Prepocessing of Visual and Audio Information Based on a Fuzzy Network

Shahnaz N. Shahbazova

Azerbaijan State University of Economics (UNEC),
Department of Digital Technologies and Applied Informatics
194 Murtuza Mukhtarov, Baku, Azerbaijan, AZ1001
Emails: shahbazova.shahnaz@unec.edu.az
Corresponding Author
ORCID ID: 0000-0002-9898-6829

Abstract: In this paper we present the results of a study related to the development of the most effective representation and organization of libraries of visual and sound objects in terms of the absence of any restrictions and template optimizations. It is based on the idea of a dictionary concept that contains a brief description of stored visual and sound information. Consequently, providing associative communication between information objects, up to the possibility of creating an over-subject level abstract representation.

Keywords: visual and audio information, virtual environment, contrasting boundaries, objects-sources of sounds, dictionary concept, mechanism of self-organization, transition of information, successful formation of a classification.

1 Introduction

The paper presents and justifies the method of entering visual and audio information into the system used in the work. A virtual environment is illustrated, which is a model of a room containing visual objects in the form of 3-dimensional objects and photographs, and sound objects in the form of objects-sources of sounds. The stage of preliminary processing of visual and audio information received by the system was described. One of the main methods of preprocessing is presented, which consists in applying the Mach effect to obtain more highlighted (contrasting) boundaries of information objects.
The non-standard principles of the formation of classification of information objects based on the mechanism of self-organization of incoming information during training (the transition of the amount of information into quality) are also revealed. A detailed example is given illustrating the successful formation of a classification graph based on a sample of correspondences between information objects with different vocabulary concepts (markers).

2 Mach effect in contrasting boundaries of information objects

The base mechanism for searching for object boundaries is based on studies of the optic nerve and one of the properties called the Mach effect [16] (1). It should be noted that the author did not find a similar principle in the literature on speech recognition and the design of a hearing aid, but those applied in the processing of sound information gave good results.

According to modern concepts the cells that make up the retina of the eye are specialized nerve cells, and they are related evolutionarily and morphologically to excitable tissue in the brain. Thus, some of the retinal nerve cells perform visual processing before the signals leave the eye [2].

Fig.1. Illustration of the visual Mach effect in border regions

Light-sensitive receptor cells on the back of the retina are in front of the black choroid, which increases contrast by eliminating internal reflection and blocking light from passing through the front side of the eyeball. The receptors are connected via bipolar cells (so-called because of their bilateral shape) to the ganglion layers of the nerve that pass outside of the eye to shape the progression of the optic nerve to the brain [1,5].

The retina contains horizontal cells that connect small clusters of receptors. When a receptor is illuminated, adjacent receptors are made less sensitive by horizontal cells, increasing local contrast. This is a preliminary form of edge detection, it causes an optical effect known as the Mach effect (see Fig. 1), this principle in a slightly modified form is used in the preliminary stage of processing incoming information.

Each rectangle contains a uniform shade (see Fig. 2, Fig. 2c), but the edge located near the darker rectangle appears lighter to humans (see Fig. 2a). Likewise, an edge located near a lighter rectangle appears darker [19-22].

Since computer systems do not perceive this property (natural for humans) (see Fig. 2b), during pre-processing, visual and audio information is subjected to mathematical recalculation to obtain an imitation of the effect (see Fig. 2b).

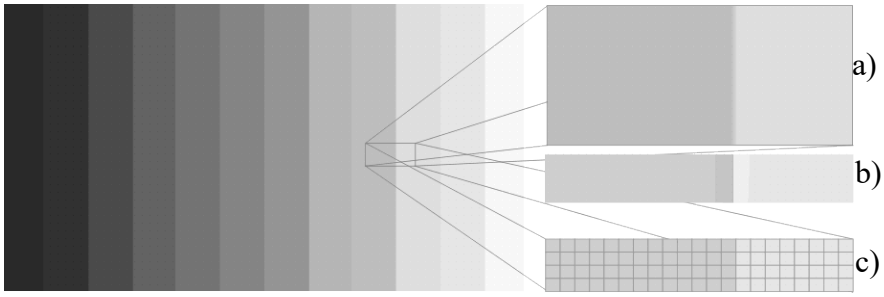


Figure 2. Computer simulation of the Mach effect

The processing difference is that visual information is primarily processed in space, while auditory information is primarily processed in time (see Fig. 3a).
Thus, the information becomes more contrasting:

- visual – by shadow or color transitions;
- sound - based on volume or rhythm transitions.

