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

Cellular Neural Networks for Object-Oriented Coding Schemes: Motion Estimation

Giuseppe GRASSI and Luigi Alfredo GRIECO

Dipartimento di Ingegneria dell'Innovazione, Università di Lecce, 73100 Lecce – Italy
giuseppe.grassi@unile.it; alfredo.grieco@unile.it

Abstract: Image analysis algorithms are of great interest in the context of object-oriented coding schemes. By exploiting Cellular Neural Network (CNN) architecture, in this paper a new analogic algorithm for obtaining the *motion estimation* is illustrated. Simulation results, carried out for *Miss America* and *Claire* video sequences, confirm the validity of the CNN-based approach developed herein.

Keywords Cellular Neural Networks, Motion Estimation, Object-oriented Image Analysis

1. Introduction

The object-oriented coding approach represents an interesting and sophisticated method for video coding [1]-[2]. It requires a strong image analysis stage, which consists of four steps: i) segmentation; ii) motion estimation; iii) image synthesis; iv) consistency observation [1]. The result of such operations is that objects of great importance for the image understanding are detected and coded using special parameter sets, whereas still background objects are not included in the coding procedure. This approach enables the transmission rate to be greatly reduced. Since object-oriented coding schemes require very powerful devices, it can be promising to exploit the great computational power offered by the CNN Universal Machine [2]. Referring to CNN-based approaches, the results available in literature only concern the development of analogic algorithms for the segmentation stage [1]-[2]. Therefore, the aim of this paper and of the companion one [3] is to bridge the gap by developing the remaining stages of the object-oriented image analysis scheme. In particular, this paper presents a new analogic CNN algorithm for *motion estimation*, whereas the companion paper [3] focuses on the *image synthesis* and *consistency observation* steps. The paper is organized as follows. In Section 2 the analogic algorithm for implementing the *motion estimation* step is described. Such stage mainly consists in modeling and evaluating the geometrical and motion properties of the objects. Finally in Section 3 simulation results, carried out for *Miss America* and *Claire* video sequences, show the effectiveness of the proposed CNN algorithm.

2. CNN Algorithm for Motion Estimation

2.1. CNN Model

Herein the following two-dimensional CNN array is considered:

$$\dot{x}_{ij}(t) = -x_{ij}(t) + \sum_{kl\in N_r} A_{ij,kl}\, y_{kl}(t) + \sum_{kl\in N_r} B_{ij,kl}\, u_{kl}(t) + \\ + \sum_{kl\in N_r} \hat{A}_{ij,kl}(y_{kl}) + \sum_{kl\in N_r} \hat{B}_{ij,kl}(u_{kl}) + I_{ij} \qquad (1)$$

where $x_{ij}(t)$ is the state, $y_{ij}(t)$ is the output, $u_{ij}(t)$ is the input, $A_{ij,kl}$ and $\hat{A}_{ij,kl}$ are the linear and nonlinear feedback operators, respectively, $B_{ij,kl}$ and $\hat{B}_{ij,kl}$ are the linear and nonlinear control operators, respectively, I_{ij} is the bias whereas f is the usual piece-wise linear output function. Five different images can describe a CNN layer, that is, the input, state, output, bias and mask images. These images are indicated with $\boldsymbol{U}$, $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{I}$ and $\boldsymbol{M}$, respectively (black=+1; white=-1; gray-scale within [-1,1]). In particular, the mask image is a binary image specifying whether the correspond cell is in active state (mask pixel=+1) or inactive state (mask pixel = -1).

2.2. Motion Estimation

The proposed CNN algorithm (Fig. 1) consists in determining in the frame $\boldsymbol{I}_{i+n}$ the position of those objects belonging to frame $\boldsymbol{I}_i$, which represent the result of the segmentation technique described in [1]. The position of these objects in the frame $\boldsymbol{I}_{i+n}$ can be computed by moving each object (belonging to frame $\boldsymbol{I}_i$) in a pxq-pixel window and by comparing the result with the frame $\boldsymbol{I}_{i+n}$. Finally, the position of the object that corresponds to the smallest error is stored for the next stage.
The first part of proposed algorithm is based on CNN templates called INVERSION, ADDITION, ABSOLUTE VALUE and

DIFFUSION (for computing the mean absolute error). Successively, in order to move the object and its mask within the frame $\boldsymbol{I}_i$, the templates UP, RIGHT,DOWN and LEFT are considered. All these templates can be found in the CNN Software Library (CSL) developed in Budapest [4]. If the segmented objects detected in frame $\boldsymbol{I}_i$ are N_i, the result of the CNN motion estimation algorithm are the compensated images $\boldsymbol{Y}^{COMP}_{(OBJ)\,i}$ and their corresponding masks $\boldsymbol{M}^{COMP}_{i}$ for i=1, … N_i.

