic systems for subject areas in which visual images, especially holistic ones, play a major role, should change the approach to constructing models. The knowledge base should integrate

both linguistic and image components. On the basis of intermediate conclusions presented to users and necessarily including images, a decision will be made regarding the further search and formation of the final hypothesis.

REFERENCES

[1]   Paivio A. Imagery and verbal processes. New York, Holt, Rinehart & Winston, 1971.

[2]   Kobrinskii B.A. Approaches to the construction of cognitive linguistic-image models of knowledge representation for medical intelligent systems. Sci Tech Inform Process 2016; 43: 289-295.

[3]   Kobrinskii B.A. Sequences of Images in Intelligent Systems. Sci Tech Inform Process 2010; 37: 328-335.

[4]   Stadler B.M.R., Stadler P.F., Wagner G.P., Fontana W. The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change. J Theor Biol 2001; 213: 241-274. DOI: 10.1006/jtbi.2001.2423

[5]   Richardson J.T.E. Imagery. Abingdon: Taylor & Francis Group, Psychology Press, 1999.

[6]   Solso R.L., MacLin O.H., MacLin M. Cognitive Psychology (8th ed.). Yorkshire, Pearson, 2008.

[7]   Pawlak Z. Rough set theory and its applications. J Telecom Inform Technol 2002; 3: 7-10.

[8]   Zhang Q., Xie Q., Wang G. A survey on rough set theory and its applications. CAAI Transact Intell Technol 2016; 1: 323-333 (open access). DOI: 10.1016/J.TRIT.2016.11.001.

[9]   Voinov A.V. The role of similarity judgment in intuitive problem solving and its modeling in a sheaf-theoretic framework. In: L. Wang, S. Halgamuge and X. Yao, eds. FSKD'02: Proceedings of the 1st International Conference on Fuzzy Systems and Knowledge Discovery. Vol.2. Singapore: Orchid Country Club, November 18-22, 2002. P.753-757.

[10]   Hüllermeier E. Similarity-based inference as evidential reasoning. Int J Approx Reason 2001; 26: 67-100. Doi: 10.1016/S0888-613x(00)00062-1

[11]   Esteva F., Garcia P., Godo L. and Rodrýguez R. A modal account of similarity-based reasoning. Int J Approx Reason 1997; 16: 235-260.

[12]   Yang L., Jin R., Mummert L., et al. A Boosting Framework for Visuality-Preserving Distance Metric Learning and Its Application to Medical Image Retrieval. IEEE Trans Pattern Anal Mach Intell 2010; 32: 30–44. DOI: 10.1109/TPAMI.2008.273.

[13]   Baraitser M. and Winter R.M. London Dysmorphology Database, London Neurogenetics Database & Dysmorphology Photo Library on CD-ROM. 3rd ed. Oxford: Oxford University Press, 2001.

[14]   Aymé S. Orphanet, an information site on rare diseases. Soins 2003; 672:46-47.

[15]   Hammond P., Hutton T.J., Allanson J.E., et al., Discriminating Power of Localized Three-Dimensional Facial Morphology. Am J Hum Genet 2005; 77:999-1010. DOI: 10.1086/498396

[16] Vardell E., Bou-Crick C., VisualDx: A Visual Diagnostic Decision Support Tool. Med Ref Serv Q 2012; 31:414-424. DOI: 10.1080/02763869.2012.724287

[17] Kuru K., Niranjan M., Tunca Y., et al. Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics. Artif Intell Med 2014; 62: 105–118.

[18] Seitinger A., Fehre K., Adlassnig K.P., et al. An Arden-Syntax-based clinical decision support framework for medical guidelines – Lyme borreliosis as an example. Stud Health Technol Inform 2014; 198: 125-132.