The figure (see Figure 3b) illustrates the difference in the wearable visual information to create the Mach effect.

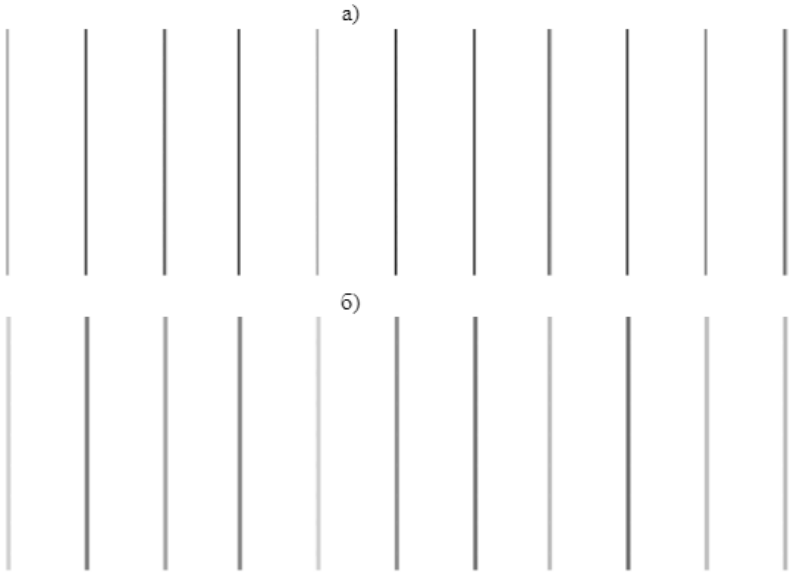


Fig. 3. Changes made to the drawing algorithm for simulating the Mach effect (coverage radius: a – 5 pixels, b – 10 pixels)

