Hybrid-Maximum Neural Network for Depth Analysis from Stereo-Image

Łukasz Laskowski

Technical University of Czestochowa, Department of Computer Engineering,
Al. A.K. 36, 42-200 Czestochowa, Poland

Abstract. In present paper, we describe completely innovation architecture of artificial neural nets based on Hopfield structure for solving of stereo matching problem. Hybrid neural network consists of classical analogue Hopfield neural network and maximal neural network. The role of analogue Hopfield network is to find of attraction area of global minimum, whereas maximum network is to find accurate location of this minimum. Presented network characterizes by extremely high rate of working with the same accuracy as classical Hopfield-like network. It is very important as far as application and system of visually impaired people supporting are concerned. Considered network was taken under experimental tests with using real stereo pictures as well simulated stereo images. This allows on calculation of errors and direct comparison to classic analogue Hopfield neural network. Results of tests have shown, that the same accuracy of solution as for continuous Hopfield-like network, can be reached by described here structure in half number of classical Hopfield net iteration.

Keywords: Hopfield, neural networks, stereovision, depth analysis, hybrid network.

1 Introduction

Sight is the sense that people make most use of it in everyday life. Eyes give us possibility to estimating distance to nearby object, recognizing, reading, predicting of moving object position and so on. All of these guarantee safe and convenient life. People, who lost or damaged this sense are not able to live independently. It is a multi-disciplinary effort to develop devices for individuals whom happened to loose their sight. One of solution can be portable binocular vision system, which makes possible in orientation and mobility for blind users. The system can be based on stereovision [1]. The advantages of stereovision include ease of use, non-contact, non-emission, low cost and flexibility. For these reasons, it focuses attention of scientists to develop of stereo-vision methods.

Stereovision is natural way of determination of distance by human. The simplified model of human sight can be present as two parallel cameras, and this model (named parallel stereovision system) is going to be considered in present paper. In stereovision system (two displaced parallel cameras) expression on distance to given world point can be written as eqn.1.

L. Rutkowski et al. (Eds.): ICAISC 2010, Part II, LNAI 6114, pp. 47–55, 2010.

$$z = \frac{df}{x_l' - x_r'}. \tag{1}$$

Where x_l' and x_r' are positions of point's image on planes of left and right cameras, f denotes the focal length of the cameras, and d represents the distance between two cameras.

Depth analysis [2] depends on determination of third dimension of each point in observed scene. Third dimension of point is inversely proportional to difference in position of image of this point on plane left and right (disparity). Problem seems to be trivial and is trivial for one point. Real scene contains from large amounts of points, and the complexity of the correspondence problem depends on the complexity of the scene. There are constraints and schemes that can help reduce the number of false matches, but many unsolved problems still exist in stereo matching [3]. Main problems are occlusion, discontinuity of depth, discontinuity of periphery as well as regularity and repetitivity. For this reasons stereo matching problem is one of the most complex problems in computer vision.

There were purposed few types of algorithms for solving this problem [4]. Main of this was: Feature based algorithms, Phase based algorithms, energy based algorithms and area based algorithms. Nowadays algorithms in original form are used rarely only to very basic problems. Scientists try to merge different types of solution in order to pull out as many advantages from all types of algorithms as possible and avoid disadvantages. Energy can be minimized also by using Hopfield-like Neural Nets [5], [6]. The ability of the Hopfield network to solve optimization problems relies on its steepest descent dynamics and guaranteed convergence to local minima of the energy landscape. This kind of system was used in stereo-matching problem. Both types of Hopfield-like network were used: continuous and discrete.

Looking on the state-of-the-art-of stereovision matching with using of Hopfield-like networks one can has the impression that this domain is well investigated. However this subject is so wide and complicated, that there are still possibility of improve efficiency such systems or working out better architecture of nets. Author tried to use mentioned above solutions to stereo matching process. Each time error of network working (calculation of errors was described in farther chapters of present work) and number of iteration were noted. Unfortunately neither of presented above network works correctly. Error of working was very high (each time above 30%) what practically eliminate those methods to solving of stereo matching problem. It is the reason of looking for new solution and new architecture of neural networks. Author tried to increase of working ratio with no loosing of solution accuracy.

