IEICE Proceeding Series

Detection of Spatial-Temporal Events with Delayed CNN Templates

András Horváth, Tamás Roska

Vol. 2 pp. 378-381
Publication Date: 2014/03/18
Online ISSN: 2188-5079

2013 International Symposium on Nonlinear Theory and its Applications
NOLTA2013, Santa Fe, USA, September 8-11, 2013

Detection of Spatial-Temporal Events with Delayed CNN Templates

András Horváth[†] and Tamás Roska[†]

[†]Faculty of Information Tecnhology, Pazmany Peter Catholic University, Budapest, Hungary
1444 Pf. 278, Budapest, Hungary

Abstract—Spatial-temporal event detection is a crucial task in machine-vision and it is usually difficult to be handled efficiently with current algorithms and devices. The human vision system can solve extremely difficult task with simplicity and low power consumption, like looming or detection of moving objects with given speed and direction. In this article we show how cellular neural networks with delayed type templates are capable of detecting certain spatial-temporal events. The detection is done by using continuous dynamics without cutting the input flow into frames. We can observe similar structures, the analogy of delayed type templates in the retina as well, which performs well and efficiently in image processing tasks. Delayed type templates can provide us with even more flexibilities and possibilities in new applications including frameless detection of motion features.

1. Introduction

The detection of spatial features on still images is well developed and one can find many libraries for solving various image processing tasks [1]. These libraries and operations can be found for cellular neural networks (CNNs) as well.

However it is impossible to find operations for spatial-temporal detection, not only for CNNs but also for regular computers. Most commonly these tasks are solved by frame-by-frame algorithms, which usually handle differences between the consecutive processed frames. They are usually not identifying spatial-temporal dynamics, only spatial characteristics and derive back the temporal features to the differences of the frames. One can find practical solutions but they are all depending on temporal frequencies (if the frequency of the input-sampling changes the result can change as well) and are usually not robust. If one of the frames are missing in the input-flow the results can be very different. This shows that these methods are not identifying real spatial-temporal characteristics and are applied in practice because of the fact, that quantized samples can be handled in regular, transistor-based computers relatively efficiently.

In the human vision system we can find some marvelous examples of spatial-temporal detections. The detection of looming objects takes place in the retina [2], and the operation of the retina is continuous, not frame-based. This detection is done by continuous dynamics and with extremely low power consumption.

Many similarities between the operation of the retina and cellular neural networks have been shown [3]. However the template operations in CNNs happen with the same speed in every direction and on cell, meanwhile in the retina the communication time between rod, cones and ganglion cells is different. Considering this similarities we will examine certain type of delayed-CNNs and show how simple spatial temporal dynamics can be used to identify continuous dynamics on these delayed type CNN networks. Delayed type templates can provide us with even more flexibilities and possibilities in new applications including frameless detection of motion features.

In section 2 we describe delayed type CNNs and also a simplified version of the general theorem which cover most of the practical networks (including nanoscale devices) and similarities to the retina. In section 3 we show how a class of delayed type templates can be designed to detect given spatial-temporal motifs. In section 4 we show a possible application on frameless detection of motion features used for gesture recognition and control. In section 5 we conclude our results.

2. Delayed type CNNs

Great attention have been paid to CNNs in recent years and many publications were presented regarding image processing[4]. The applications clearly show that these networks can solve image and signal processing tasks efficiently and with low power consumption.

Although delayed type templates have been investigated mathematically[5], most papers investigate stability criteria (with Ljapunov methods and linear matrix inequality techniques)[6], [7]. But the applicability of such architectures was not investigated properly.

To describe a delayed type CNN architecture we require an $M \times N$ grid, containing $n = MN$ the CNN elements: The behavior of the ith element ($i = 1, 2, 3 \ldots n$) is defined by the following delayed differential equation:

$$\begin{aligned}\frac{dx_i(t)}{dt} &= -\gamma_i x_i(t) + \textstyle\sum_{j\epsilon S_i} A_{ij} y[x_j(t)] + \sum_{j\epsilon S_i} A^D_{ij} y[x_j(t - \tau_{A_{ij}}(t))]) \\ &\quad + \textstyle\sum_{j\epsilon S_i} B_{ij} u_j(t) + \sum_{j\epsilon S_i} B^D_{ij} u_j(t - \tau_{B_{ij}}(t))) + Z_i \end{aligned} \tag{1}$$