3. Simulation Results

3.1. Motion estimation

We start by reporting in Fig. 2(a) and Fig. 2(c) the third ($\boldsymbol{I}_3$) and the seventh ($\boldsymbol{I}_{3+4}$) frame of the *Miss America* video sequence, respectively, whereas in Fig.2(b) the result of the segmentation stage is illustrated (N_3 = 9, that is, 9 segmented objects). By applying the motion estimation algorithm to the ninth object, with a 15x15-pixel window, the selected images reported in Figure 2(d)-2(h) are obtained. Namely, the iteration that corresponds to the minimum error gives the image reported in Fig. 2(f), whereas the final result of the algorithm, that is, the *compensated object image* and the *compensated mask* are reported in Fig. 2(g) and Fig. 2(h), respectively.

3.2. Comparisons and performances

At first, we apply the motion estimation algorithm to the first 50 frames of the *Miss America* sequence, starting from frame 3 with a 15x15-pixel window. Successively, we apply the algorithm to the first 50 frames of the *Claire* sequence, with a 19x19-pixel window. We start by frame 3, where 7 objects are found (see Fig. 3). Notice that, since the motion in *Claire* is larger than the motion in *Missa*, a larger search-window has been chosen. The results of the comparison are reported in Fig. 4, where for each sequence the gains in dB are reported. Such gains have been obtained by computing the energies of the *difference image* with motion compensation and without motion compensation. The improvements in dB clear highlight the performances of the proposed algorithm in estimating the scene transformation.

4. Conclusion

This paper has focused on object-oriented image analysis using CNN architecture. In particular, a new analogic CNN algorithm for *motion estimation* has been presented. Simulation results for *Miss America* and *Claire* video sequences have shown the effectiveness of the algorithm. Notice that the results obtained herein make perceive the possibility of applying the proposed Motion Estimation algorithm to artificial vision via stereo processing systems, where the basic task is to detect moving objects in real time. In particular, it is necessary to detect objects that enter or exit the scene, as well as objects that, by moving through the scene, are potentially dangerous for the motion of the observer.

Acknowledgements

This paper has been supported by MURST under grant PRIN 2001 (P.I.: Prof. Mauro Forti, University of Siena, Italy).

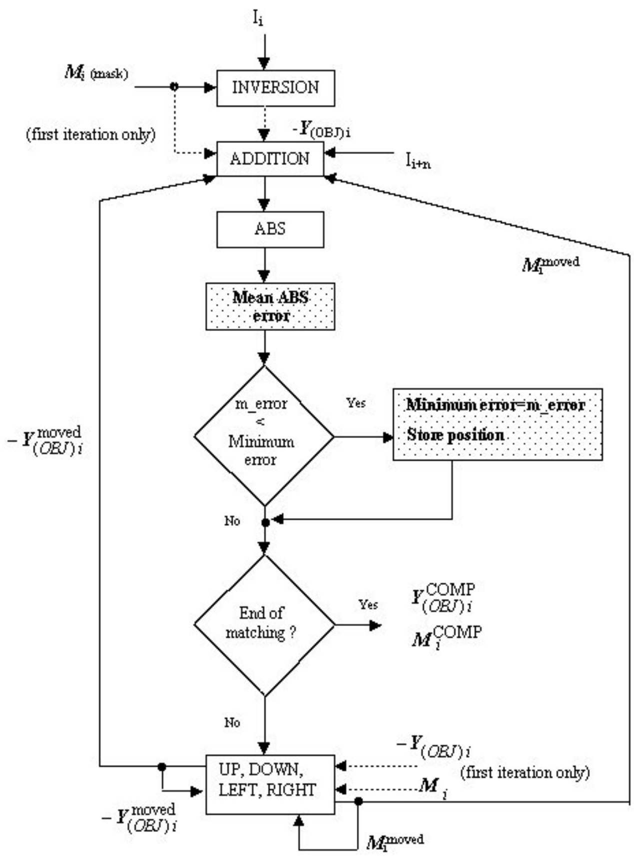

Figure. 1. Block diagram of the motion estimation algorithm.