In present paper completely innovative architecture of network to solving stereo matching problem is presented. Due to using of two types of networks: classical analogue Hopfield like network and maximum network described here structure is working much faster then classical Hopfield net with satisfied accuracy. Its efficiency was confirmed in tests on real and simulated pictures (what allowed on error calculation).

2 Analogue Hopfield Neural Network for Stereo Matching Problem

Stereo matching problem can be casted as an optimization task, where an energy function, representing constrains on the solution, has to be minimized. The optimization problem then can be performed by means of the Hopfield neural network. The most accurate solution can be obtained by analogue Hopfield-like neural net. The proposed network consists of $n \times n$ neurons for one epipolar line in image. It is easy to note that target system will consist of n networks working pararely - each network will realize stereo-matching problem for one epipolar line. n is dimension of images ($width = height = n$). Each neuron neu_{ik} with potential v_{ik} is responsible for fitting i-point in right image to k-point in left image. The highest external potential of neu_{ik} is, the better fittingo of points. In final configuration only for corresponding points, i in right image to k in left image potential of neu_{ik} will be equal 1, for rest point external potential of neurons will be equal 0. Very convenient is to represent neurons an a matrix, named Fitting Matrix (FM), where number of row represent i index, and number of column represents k index.

The equation of motion of the Hopfield model must be discretized by means of a numerical method. In this case Euler discretization was used:

$$u_{ik}(t+1) = u_{ik}(t) + \Delta t\Big(\sum_{j}\sum_{l} t_{ik,jl} v_{jl}(t) + I_{ik} - \frac{u_{ik}(t)}{\tau}\Big) \tag{2}$$

where Δt is time step, τ is a positive constant (interpreted as neuron relaxation time). In presented design value $\Delta = 10^{-3}$ have been chosen, which have been determined to be small enough for the Euler rule to provide enough accuracy.

3 The Energy Function

In presented here method crucial is the energy function [7]. Network proceeds minimization of this function until it finds minimum. This means that solution of problem was found.

Minimization of energy function must secure following criterions:

1. For couples of correlated points (i, k) and (j, l) in given epipolar line, where i and j are numbers of point in right image, k and l are numbers of point in left image, Correlation Coefficient $C_{ik,jl}$ should have as high value, as possible - term of *Correlation*;
2. Assigning must be reciprocally unique - term of *Uniqueness*;
3. Sequence of assigning in areas must be kept - term of *Area Sequence*;
4. Continuity of assigning in areas must be kept - term of *Continuity*;
5. Global sequence of assigning must be kept - term of *Global Sequence*;

Taking all these terms into consideration, energy function can be expressed in form of equation:

$$\begin{aligned} \mathrm{E}= & -a\sum_i\sum_k\sum_j\sum_l C_{ik,jl}v_{ik}v_{vjl} \\ & +b\big(\sum_i\sum_k\sum_{l\neq k} v_{ik}v_{jl} + \sum_i\sum_k\sum_{j\neq i} v_{ik}v_{jl}\big) \\ & +c\sum_i\sum_k\sum_{l\leq k} v_{ik}v_{(i+1)l}\sigma_{i,i+1} + d\sum_i\sum_k\sum_l v_{ik}v_{(i+1)l}\sigma_{i,i+1}\xi_{ik,jl} \end{aligned} \quad (3)$$

where a, b, c, d, e are weight coefficients of each energy components. Having an energy function given as eqn. 3 interconnection weights and external currents are given following equation:

$$\begin{cases} t_{ik,jl} = aC_{ik,jl} - b\delta_{ij}(1-\delta_{kl}) - b\delta_{kl}(1-\delta_{ij}) - c\rho_{l<k}\delta_{(i+1)j}\sigma_{i,i+1} \\ \qquad\quad -d\delta_{(i+1)j}\sigma_{i,i+1}\xi_{ik,jl} \\ I_{ik} = 0 \end{cases} \quad (4)$$

where δ_{ij} is Kronecker delta, sign $\rho_{l<k}$ is defined by following equation:

$$\rho_{i<k} = \begin{cases} 0 \text{ for } l > k; \\ 1 \text{ for } l \leq k; \end{cases} \quad (5)$$

Values, given by eqn. 4, are basis of working of continuous Hopfield neural network.