The state of the CNN cells is defined by: $x(t) = [x_1(t), x_2(t), x_3(t) \ldots x_n(t)]^T \epsilon \mathbb{R}^n$, the state vector of the system with initial conditions: $x_i(s) = \theta_i(s)(s \epsilon [-\tau, 0], i =$

$1, 2, 3 \dots n)$ $\gamma_i (i = 1, 2, 3 \dots n)$ are the state coefficients with $\gamma_i > 0$. Representing the rate in which the cell will return to resting potential ($x_i = 0$) in isolated state. In most papers, also in this one γ_i are the same for all i, and considered unity. $y(t) = [y_1(t), y_2(t), y_3(t) \dots y_n(t)]^T \epsilon \mathbb{R}^n$ are the output values of the cells. $y(t)$ has to be continuous, bounded and monotonically increasing on R, satisfying the Lipschitz condition. In this case the nonlinearity is the following:

$$y_i(t) = \frac{1}{2} |x_i(t) + 1| - \frac{1}{2} |x_i(t) - 1| \qquad (2)$$

$u_i(t)$ is the input of cell i in time t. Z_i is the bias of cell i (considered constant). A_{ij} and A^D_{ij} are defining the feedback and delayed feedback matrices for cell i. B_{ij} and B^D_{ij} are defining the feedforward, input and delayed input matrices for cell i. S_i is the neighborhood, the sphere of influence of cell i, S_i contains the set of those cells which will directly influence the behavior of cell i. If the CNN is homogeneous the values of template A, A^D, B, B^D are not depending on i only on j. The transmission delays are defined as: $\tau_{A_{ij}}(t)$ and $\tau_{B_{ij}}(t)$. If the delays are not changing in time we can define them as $\tau_{A_{ij}}$ and $\tau_{B_{ij}}$.

2.1. Delayed type CNNs in practice

Practical implementations of CNNs are made by coupling similar cells in a cellular network. The cells can be electrical[8], chemical[9], spin-torque oscillators[10] etc... If we consider our cellular network as a set of elements where the coupling strength and delay will depend on physical properties, like the distance between the elements, we can derive much simpler dynamics defining this subset of delayed type CNNs. These networks are especially typical in the nano-scale using nano-magnets or spin torque oscillators as cells, because here the distance between the cells is the only parameter that will determine the connections and in the nano-scale only these type of connections can be implemented efficiently.

To simplify our notation we can define two matrices A_c and A_d, where A_c will define the coupling strengths and A_d will define the coupling delays (τ).

Similarly we can do the same for the input templates: B_c and B_d will determine the coupling. In this case we do not have to use two different matrices, like feedback and delayed feedback template, because if element a_{ij} is nonzero in one matrix it has to be zero in the other, this way we can melt these two matrices into one.

$$\begin{aligned} \tfrac{dx_i(t)}{dt} &= -\gamma_i x_i(t) + \textstyle\sum_{j \epsilon S_i} A_{c_{ij}} y[x_j(t - \tau_{A_{d_{ij}}}(t)]) \\ &+ + \textstyle\sum_{j \epsilon S_i} B_{cij} u_j(t - \tau_{B_{dij}}(t))) + Z_i \end{aligned} \qquad (3)$$

3. Template design for delayed type CNN templates

In this section we show how templates can be designed for delayed type CNNs defined in section 2.1. Similarly to the driving point plot in regular CNNs, in certain cases we can design templates for given spatial-temporal functionality.

If the delay between elements is one directional, which means only one of the elements will affect the other and there are no retroactive coupling, the templates can be designed easily. In case of homogeneous CNNs this means that if a template value in A_d is non-zero, the opposite value (mirrored to the central element of the template) has to be zero. In case of heterogeneous CNNs this condition is fulfilled if the connectivity graph of the cells (an edge is drawn from node i to node j, if i is coupled to j with a given nonzero tau) forms a directed acyclic graph. Although this constraint may seem strict at first glance, we will see in section 4 that many of the practical tasks can be solved with templates fulfilling these conditions.

If the conditions above are fulfilled we can define separated time-plains for every element and we can start template design from the element which is not determined by any other cells and continue the design on the other, consecutive plains. Although in theory this method is feasible, in practice the iterations define driving-point plots for every time-plains, which can mean in practice a few thousand different driving-point plots and template designs. In practice we can design the templates by computer programs which can easily handle a few thousand driving-point plots.

In the following section we will show simple examples how templates can be designed for detecting movement with given speed and direction and how templates can be designed to detect different spatial-temporal trajectories.