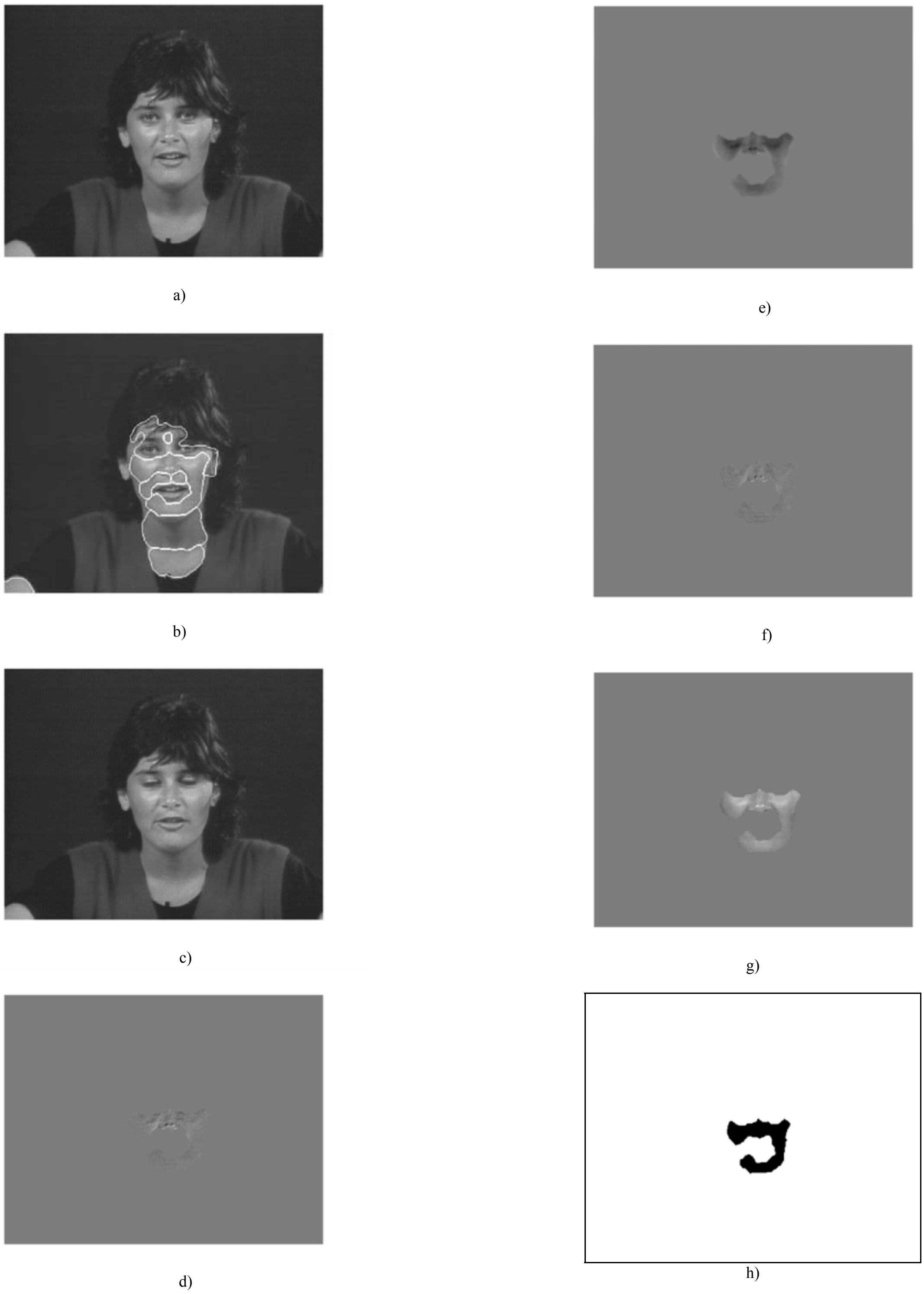

Figure. 2. Motion estimation algorithm for *Miss America*: a) frame 3; b) segmented frame 3; c) frame 7; d) image corresponding to *error* = 0.0464; e) image corresponding to *error* = 0.1535; f) image corresponding to *error* = 0.0343 (minimum error); g) *compensated object image*; h) *compensated mask*.

Figure. 3. Segmented frame 3 of *Claire*.

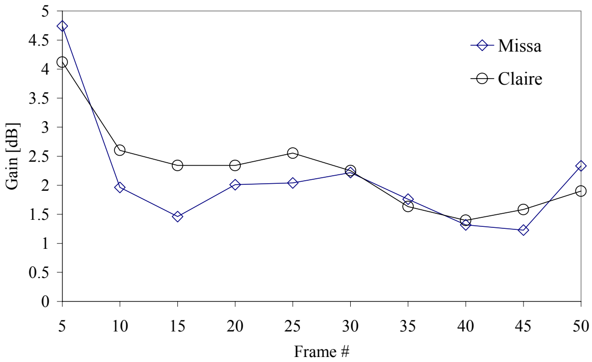

Figure. 4. *Missa* and *Claire* video sequences: gains in dB.

References

[1] A. Stoffels, T. Roska and L. O. Chua, "On object-oriented video-coding using the CNN Universal Machine," IEEE Trans. on CAS, part I, Vol. 43, No. 1, pp. 948-952, 1985.

[2] A. Stoffels, T. Roska and L. O. Chua, "Object-oriented image analysis for very-low-bitrate video-coding systems using the CNN Universal Machine," Int. J. Circuit Theory and Applications, Vol. 25, pp. 235-258, 1997.

[3] G. Grassi and L. A. Grieco, "Object-Oriented Image Analysis via Analogic CNN Algorithms – Part II: Image Synthesis and Consistency Observation," in Proceedings of 7th IEEE Int. Workshop on Cellular Neural Networks and Their Applications, Frankfurt/Main, Germany, July 22-24, 2002.

[4] CNN Software Library, Budapest, 2000, available at http://lab.analogic.sztaki.hu/.