4 Architecture of Hybrid-Maximum Neural Network

Disadvantage of analogue Hopfield neural network is long time of computation. Rapidity of working is very important as far as target system is concerned - it should work in real-time. Much faster is a maximum neural network. Additional advantage of maximum network is automatically meeting of term of uniqueness. This type of neural structure was introduced by Takefuji and al in [8]. Maximum neural network was defined as discrete Hopfield-like network with specific activation function: only neuron with the highest value of internal potential (in some group) is activated, rest of neurons have low potential. The maximum activation function for stereo matching problem can be formulated as follow:

$$f(u_{ij}) = \begin{cases} 1 \text{ if } u_{ij} = max(u_{i1}, u_{i2}, ..., u_{in}) \\ 0 \text{ otherwise}; \end{cases} \qquad i,j = 1,...,n. \quad (6)$$

This kind of neural network found application in optimization problems [9]. Unfortunately in original form maximum network is not fit to solving of stereo-matching problem. The reason is the same as for discrete Hopfield neural network - stereo matching problem is too complex and network is trapped in local minima.

However it is possible to join precision of working of analogue Hopfield network with rapidity of maximum neural network's working. Presented here hybrid neural network contains of analogue Hopfield network and maximum neural network. An architecture of discussed here neural network was depicted on fig. 1.

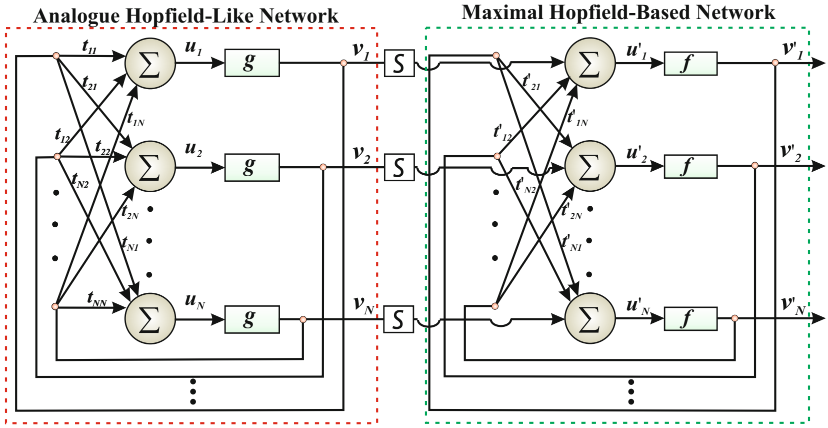

Fig. 1. An architecture of hybrid - maximum neural network

In first stage of working analogue Hopfield neural network is looking for attraction area of global minimum. After finding attraction area of global minimum, network is switched to maximum mode thanks to block of switched function S (see fig. 1). Switching follows after given number of iteration (determined empirically). In maximum mode network evaluates fast towards towards global minimum and term of uniqeness is kept automatically thanks to maximum activation function. It must be stressed, that the form of energy function for maximum network is little different, as on eqn. 4 - there is only one term of uniqeness, second is realized by activation function.

5 Experimental Results

The proposed method was implemented on a Personal Computer with Pentium IV-2.80 GHz CPU and 2 GB SDRAM.

Helpful to analysis of network working is the neurons activity map. It can be interpreted as a graphical form of fitting matrix for investigated line - white points mean neurons with high potentials, black points correspond to neurons with low potentials. Intermediate colors correspond to values between 0 and 1. Thanks to neurons activity map the dynamics of neural network can be observed (map is updated with each iteration).

The resolution of input stereo-images is 100 $\times$100. Images are calibrated in order to find corresponding lines, before starting stereo matching procedure. This process allows on scanning of pictures line-by-line, what decreases complexity of method.

To verify the efficiency of the proposed method, an experiment was performed using both simulated and real images. Using simulated images allowed on error calculation (only in this case error is calculated).