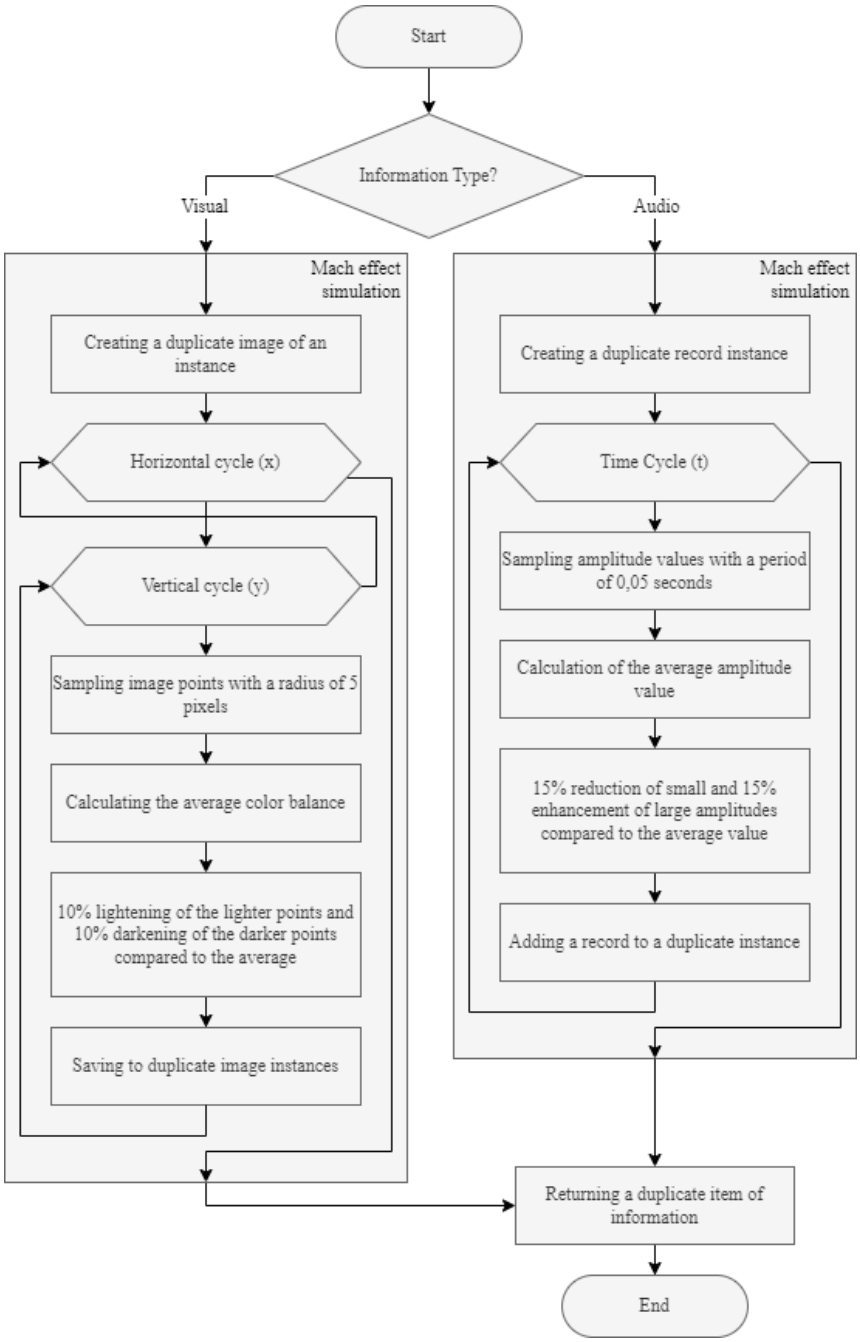
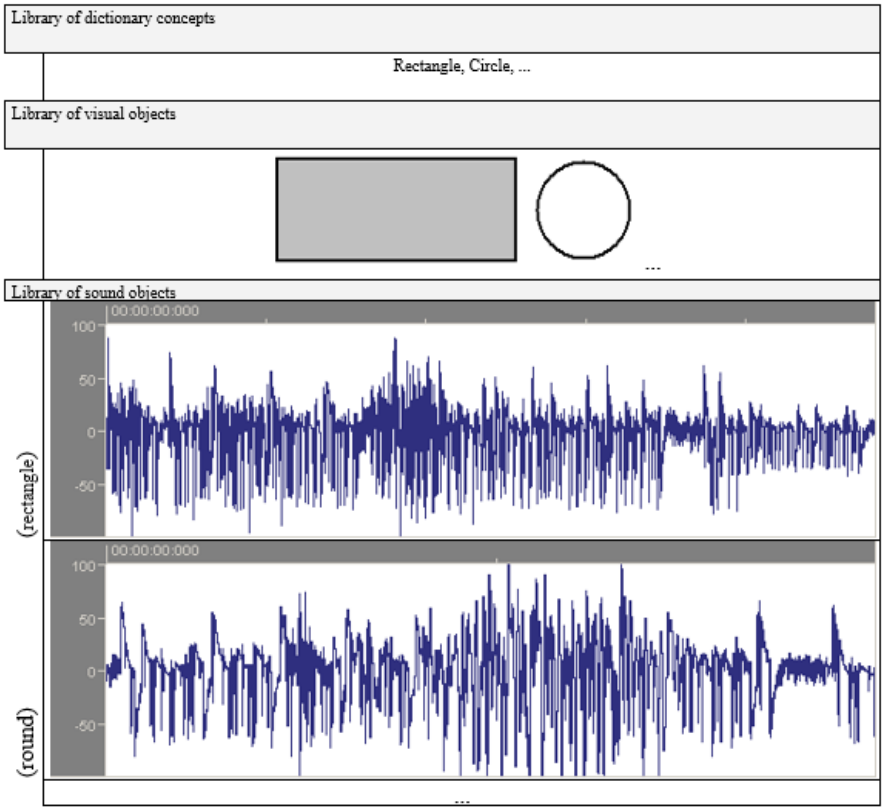


Fig. 4. Information transformation algorithm simulating the Mach effect

3 Organization of Object Libraries

One of the reliable conditions for ensuring the effectiveness of intelligent systems is the quality of implementation of basic knowledge and databases, which create the basis for expert knowledge and decision-making tools.

Table 1
Arrays of object groups



In this work, the system of knowledge bases and databases was not implemented in the classical style since the problem being solved had to remain free from any template solutions in order to carry out further evolutionary development and complication of the system [25-27].

At the heart of an object library, which represents the basis of expert knowledge, is the idea of a dictionary (or vocabulary) concept.

The dictionary concept contains an associative description of stored visual and audio information, allowing the introduction of a new abstract level that, while preserving the semantic meaning of the information received, frees it from the

redundancy associated with the actual information capacity of the object under study[23-24].