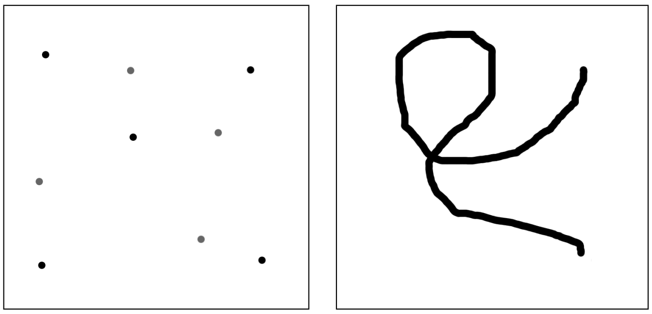
3.1. Detecting objects with a given speed

In this section we will show how we can design a delayed template that identify objects moving slower than a given speed in a given direction. We assume for simplicity that the objects are black and are moving in front of a white background. Let be the direction of the movement a diagonal, south-east direction and the threshold velocity of the detection v. We will note τ_v the time required for the object to move from one cell to its neighbor with speed v.

Our task is to design a template which will excite the cell only if the object is 'in the cell' in the current state and was τ_v time ago in the neighboring cell defined by the direction of the movement. The simplest B template detecting this spatial-temporal dynamics will be:

$$B_d = \begin{bmatrix} \tau_v & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} B_c = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} Z = -1 \qquad (4)$$

The cell with these templates will be excited only if its current input is black and its neighbor was black τ_v time ago. It is easy to see that this template implements the desired functionality and detects all objects which moved slower than the given v velocity. We can also adjust our template, that after a cell got excited it will remain in excited state regardless from the input-flow. We can also de-



(a) Frame from the input flow

(b) Frame from the output flow

Figure 1: As it can be seen on the image only the object with the given properties: given trajectory (direction, speed and intensity) was detected from the nine moving objects. The left figure shows one frame from the input-flow with 9 objects, meanwhile the right figure shows the detected trajectory (when all the excited cells remain in their state) covered by the upper-right black object.

sign a template that will excite the cells only for a predefined time, and after this the cells will return to their normal state.

3.2. Detecting objects with a given trajectory

Based on the method and example defined in the previous section and with the aid of computer programs we can easily define heterogeneous CNNs, which will identify previously given trajectories only.

For a simulation setup we have recorded 64 different trajectories (with different speed, path and object characteristics) and we have designed heterogeneous CNNs which were able to distinguish between these trajectories. The trajectories are defined as chains of CNN cells and a cell can only become excited if the previous cell in the trajectory is already excited.

We can design the last pixels in every trajectory in a way that they will maintain their excited state. This way it is enough to read out only one pixel from the network to identify whether a trajectory occurred or not.

An example of the trajectory detection can be seen on Figure 1.

These two simple examples show how delayed type CNN templates can be designed for given functionalities. In the next chapter we will show two simple examples how delayed type template can be applied in practice.

4. Applications

In this section we show two practical applications: velocity detection and gesture detection. Both examples are using frameless detection, where continuous dynamics are identified by continuous dynamics, none of the methods are sensitive to quantization.

(a) Input Image

(b) Image from the delayed CNN network

Figure 2: As it can be seen on the image the positive and negative edges can be detected easily and the spatial difference between these edges is proportional to the speed of the object. This implies a speed calculation with the usage of continuous dynamics. Both images contain 64 × 64 pixels.

Velocity detection

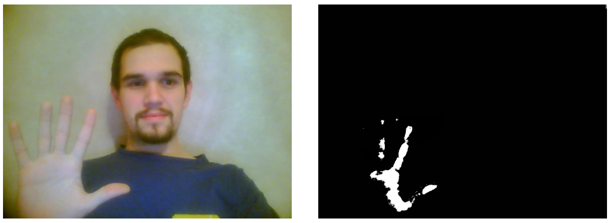
It is a known fact, that Sobel kernel based edge detection ('optimal edge detector' in CNN template libraries) creates a double edge at detected object. One of the edges will be created on one side of the object while a negative edge will appear on the other side, due to the fact, that the signs of the coefficients in the template are the opposite.

We can use this simple fact to detect the velocity of an object. If we alter the delay template accordingly, we can delay the appearance of the negative edge on the image by a given time (τ_e):

$$B_d = \begin{bmatrix} \tau_e & 0 & 0 \\ \tau_e & 0 & 0 \\ \tau_e & 0 & 0 \end{bmatrix} B_c = \begin{bmatrix} -0.11 & 0 & 0.11 \\ -0.28 & 0 & 0.28 \\ -0.11 & 0 & 0.11 \end{bmatrix} \quad (5)$$

Using these templates we will get an output video (again a continuous flow), where the appearance of the negative edge will be in the position, where the object was τ_e time ago. If we process this output-flow, we can easily identify the positions of the edges (both positive and negative edges can be found by regular CNN templates) and from the results we can easily calculate an edge based optical-flow for our object. From the optical-flow we can easily derive the velocity of the object. An example of speed calculation can be found on figure 2.

