Advances in Computational Intelligence

11th International Work-Conference on Artificial Neural Networks, IWANN 2011
Torremolinos-Málaga, Spain, June 8-10, 2011
Proceedings, Part I
Organization

IWANN 2011 Organizing Committee

Honorary Chairs
Alberto Prieto
Francisco Sandoval

University of Granada
University of Malaga

Conference Chairs
Joan Cabestany
Ignacio Rojas
Gonzalo Joya

Polytechnic University of Catalonia
University of Granada
University of Malaga

Technical Program Chairs
Francisco Garcia
Miguel Atencia

University of Malaga
University of Malaga

Satellite Worshops Chairs
Juan M. Corchado
Jose Bravo

University of Salamanca
University of Castilla la Mancha

Publicity and Publication Chairs
Pedro Castillo
Alberto Guillen
Beatriz Prieto

University of Granada
University of Granada
University of Granada

IWANN 2011 Program Committee

Plamen Angelov
Cecilio Angulo
A. Artes Rodriguez
Antonio Bahamonde
R. Babuska
Sergi Bermejo
Piero P. Bonissone
Andreu Catala
Gert Cauwenberghs
Jesus Cid-Sueiro
Rafael Corchuelo

University of Lancaster
Polytechnic University of Catalonia
University of Carlos III, Madrid
University of Oviedo
Delft University of Technology
Polytechnic University of Catalonia
GE Global Research
Polytechnic University of Catalonia
University of California, San Diego
University of Carlos III, Madrid
University of Seville
<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Óscar Cordón</td>
<td>European Centre for Soft Computing</td>
</tr>
<tr>
<td>Carlos Cotta</td>
<td>University of Malaga</td>
</tr>
<tr>
<td>Marie Cottrell</td>
<td>University of Paris I</td>
</tr>
<tr>
<td>Alicia D’Anjou</td>
<td>University of the Basque Country</td>
</tr>
<tr>
<td>Luiza De Macedo Mourelle</td>
<td>State University of Rio de Janeiro (UERJ)</td>
</tr>
<tr>
<td>Dante Del Corso</td>
<td>Polytechnic of Turin</td>
</tr>
<tr>
<td>Angel P. del Pobil</td>
<td>University of Jaume I, Castellon</td>
</tr>
<tr>
<td>Richard Duro</td>
<td>University of A Coruña</td>
</tr>
<tr>
<td>Marcos Faundez-Zanuy</td>
<td>Polytechnic University of Mataro</td>
</tr>
<tr>
<td>J. Manuel Ferrández</td>
<td>Polytechnic University of Cartagena</td>
</tr>
<tr>
<td>Kunihiko Fukushima</td>
<td>Takatsuki, Osaka</td>
</tr>
<tr>
<td>Chistian Gamrat</td>
<td>CEA, Gif sur Yvette</td>
</tr>
<tr>
<td>Patrik Garda</td>
<td>University Paris Sud, Orsay</td>
</tr>
<tr>
<td>F. Javier Gonzalez Cañete</td>
<td>University of Malaga</td>
</tr>
<tr>
<td>Karl Goser</td>
<td>University of Cordoba</td>
</tr>
<tr>
<td>Manuel Graña</td>
<td>University of the Basque Country</td>
</tr>
<tr>
<td>Anne Guerin-Dugue</td>
<td>Institut National Polytechnique de Grenoble</td>
</tr>
<tr>
<td>Hani Hagras</td>
<td>University of Essex</td>
</tr>
<tr>
<td>Alister Hamilton</td>
<td>University of Edinburgh</td>
</tr>
<tr>
<td>Jeanny Hérault</td>
