International Symposium on Nonlinear Theory and its Applications
Xi'an, PRC, October, 7-11, 2002

NEW MOTION SEGMENTATION APPROACH BASED ON CELLULAR NEURAL NETWORKS IN COMPLEX IMAGE SEQUENCES

MENG Shukai, MO Yulong, CHENG Yuning

Dept. of Communication Engineering, Shanghai University
No. 149 Yanchang Rd., Shanghai, China, 200072
Phone: +86-21-56332291, Fax: +86-21-56336908
Email: skmeng@263.net, yulongmo@online.sh.cn

Abstract

In this paper, a new approach to motion segmentation based on cellular neural networks (CNN) is proposed. Difference merged image algorithm based on CNN is presented and applied to motion segmentation of complex image sequences. In order to realize the algorithm, we design three CNN templates, negative template, addition template and composition template, which can be used in Cellular Neural Network Universal Machine. Comparing with traditional algorithms, difference merged image algorithm can obtain more motion information and increase the accuracy of segmentation, because it is directly used in gray-scale image sequences without adopting threshold processing as the first step. Finally, we show the experiment results of motion segmentation, which prove the efficiency and accuracy of this approach.

1. Introduction

The video coding standard MPEG-4 is enabling content-based function, which needs for the segmentation of semantically meaningful motion objects in complex image sequences. In recent years, many good approaches to motion segmentation of image sequences are presented, but most of them can not satisfy the real-time demand that is important to video processing, for they are limited by the computing capability of some hardware and algorithms themselves. So, it is a feasible solution to adopt new hardware that has high computing speed and good algorithms that match the hardware.

Cellular neural networks are basically large scale nonlinear analog circuits, composed of locally connected cells [1], which have outstanding capabilities when used in image and video processing. CNN is very well suitable for VLSI implementations, and the so-called CNN Universal chip has very high computing speed, for instance, the CNN Universal chip CNNUC3 [3] [5] can arrive at 1 Tera (10^{12}) analog instructions per second. For many video processing applications, this immense speed surpassed the computing power of many current supercomputers and can satisfy the real-time demand of video.

The CNN architecture is described in [5], and we only recall the main results. Here, each cell can be described by the CNN equation:

$$\dot{x}_{ij} = -x_{ij} + z_{ij} + \sum_{kl\in S_{ij}(r),kl\neq ij} a_{kl} f(x_{kl}) + \sum_{kl\in S_{ij}(r),kl\neq ij} b_{kl} u_{kl} \qquad (1)$$

and the output equation

$$y_{ij} = f(x_{ij}) = \frac{1}{2}(|x_{ij}+1| - |x_{ij}-1|) \qquad (2)$$
$$i = 1, 2, \ldots, M, \ j = 1, 2, \ldots, N$$

where $u\in R^u$ is input, $z\in R^z$ is threshold, $x\in R^x$ is state and $y\in R^y$ is output, the "dot" denotes the time derivative, $S_{ij}(r)$ is sphere of influence whose radius is r. For a space-invariant CNN with 3×3 sphere of influence ($r = 1$), nine control coefficients b_{kl} and nine feedback coefficients a_{kl} compose two 3×3 matrices B and A, and the two matrices B and A are called the control template and the feedback template, and z_{ij} is threshold.

First, as the core of the approach, difference merged image algorithm based on CNN will be introduced in detail in section 2. And then, in section 3, negative template, addition template and composition template are shown, which are necessary to constitute the algorithm. Finally, in order to compare with the traditional algorithm in [4] and prove the efficiency and accuracy of this approach, the experiment results are shown in section 4.

2. Difference merged image algorithm based on CNN

Difference merged image algorithm is developed from difference image algorithm. For this algorithm is based on CNN and can be used in Cellular Neural Network Universal Machine, it makes this algorithm has high speed in motion segmentation. On the other hand, comparing with the traditional algorithm in [4], the novel character of this algorithm is without threshold processing as the first step and the difference operation is directly used in gray-scale image sequence, which can increase the efficiency and accuracy of motion segmentation and overcome the shortcoming of the traditional algorithm that make some original information in gray-scale images lost. So we can use this algorithm to segment motion objects in complex image sequences. This point will be proved in section 4.