Other feature based detections, like corner or similarity based features can also be derived with similar delayed templates.



(a) Input Image

(b) Image from the delayed CNN network

Figure 3: As it can be seen on the image the positive and negative edges can be detected easily and the spatial difference between these edges is proportional to the speed of the object. This implies a speed calculation with the usage of continuous dynamics. Both images contain 64 × 64 pixels.

Gesture detection

Based on the previous examples and methods we have generated a setup, which is able to detect hand movement according to predifined trajectories. The different spatial movements and the velocity of the movement are identified by delayed type CNN architectures and it can translate it to stored commands. The system in its current state can distinguish between four different directions (the maximal component of the velocity vector) and three different speeds. This can define a user interface containing 12 different gestures identified by a homogeneous delayed type CNN.

With the usage of heterogeneous CNNs even more different trajectories could be distinguished. Example frames from the gesture detection can be seen on figure 3.

5. Conlusion

We have described the theory of delayed type CNN templates and also introduced a method with example for designing simple delayed type CNN templates. We have showed some simple examples for the detection of spatial-temporal characteristics, such as speed and trajectory detection. We also implemented practical applications like a user interface based on a delayed type CNN for control.

The novelty of our method is frameless processing, which differs from common image processing techniques. We do not process spatial temporal signals as sequences of frames by discrete operations or processing individual frames and then combining or subtracting them from each other. Our detection is a result of interacting continuous dynamics combined with the dynamics of our device. The detection does not need temporal quantization and it can be applied in case of extremely fast events as well. We think this kind of detection can be applied in those cases where there are no inherent clock cycles, like in most of the currently researched nano devices, like spin-torque oscillators.

Acknowledgments

The support of the grants TÁMOP-4.2.1.B-11/2/KMR-2011-0002 and TÁMOP-4.2.2/B-10/1-2010-0014 are gratefully acknowledged.

References

[1] RJ Schalkoff , "Digital image processing and computer vision" *John Wiley & Sons, Inc. Publishers - New York*, 1989

[2] T. Fulop; A. Zarandy , "Bio-inspired looming direction detection method" *IEEE 13th International Workshop on Cellular Nanoscale Networks and their Applications*, pg. 1-6, August 2012, Turin, Italy

[3] Frank Werblin, Tamas Roska, Leon O. Chua , "The analogic cellular neural network as a bionic eye" *International Journal of Circuit Theory and Applications*, Volume 23, Issue 6, pages 541-569, November-December 1995

[4] Chua L.O., Roska T., "Cellular Neural Networks and Visual Computing: Foundations and Applications" *University Press, Cambridge*,2002

[5] T. Roska , "Cellular neural networks with nonlinear and delay-type template elements" *Proceedings of IEEE International Workshop on Cellular Neural Networks and their Applications*, pages 12-25, 1990

[6] Cao J. , "A Set of Stability Criteria for Delayed CNNs" *IEEE Transactions on Circuits and Systems I: Regular Papers*, Volume 48, pages 494-498, 2001

[7] Athanasios I. Margaris, "On the global stability of time delayed CNNs" *Central European Journal of Computer Science*, Volume 1, Issue 1, pp 67-100, March 2011

[8] Rodriguez-Vazquez, A, Espejo, S. ; Dominguez-Castron, R. ; Huertas, J.L. ; Sanchez-Sinencio, E. "Current-mode techniques for the implementation of continuous- and discrete-time cellular neural networks" *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Volume 40, Issue 3, pp 132-146, March 1993

[9] T. Asai, Y. Kanazawa, T. Hirose, Y. Amemiya "Analog Reaction-Diffusion Chip Imitating Belousov-Zhabotinsky Reaction with Hardware Oregonator Model" *Int. Journal of Unconventional Computing*, Vol. 1, pp. 123?147, 2005

[10] Gyorgy Csaba, Matt Pufall, Dmitri Nikonov, George Bourianoff, Andras Horvath, Tamas Roska, Wolfgang Porod "Spin Torque Oscillator Models for Applications in Associative Memories" *13th IEEE international workshop on cellular nanoscale networks and applications*, Turin, Italy, 2012