<td>GIPSA-Lab, INPG, Grenoble</td>
</tr>
<tr>
<td>Luis Javier Herrera</td>
<td>University of Granada</td>
</tr>
<tr>
<td>Francisco Herrera</td>
<td>University of Granada</td>
</tr>
<tr>
<td>Cesar Hervás</td>
<td>University of Cordoba</td>
</tr>
<tr>
<td>Tom Heskes</td>
<td>Radboud University Nijmegen</td>
</tr>
<tr>
<td>Pedro Isasi</td>
<td>University of Carlos III, Madrid</td>
</tr>
<tr>
<td>Simon Jones</td>
<td>University of Loughborough</td>
</tr>
<tr>
<td>Christian Jutten</td>
<td>GIPSA-lab/DIS - CNRS - Grenoble University</td>
</tr>
<tr>
<td>Kathryn Klemic</td>
<td>Yale University</td>
</tr>
<tr>
<td>Amaury Lendasse</td>
<td>Helsinki University of Technology</td>
</tr>
<tr>
<td>Kurosh Madani</td>
<td>University of Paris XII</td>
</tr>
<tr>
<td>Jordi Madrenas</td>
<td>Polytechnic University of Catalonia</td>
</tr>
<tr>
<td>Luís Magdalena</td>
<td>ECSC Mieres</td>
</tr>
<tr>
<td>Dario Maravall</td>
<td>Polytechnic University of Madrid</td>
</tr>
<tr>
<td>Bonifacio Martín Del Brio</td>
<td>University of Zaragoza</td>
</tr>
<tr>
<td>Francesco Masulli</td>
<td>University of La Spezia, Genoa</td>
</tr>
<tr>
<td>Jose M. Molina</td>
<td>University of Carlos III, Madrid</td>
</tr>
<tr>
<td>Augusto Montisci</td>
<td>University of Cagliari</td>
</tr>
<tr>
<td>Claudio Moraga</td>
<td>European Centre for Soft Computing</td>
</tr>
<tr>
<td>Juan M. Moreno</td>
<td>Polytechnic University of Catalonia</td>
</tr>
<tr>
<td>Klaus-Robert Muller</td>
<td>FIRST, Berlin</td>
</tr>
<tr>
<td>Jose Muñoz</td>
<td>University of Malaga</td>
</tr>
<tr>
<td>Alan F. Murray</td>
<td>Edinburgh University</td>
</tr>
<tr>
<td>Jean-Pierre Nadal</td>
<td>Normal Superior School, Paris</td>
</tr>
</tbody>
</table>
Nadia Nedjah
Erkki Oja
Madalina Olteanu
Julio Ortega
Kevin M. Passino
Witold Pedrycz
Francisco Pelayo
Vincenzo Piuri
Hector Pomares
Carlos G. Puntonet
Leonardo Reyneri
Eduardo Ros
Ulrich Rueckert
Eduardo Sanchez
Jordi Solé-Casals
Peter Szolgyay
John Taylor
Carme Torras
I. Burhan Turksen
Mark Van Rossum
Marley Vellasco
Alfredo Vellido
Michel Verleysen
Thomas Villmann
Changjiu Zhou
Ahmed Zobaa
Pedro Zufiria
State University of Rio de Janeiro
Helsinki University of Technology
University of Paris I
University of Granada
The Ohio State University USA
University of Alberta
University of Granada
University of Milan
University of Granada
University of Granada
University of Granada
University of Paderborn
LSI, EPFL
University of Vic
Pazmany Peter Catholic University
Kings College London, UK
Polytechnic University of Catalonia
TOBB Econ Technol. University, Ankara
University of Edinburgh
Pontif. Catholic University of Rio de Janeiro
Polytechnic University of Catalonia
Catholic University of Louvain-la-Neuve
University of Leipzig
Singapore Polytechnic
University of Cairo
Polytechnic University of Madrid