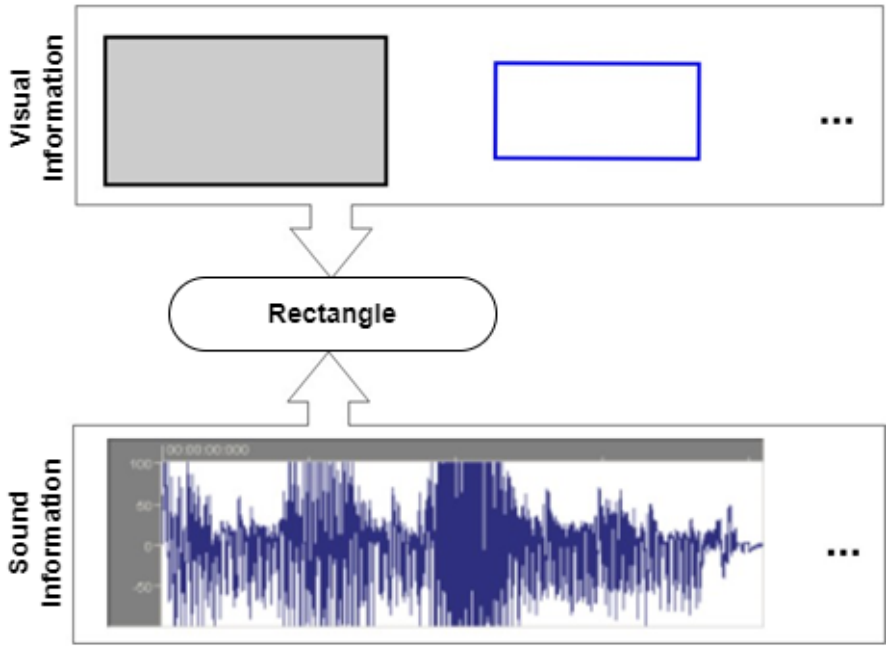


Fig. 5. Associative elements of objects

Memory elements are a compilation of visual and auditory objects with vocabulary concepts (see Figure 5) representing brief information (a marker) about an object of information and providing an associative link between objects. Physically, visual and sound objects are short video fragments (as the most universal format) usually lasting up to 3 seconds, but for static visual objects one frame is enough, and for some sound objects it stretches up to 10-15 seconds. Basically, duration restrictions are related, on the one hand, to the limited performance of aircraft, and on the other hand, to expediency, since the more diverse the record about an object, the less likely it is for clear recognition [4,6-8].
In the existing conversion algorithm, the primary information is in any case stored with the appropriate marking (as a new copy), leading to the rapid filling of the disk drives available to the system, while the human intellect “by default” forgets the features of the object and sound (except for special cases related to effects memory or increased interest), briefly remembering only the fact of the received information.
In the presented work, semantic concepts are singular nouns of visual and audio objects, for example, rectangle, bear, table, chair, computer, etc., thus fully providing the entire range of visual and audio information required in research [9-13].

In the ongoing work on the system, quite successful steps are being taken on the way to the use of non-material dictionary concepts in the system. At the moment, the main emphasis is on the use of verbs, which represent a new above-the-subject level of recognition, but due to the qualitatively new level of computational costs, sufficient stable results [28].

Conclusion

The paper describes the stage of preliminary processing of the visual and sound information received by the system. One of the main methods of preliminary processing is presented, which consists of applying the Mach effect to obtain more distinct (contrast) boundaries for information objects.

References

[1] Findeisen W., Malinowski K. Two-level control and coordination for dynamical systems. Arctium automatiki i telemechaniki. T. XX1V, N1, p.3-27.

[2] Foltyniewicz R. Efficient High Order Neural Network for Rotation, Translation and Distance Invariant Recognition of Gray Scale Images. Lecture Notes in Computer Science - Computer Analysis of Images and Patterns, 1995, pp. 424-431.

[3] Frei W. Image Enhancement by Histogram Hyperbolization // Computer Graphics and Image Processing. – 1977. – Vol. 6. № 3. – P. 286-294.

[4] Fukushima, K. (1988). Neocognition: a hierarchical neural network capable of visual pattern recognition. Neural Networks 1 (2), Стр. 119-130.

[5] Gonzalez R.C., Fittes B.A. Gray–Level Transformation for Interactive Image Enhancement // Mech. Mach. Theory. – 1975. – V.12. – P. 111-112.

[6] Gonzalez R.C., Wintz P. Digital Image Processing // Addison–Wesley. Reading. – Massachusetts, 1987. – 505 p.

[7] Gutta S. and Wechsler H. Face recognition using hybrid classifiers. Pattern Recognition 1997, Vol. 30, pp. 539-553.

[8] Hall E.L. Almost Uniform Distribution for Computer Image Enhancement // IEEE Trans. Computers. – 1974. – Vol. 23. № 2 – P. 207-208.

[9] Kaiser G., Kaplan S. Parallel and distributed Incremental Attribute Evaluation Algorithms for Multiuser Software development Environments. ACM Transactions on Software Engineering and Methodology, vol.2, No.1, Jan. 1993, 47 - 92.