Difference merged image algorithm based on CNN should follow the conditions:
(1) The background of an image sequence is relatively static. (2) The optical flow of adjacent frames little changes.

We use $f(x, y, i)$ to denote the current frame in a complex image sequence, $f(x, y, i\text{-}k)$ denotes the No. k frame previous to the current frame, and $f(x, y, i+l)$ denotes the No. l frame next to the current frame. Difference merged image algorithm based on CNN can be described as:

$$P = SEG[AND[FILL[EDGE[D1]],\ FILL[EDGE[D2]]]] \quad (3)$$

where

$$\begin{aligned} D1 &= ADD[f(x,y,i-k),\ REV[f(x,y,i)]] \\ &= f(x,y,i-k) - f(x,y,i) \end{aligned} \quad (4)$$

$$\begin{aligned} D2 &= ADD[f(x,y,i),\ REV[f(x,y,i+l)]] \\ &= f(x,y,i) - f(x,y,i+l) \end{aligned} \quad (5)$$

and $REV[\cdot]$ denotes the operation of negative template, $ADD[\cdot]$ is the operation of addition template, $EDGE[\cdot]$ is the operation of edge detection template, $FILL[\cdot]$ is the operation of patch maker template, $AND[\cdot]$ is the operation of logical AND template, $SEG[\cdot]$ is the operation of composition template. P denotes the motion objects segmented from the current frame $f(x, y, i)$, $D1$ and $D2$ denote the two difference images. The flow chart of this algorithm is shown in Fig. 1. The input of the algorithm is three continuous frames $f(x, y, i\text{-}k)$, $f(x, y, i)$ and $f(x, y, i+l)$ in an image sequence, and the output is motion objects in the current frame $f(x, y, i)$.

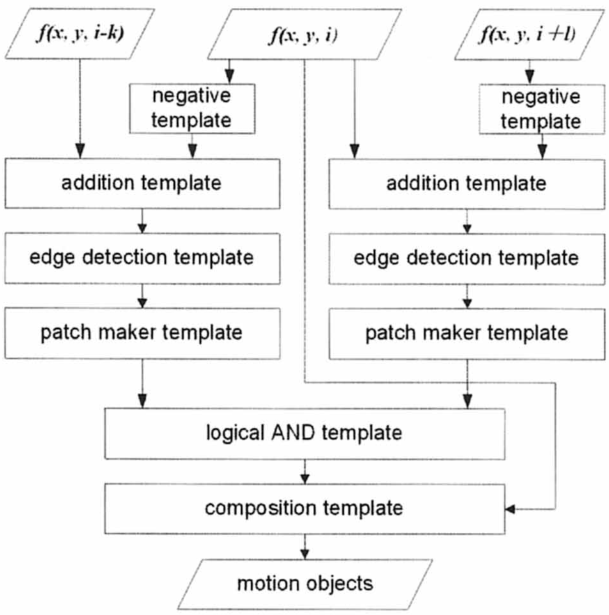


Fig. 1 the algorithm flow chart

First, through negative template and addition template processing, the two difference images $D1$ and $D2$ are calculated. And then, after processed by edge detection, patch maker and logical AND template, the two difference images are merged into one binary image which only contains the region of motion objects. Finally, the motion objects in current frame can be segmented by composition template.
Among those CNN templates, edge detection template, patch maker template and logical AND template are introduced in [5] [6]. Here, we will not review them. In section 3, we will mainly discuss negative template, addition template and composition template.

3. CNN templates of the algorithm

3.1 Negative CNN template

For CNN operation has no subtraction operation directly, in order to realize the difference operation between the neighborhood frames, we present negative CNN template. According to equation (2), the output $y_{ij} \in [-1, 1]$ is different from the traditional gray-scale bitmap image whose range is [0, 255]. When a gray-scale image is processed by CNN, the input image $u_{ij} \in [0, 255]$ have to be mapped into [-1, 1]. Under this condition, the function of this template is to change the input gray-scale image to the negative image of itself as the output, that is to make $y_{ij} = -u_{ij}$.
The input of this template is a given gray-scale image u_{ij}; the initial state is $x_{ij}(0) = 0$; the output is $y_{ij} = -u_{ij}$. The template is described as:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad z_{ij} = 0 \quad (6)$$