ISCIF 2011 Program Committee

José M. Molina (Co-chair) Universidad Carlos III (Spain)
Juan M. Corchado (Co-chair) University of Salamanca (Spain)
Jesús García (Co-chair) Universidad Carlos III (Spain)
Javier Bajo (Co-chair) Pontifical University of Salamanca (Spain)
James Llinas (Co-chair) University of Buffalo (USA)
Sara Rodríguez University of Salamanca (Spain)
Juan F. de Paz University of Salamanca (Spain)
Carolina Zato University of Salamanca (Spain)
Fernando de la Prieta University of Salamanca (Spain)
Miguel Angel Patricio Universidad Carlos III (Spain)
Antonio Berlanga Universidad Carlos III (Spain)
Juan Gómez Universidad Carlos III (Spain)
José María Armingol Universidad Carlos III (Spain)
Moises Sudit University of Buffalo (USA)
Optimization Algorithms in Graphic Processing Units
Antonio Mora University of Granada
Maribel García-Arenas University of Granada
Pedro Castillo University of Granada

Computing Languages with Bio-inspired Devices
M. D. Jimenez-Lopez University of Rovira i Virgili
A. Ortega De La Puente Autonomous University of Madrid

Computational Intelligence in Multimedia
Adriana Dapena University of A Coruña
Julio Bregáins University of A Coruña
Nicolás Guil University of Malaga

Biologically Plausible Spiking Neural Processing
Eduardo Ros University of Granada
Richard R. Carrillo University of Almeria

Video and Image Processing
Enrique Domínguez University of Malaga
José García University of Alicante

Hybrid Artificial Neural Networks: Models, Algorithms and Data
Cesar Hervás University of Cordoba
Pedro Antonio Gutiérrez University of Cordoba

Advances in Machine Learning for Bioinformatics and Computational Biomedicine
Paulo J.L. Lisboa Liverpool John Moores University
Alfredo Vellido Polytechnic University of Catalonia
Leonardo Franco University of Malaga

Biometric Systems for Human–Machine Interaction
Alexandra Psarrou University of Westminster
Anastassia Angelopoulou University of Westminster
C.M. Travieso-Gonzlez University of Las Palmas de Gran Canaria
Jordi Solé-Casals University of Vic
### Computational Intelligence in Multimedia Processing

A Novel Strategy for Improving the Quality of Embedded Zerotree Wavelet Images Transmitted over Alamouti Coding Systems .............................. 489  
Josmary Labrador, Paula M. Castro, Héctor J. Pérez–Iglesias, and Adriana Dapena

Applying Data Mining Techniques in a Wyner-Ziv to H.264 Video Transcoder ........................................................................................................... 497  
José Luis Martínez, Alberto Corrales-García, Pedro Cuenca, and Francisco José Quiles

On the Use of Genetic Algorithms to Improve Wavelet Sign Coding Performance ........................................................................................................... 505  
Ricardo García, Otoniel López, Antonio Martí, and Manuel P. Malumbres

Kernel-Based Object Tracking Using a Simple Fuzzy Color Histogram .......................................................................................................................... 513  
Juan Villalba Espinosa, José María González Linares, Julián Ramos Cózar, and Nicolás Guil Mata

---

### Biologically Plausible Spiking Neural Processing

Isometric Coding of Spiking Haptic Signals by Peripheral Somatosensory Neurons .............................................................................................................. 528  
Romain Brasselet, Roland S. Johansson, and Angelo Arleo

Context Separability Mediated by the Granular Layer in a Spiking Cerebellum Model for Robot Control ................................................................. 537  
Niceto R. Luque, Jesús A. Garrido, Richard R. Carrillo, and Eduardo Ros

Realistic Modeling of Large-Scale Networks: Spatio-temporal Dynamics and Long-Term Synaptic Plasticity in the Cerebellum .............................. 547  
Egidio D’Angelo and Sergio Solinas

Event and Time Driven Hybrid Simulation of Spiking Neural Networks .................................................................................................................. 554  
Jesus A. Garrido, Richard R. Carrillo, Niceto R. Luque, and Eduardo Ros

---

**Author Index** ........................................................................................................ 563
Computational Intelligence in Multimedia Processing