An operational procedure for solving the stereo matching problem is summarized as follows:

1. Assume number of image epipolar line $h = 0$;
2. Assume maximum number of iteration it_{max} enough to finding attraction area of global minimum;
3. IEF mapping into Analogue Hopfield Network:
   (a) Compute external inputs of neurons and their interconnection strengths using eqn. 4 (with keeping of symmetrical interconnection strength's matrix);
   (b) Initialize states of neurons in heuristic way - assume potentials v_{ik} proportional to correlation coefficients C_{ik};
4. Continuous Hopfield Network updating procedure for energy minimization (working in continuous Hopfield mode):
   (a) For each neuron compute the internal potential, with using of eqn. 2;
   (b) For each neuron compute the external potential using sigmoidal activation function;
   (c) If number of iteration is equal it_{max}, go to (5), else go to (a);
5. Energy function mapping into Maximum Network:
   (a) Compute external inputs of neurons and their interconnection strengths (with keeping of symmetrical interconnection strength's matrix);
   (b) Assume states of neurons the same as at the end of working of continuous Hopfield network;
6. Maximum Network updating procedure for energy minimization (working in maximum mode):
   (a) For each neuron compute the internal potential;
   (b) For each neuron compute the external potential using eqn. 6;
   (c) If changes of internal potentials for each neurons are equal zero go to (7), else go to (a);
7. If present epipolar line is not last one, increment number of line $h = h + 1$ and go to (2), else go to (8);
8. End simulation.

There was assumed, that attraction area of global minimum usually reached after 50 iterations (empirically confirmed). In maximum mode stable state is reached after at most 20 iterations, and this limit of iteration was assumed in order to have possibility of confirmation of results reached for different stereo-images.

Results of stereo matching process carried out by Hybrid Maximum Network can be seen below, for simulated images (fig. 2), and for real images (fig. 3).

In first row stereo pictures, used for stereo matching process and obtained depth map were shown. Second row shows neurons activity maps for 80 scanning line (arbitrary assumed) in iterations (number of "n") of neural net working in analogue Hopfield mode. The last row presents the same sequence, repeated for network working in maximum mode.

The error of stereo matching process can be calculated only for simulated pictures (possibility of neurons activity map determination). Results of Hybrid Maximum Network's working was juxtaposed with results of stereo matching

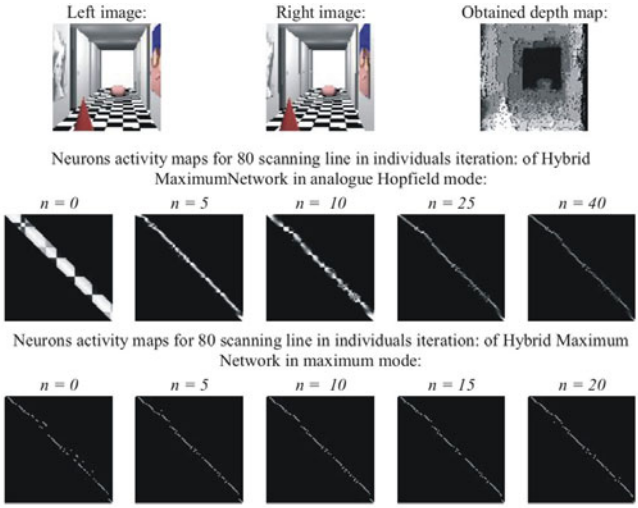

Fig. 2. The result of stereo matching process caried out by Hybrid-Maximum Neural Network for simulated stereo-images

process done by analogue Hopfield-like network and Maximum network. Considered here network types worked in the same conditions, with using the same fitting coefficients. Errors of working (δd) and iteration numbers, necessary to reaching stable state amounted:

– $\delta d = 20.04\%$, 60 iterations for Hybrid Maximum network,
– $\delta d = 19.89\%$, 83 iterations for analogue Hopfield-like network,
– $\delta d = 68.52\%$, 54 iterations for Maximum Network.