3.2 Addition CNN template

To each corresponding pixel of two same-sized gray-scale images, addition template can realize add operation, and what we obtained is still a gray-scale image. But different from the ordinary add operation, the output of the template is defined by equation (2). In this algorithm, we can adopt this template and negative CNN template to produce the two difference images $D1$ and $D2$. Because two difference images are gray-scale images and most motion information of objects is saved in them, it is very effective to improve the accuracy of motion segmentation compared with the algorithm in [4], which is the novel character of this algorithm It will be proved in section 4.
The input of addition template is first gray-scale image $u_{ij} = f_1(x, y)$; the initial state is second gray-scale image $x_{ij}(0) = f_2(x, y)$; the output is the sum of first image

and second one, that is $y_{ij} = f_1(x,y) + f_2(x,y)$. Addition CNN template can be described as:

$$A = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{Bmatrix}; \quad B = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0 \end{Bmatrix}; \quad z_{ij} = 0 \tag{7}$$

3.3 Composition CNN template

In order to segment motion objects from the current frame, composition CNN template is presented. It composes the result calculated from logical AND template and the original image of the current frame to segment and reconstruct an image that only contains the motion objects in the current frame. And the output of this template is a gray-scale image.
The input of this template is the current frame *f(x, y, i)*; the initial state is the result calculated from logical AND template; the output is the motion objects segmented from the current frame *f(x, y, i)*, that is *P* in equation (3). Composition CNN template is described as:

$$A = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{Bmatrix}; \quad B = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0 \end{Bmatrix}; \quad z_{ij} = -0.5 \tag{8}$$

4. Experiment results

Under PC with Intel Pentium 4 1.7G CPU and Windows 98 operating system, we use the simulate software—CNN Visual Mouse Platform to prove the algorithm we proposed. The experiment image sequence is standard image sequence "children", and it is a complex image sequence, which not only contains multiple rigid and non-rigid motions objects but also have the different motion directions and the complex background. From this image sequence, we choose No. 170 frame as the current frame *f(x, y, i)*, No. 165 frame as the previous frame *f(x, y, i-k)* and No. 175 frame as the next frame *f(x, y, i+l)*. The goal of experiment is to segment the motion objects from No. 170 frame. The experiment results of difference merged image algorithm based on CNN are shown in Fig. 2.
According to Fig. 2, we can see that the two difference images Fig. 2 (d) and Fig. 2 (e) are still gray-scale images, which obviously contain the information of negative images, and the contour of motion range is clear and complete. The relative static background range is shrouded by zero (where the range of gray-scale is [-1, 1]). So, we said this algorithm extract the motion information directly from gray-scale image sequences. Fig. 2 (f) is the final result of the experiment, that is the motion objects segmented from Fig. 2 (b). As concerning the segmentation of semantically meaningful motion objects, this experiment proves the algorithm we present has a good capability of motion segmentation in complex image sequences.

(a)

(b)

(c)

(d)

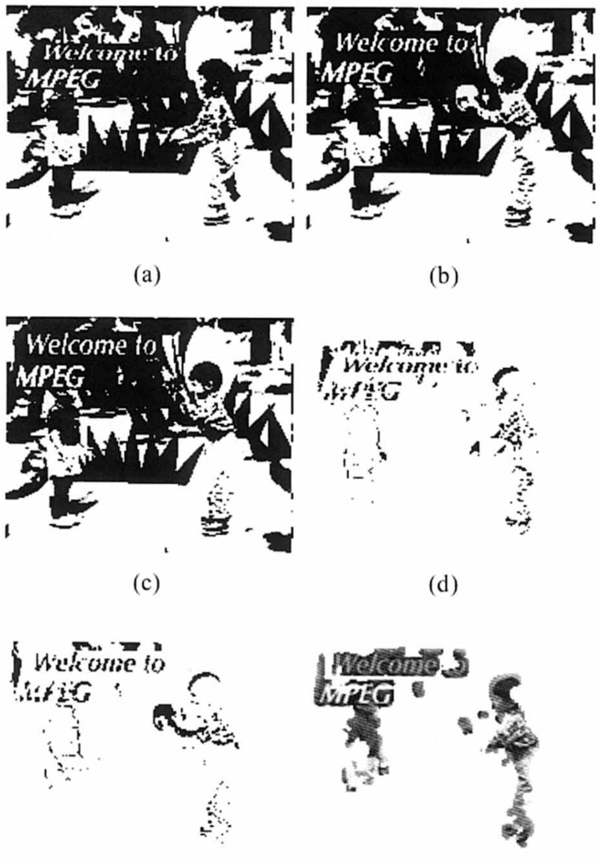
(e)