[10] Ketcham D.J. Real Time Image Enhancement Technique // Proceedings SPIE/OSA Conference on Image Processing. – Pacific Grove, California. – 1976. – Vol. 74. – P.120-125.

[11] Kralik J., Stiegler P., Vostry Z., Zavorka J. Modelovani dynamiky

rozsahlych siti. Praha, Akademia, 1984, 364p.

[12] Krzyzak A., W. Dai, C.Y.Suen. “Unconstrained Handwritten Character Classification Using Modified Backpropagation Model,” Proc. 1st Int. Workshop on Frontiers in Handwriting Recognition, Montreal, Canada, Стр.155-166, 1990.

[14] LeCun Y., O.Matan, B.Boser, J.S.Denker, D.Henderson, R.E.Howard, W.Hubbard, L.D.Jackel, H.S.Baird. “Handwritten Zipcode Recognition With Multilayer Networks,” Proc. of International Conference on Pattern Recognition, Atlantic City, 1990.

[15] Lee, Y. (1991) Handwritten digit recognition using K neares-neighbor, radial-basis function, and back-propagation neural networks. Neural Computation 3, 440-449.

[16] Lilley C., F. Lin, W.T. Хьюитт, T.L.J.Howard, Цвет в Компьютерной Графике, Отделение компьютерной графики, Отдел Информатики, Университет Манчестера, Перевод в 1998-1999 году Мачехиной Л.Р., Анисимов С.Ю., anis@anis.kchepetsk.ru. http://pmg.org.ru/basic3d/, http://pmg.org.ru/basic3d/colors_ru.zip.

[17] Mariton M., Drouin M., Abou-Kandil H., Duc G. Une nouvelle method de decomposition-coordination. 3 e partie: Application a la commande coordonnees-hierarchisee des procesus complexes. "APII", 1985, N3, p.243-259.

[18] Milanova M., Almeida P. E. M., Okamoto J. and Simoes M. G. Applications of Cellular Neural Networks for Shape from Shading Problem. Lecture Notes in Artificial Intelligence - Machine Learning and Data Mining in Pattern Recognition, 1999, pp. 51-63.

[19] Jacobsen X., Zscherpel U. and Perner P. A Comparison between Neural Networks and Decision Trees. Lecture Notes in Artificial Intelligence - Machine Learning and Data Mining in Pattern Recognition, 1999, pp. 144-158.

[20] Petrou M. Learning in Pattern Recognition. Lecture Notes in Artificial Intelligence - Machine Learning and Data Mining in Pattern Recognition, 1999, pp. 1-12.

[21] Ranganath S. and Arun K. Face recognition using transform features and neural networks. Pattern Recognition 1997, Vol. 30, pp. 1615-22.

[22] Sharov S.A., Tools for computer representation of linguistic information, 1996, http://www.rriai.org.ru/~sharoff/, sharoff@mx.iki. rssi.ru.

[23] Shahbazova Sh.N., Suleymanov M.R., “Stages of recognition of probabilistic images”, Scientific notes of AzTU, 2004

[24] Shahbazova Sh.N., Bonfig K.V., Suleymanov M.R., “Issues of constructing a recognition system for non-standard images and speech in architecture

based on a hybrid neuro-fuzzy logic apparatus," 6th International Conference "Application of Fuzzy Systems and soft computing and system of analyses, decisions and control", Barcelona, Spain, 2004

[25] Tzafestas S.G. Large-scale systems modeling in distributed-parameter control and estimation. "Modeling and Simul. Eng. 10th IMACS World Congr. Syst. Simul. and Sci. Comput., Montreal, 8-13 Aug. 1982, v.3" Amsterdam e.a., 1983, p.69-77.

[26] Valentin D., Abdi H., O'Toole A. J. and Cottrell G. W. Connectionist models of face processing: a survey. IN: Pattern Recognition 1994, Vol. 27, pp. 1209-1230.

[27] Vetter T. and Poggio T. Linear Object Classes and Image Synthesis From a Single Example Image. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, pp. 733-742.

[28] Wilson I.D. Foundations of hierarhical control. "International Journal of Control" 29, N6, 1979, p.899-933. [24] A. Marjani, M. Babanezhad, and S. Shirazian, "Application of adaptive network-based fuzzy inference system (ANFIS) in the numerical investigation of Cu/water nanofluid convective flow," *Case Stud. Therm. Eng.*, vol. 22, no. November, p. 100793, Dec. 2020.