Analyzing of network working (fig. 2 and fig. 3) one can notice, that in analogue Hopfield mode fitting is ununiqueness. This can be concluded by analyzing of neurons activity maps. In the case of uniqueness of stereo matching in each row and each column of fitting matrix (its graphical form is neurons activity map) should be at very most one non-zero element, whereas in stable state few non-vanishing elements can be observed in columns and rows of fitting matrix. Because of ununiqueness it is different to stay anything about sequence in areas. Also term of keeping of depth continuity in areas can not be stayed. This mistake can be corrected in maximum working mode. Maximum activation function involve meeting of uniqueness term, what can be seen in iterations of network in maximum mode. In each line of FM only on non-zero element can be seen. Stable state is reached very fast thanks to limitation of possible neuron states configuration (maximum activation function). In comparison to analogue Hopfield network, Hybrid Maximum Network is working with the same efficiency (error about 20% in both cases), but much faster. Results of stereo matching obtained by maximum network can not be accepted for the sake of large error - above 68% what discredit this kind of network to solving stereo matching task.

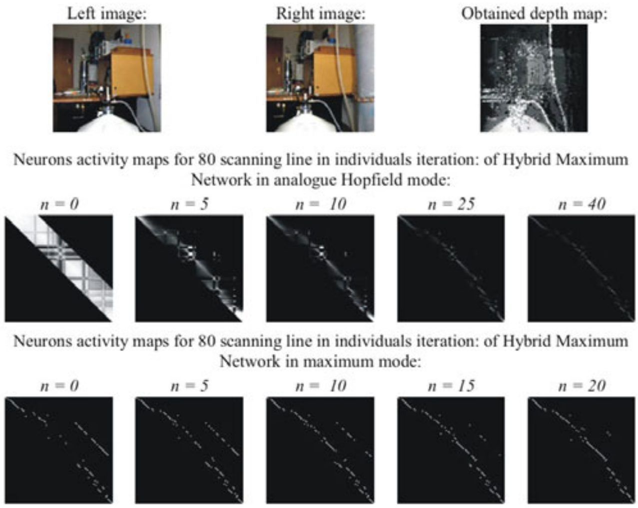

Fig. 3. The result of stereo matching process caried out by Hybrid-Maximum Neural Network for real stereo-images

6 Conclusion

This study presents using of innovative architecture of Hopfield based neural network-Hybrid Maximum Network. Introduced here network has been used to stereo matching process. The stereo correspondence problem has been formulated as an optimization task where an energy function of network is minimized. The advantage of using a Hopfield neural network is that a global match is automatically achieved because all the neurons are interconnected in a feedback loop so that the output of one affects the input of all the others. The convergence to stable state is guaranteed for continuous Hopfield-like network with continuous activation function. The parallel execution capability of this structure is also a powerful reason taking under consideration target system assisting aged people and/or visually impaired. Additionally thanks to using of maximum mode, time of computation significantly decreases.

The experimental results indicate significant gains from using of maximum mode after finding of global minimum's attraction area. A comparative analysis, performed with the classical Hopfield network, maximum network and Hybrid Maximal Network indicated beter performance of last type of network. Obtained solution of stereo correspondence problem was similar to obtained by analogue Hopfield-like network, but reached in less amount of iterations.

References

1. Aleksander, I.: Artificial vision for robots. Korgan Page (1983)
2. Barlow, H.B., Blackmore, C., Pettigrew, J.D.: The natural mechanism of binocular depth discrimination. J. Physiology 193, 327–342 (1967)

3. Faugeras, O.: Three-dimensional computer vision. A geometric viewpoint. MIT Press, Cambridge (1993)
4. Alvarez, L.: Dense Disparity Map Estimation Respecting Image Discontinuities: A PDE and Scale-Space Based Approach. Journal of Visual Communication and Image Representation 13, 3–21 (2002)
5. Hopfield, J.J., Tank, D.W.: Neural computation of decisions in optimization problems. Biological Cybernetics 52, 141–152 (1985)
6. Hopfield, J.J., Tank, D.W.: Artificial neural networks. IEEE Circuits and Devices Magazine 8, 3–10 (1988)
7. Amit, D.J., Gutfreung, H., Sampolinsky, H.: Spin-glass models of neural networks. Physical Review A2 32(2), 1007–1018 (1985)
8. Takefuji, Y., Lee, K.C., Aiso, H.: An artificial maximum neural network: a winner-take-all neuron model forcing the state of the system in a solution domain. Biological Cybernetics 67(3), 243–251 (1992)
9. Takefuji, Y., Lee, K.C.: Neural network computing for knight's tour problems. Neurocomputing 4, 249–254 (1992)