(f)

Fig. 2 The experiment results of this algorithm:
(a) No. 165 frame *f(x, y, i-k)*;
(b) No. 170 frame *f(x, y, i)*;
(c) No. 175 frame *f(x, y, i+l)*;
(d) difference image *D1*;
(e) difference image *D2*;
(f) motion objects of No. 170 frame.

In order to explain the advantage and character of this algorithm further, according to the algorithm introduced in [4] completely, we do another experiment as comparison, which has the same experiment environment and image sequence. The algorithm in [4] first adopt threshold processing to change gray-scale image sequence into binary one, and then apply accumulative difference image algorithm based on CNN to segment the obtained binary image sequence. The experiment results by using the algorithm in [4] are shown in Fig. 3.
Through threshold processing (the threshold value is 0.4), we obtain three binary images, Fig. 3 (a), Fig. 3 (b) and

Fig. 3 (c). Obviously, the information of these three binary images is less than that of the original images. It makes the difference images Fig. 3 (d) and Fig. 3 (e) contain the insufficient motion information, and can not describe the motion objects completely. So, as the final segmentation result Fig. 3 (f), it can not make us satisfied. As concerning the segmentation of semantically meaningful motion objects, the segmentation objects are incomplete.



(a) (b) (c) (d) (e) (f)

Fig. 3 The experiment results by the algorithm in [4]:
(a) binary image of No. 165 frame;
(b) binary image of No. 170 frame;
(c) binary image of No. 175 frame;
(d) difference image between Fig. 3(a) and Fig. 3(b);
(e) difference image between Fig. 3(b) and Fig. 3(c);
(f) motion objects segmented from No. 170 frame.

Obviously, comparing with the algorithm in [4], difference merged image algorithm based on CNN has more advantage in the efficiency and accuracy of segmentation. As we know, the information entropy of gray-scale images is much greater than that of binary image. For the algorithm we present directly orients the gray-scale image sequences, it means more complete motion information can be obtained and motion segmentation is more efficient and accurate.

5. Conclusion

In this paper, a new approach to motion segmentation based on CNN in complex image sequences is proposed. Difference merged image algorithm is the core of this approach, which is constructed by six CNN templates, that is negative template, addition template, edge detection template, patch maker template, logical AND template and composition template. And three of them, negative template, addition template and composition template, are introduced and designed. The algorithm directly orients the gray-scale image sequences, which is the novel character of it. Finally, through the experiment, the efficiency and accuracy of this approach has been proved.

Acknowledgements

Sincerely thanks to Ms. CHENG Yuning, Huawei Technologies Co., Ltd, for her valuable advice and great support.

References

[1] L.O. Chua, L. Yang, "Cellular Neural Networks: Theory and Applications," *IEEE Trans. on Circuits and Systems,* Vol.35, No. 10, pp.1257-1290, Oct. 1988.

[2] K. Slot, "Determination of cellular neural network parameters for feature detection of two-dimensional images," *Proc. IEEE int. Workshop on Cellular Neural Networks and their Applications*, pp.82-91, 1990.

[3] P. Foldesy, G. Linan, etc, "Object oriented image segmentation on the CNNUC3 chip," *2000 6th IEEE International Workshop on Cellular Neural Networks and Their Applications Proceedings*, pp.283-288, 2000.

[4] ANCL, BARCO, CATU, EPFL, IMSE, KUL, STM, "*First Year Progress Report*," ESPRIT PROJECT IST 1999-19007: DICTAM, pp.56-67, Feb. 2001.

[5] Leon O. Chua, *CNN: A Paradigm for Complexity*, World Scientific Publishing Company, New Jersey, June 1998.

[6] ANALOGIC COMPUTERS LTD, "*Aladdin V1.3 CNN SOFTWARE LIBRARY FOR ACE4K CHIP*," 2000