Nicolás Guił\textsuperscript{1}, Julio C. Bregáins\textsuperscript{2}, and Adriana Dapena\textsuperscript{2}

\textsuperscript{1} Departamento de Arquitectura de Computadores. Universidad de Málaga.
Boulevar Louis Pasteur.
Campus de Teatinos, 29071, Málaga, Spain
Tel.: +34-952-133327
nguil@uma.es

\textsuperscript{2} Departamento de Electrónica e Sistemas.
Universidade da Coruña
Campus de Elviña s/n
15071, A Coruña, Spain
Tel.: ++34-981-167000
\{jbregains,adriana\}@udc.es

Abstract. Computational intelligence (CI) is a well-established paradigm that incorporates characteristics of biological computers (brains) to perform a variety of tasks that are difficult or impossible to do with conventional computers. This paper reviews some of the applications of CI in multimedia processing, including shot detection in video, logotype detection, video copy detection and retrieval, and faces coding in video sequences.

1 Introduction

Multimedia processing is a very important research domain with a broad range of applications that cover techniques like object-based representation and coding, segmentation and tracking, pattern detection and recognition, multi-modal signals fusion, content-based indexing, subject-based retrieval, etc.\textsuperscript{1} The utilization of Computational Intelligence (CI) in multimedia processing has allowed the inclusion of more sophisticated methods to solve these applications than those used in traditional computation. These methods are based on characteristics of biological computers (i.e. brains) and include techniques such as rough sets, neural networks, fuzzy logic, swarm intelligence, reinforcement learning and evolutionary computation. This paper reviews some applications of CI in multimedia processing. Section \textsuperscript{2} focuses on shot detection, Section \textsuperscript{3} presents logotype detection and its applications, Section \textsuperscript{4} presents the problem of video copy detection and retrieval and, finally, Section \textsuperscript{5} focuses on face recognition and its utilization in face coding in video sequences.

© Springer-Verlag Berlin Heidelberg 2011
2 Shot Detection in Video

As video demand continues growing, fast and reliable methods for indexing purposes are needed. One of the most useful techniques in video indexing is shot detection, where a shot is defined as a sequence of frames captured from a single camera operation. Shot detection is performed by means of shot transition detection algorithms. Two different types of transitions are used to split a video into shots:

- Abrupt transitions, also referred as cuts or straight cuts, occur when a sudden change from one shot to the next one is performed in just one frame.
- Gradual transitions use several frames to link two shots together, referred by other authors as optical cuts [2].

Depending on how the shots are mixed up in the transition, there are many different types of gradual transitions. Dissolves are the most common. Fades and wipes are also frequently used. What is most important about gradual transitions is that they are often used to establish some kind of semantic information in the video. For example, dissolves have been widely used to perform scene change in video edition, where a scene is a set of shots closely related in terms of place and time. While abrupt transitions detection is a relatively easy task [3], gradual transitions detection is still an open issue, as the amount of false positives reported by the algorithms is very high for certain sequences. The main problem in gradual transitions detection is that camera operation (pan, tilt, swing, zoom, etc.) originates similar patterns to those generated by gradual transitions [4]. Thus, a method to estimate global motion in video is needed in order to discard false positives induced by camera operations [5].

The first stage of shot transition detection algorithms is the extraction of characteristics from the video streams. One or more metrics are then used to compute several parameters from the characteristics. These metrics can be based on pixel luminance, contour information, block tracking, etc. Although most of the proposed methods make use of only one metric, using several of them is recommended, as drawbacks from one metric could be compensated by the others [6], as long as the used metrics rely on different video characteristics. The computed parameters are then used to determine the occurrence of a transition. Here, data driven methods address the problem from the data analysis point of view.

On the other hand, model driven methods, based on mathematical models of video data, allow a systematic analysis of the problem and the use of domain specific constraints, which helps to improve the efficiency [7], [8]. Other authors use non-deterministic classifiers to study the computed parameters in order to perform pattern recognition, as transitions generally result in a characteristic pattern in the parameters. Using a non-deterministic classifier makes unnecessary a specifically designed pattern recognition method, which usually needs several parameters to be tuned. Also, by using a supervised classification scheme, the system is able to learn the patterns generated by different types of gradual transitions. Some shot transition detection algorithms using neural

Neural networks have been proposed to detect abrupt shot transitions using four parameters extracted from video [13]. The parameters were computed using an evolution of the classical sliding window thresholding algorithm. Jun [10] proposed a neural network based algorithm for abrupt transitions detection using only I frames from the MPEG streams. Histograms and pixel-by-pixel comparison between frames are performed. Nevertheless, this last metric is very sensitive to motion. Moreover, only news videos were used to test the proposed method. A proposal which is highly dependent on the MPEG codifier was stated by Mallikarjuna [9], reporting results difficult to compare with other proposals.

Another work, [14], presents a reliable real time approach to temporal video segmentation in MPEG compressed video. A scheme based on luminance and contour information is proposed.

3 Logotype Detection and Learning

Logotypes convey information that can be crucial to infer semantics implicit in broadcasted videos. It is a common practice for broadcasters and specific TV programs, e.g., news, talk shows, advertisements, etc., to superimpose a specific logotype on the broadcasted material. Usually such logotypes refer to the actual video content or the content creator and, as such, they can be used to support automatic semantic-based video annotation. Logotypes extracted from broadcasted videos can be annotated in a database by indicating their shape and shots or scenes where they appear. Then, video retrieval tools can be used to search for a particular TV program or to group different pieces of video material with related contents. As a consequence, accurate logotype detection can be efficiently exploited for semantic-based video classification, video retrieval, aggregation and summarization.

Many works related to logo detection in document analysis have been reported in the literature [15]. However, few approaches consider logo detection in conventional video sequences. Logo detection techniques have been used to differentiate advertisements from TV programs in [16]. This approach assumes that a logo exists if an area with stable contours can be found in the image. The authors claim that their approach does not require any supervised training and can be easily used for any type of logos without human interaction. In [17], a neural network is trained using two sets of logo and non-logo examples to detect a transparent logo. It obtains a good detection rate at the expense of a rather large training set. In [18], color outliers are used to detect pixels different from the background. No temporal information is used, thus many false detections can arise.

The work presented in [19] shows an application for logo removing. The logotype is detected by exploiting frame differences in video sequence. This procedure fails in video with low motion activity. In this case, however, the authors propose to use a logo database and to search for them using a Bayesian approach. The
detection accuracy is improved by assuming that the probability of the logos appearing in the four corners of the video frames is higher than in the center. This prior knowledge is combined with a neural network-based classifier. Other works argue correctly that logos can provide a helpful visual cue for finding related news stories. Usually, a logo is defined as the small graphic or picture that appears behind the anchor person on the screen. In [20], it is assumed that each broadcasted TV channel contains some representative semantic objects, including the channel logotype, that is displayed only during news programs. Channel logotype detection and tracking is performed to automatically classify news events in conventional broadcasting material. The approach relies on the use of logotype models stored in a database. Information about logo position and scale helps to identify the channel and the type of news. In [21], the detection of logos is used to mark news stories, as an alternative approach for tracking them. Here, from each logo, three sets of 2D Haar coefficients are computed (one for each of the RGB channels). The logos feature vector is formed by selecting the coefficients representing the overall averages and the low frequency coefficients of the three color channels. However, in all these works, logos must be known a priori.

There are related works that try to identify known brand logotypes in video data. In [22], certain variability in the logotype appearance must be allowed. In practice, due to the high computational cost of this method, only a few logotypes can be identified simultaneously. Similarly, in [23], a search for specific instances of brand logos is performed. Logo detection is achieved by exploiting homogeneously colored regions surrounding large intensity frame differences.

In [24], a framework to support semantic based video classification and annotation is described. The backbone of the proposed framework is a technique for logotype extraction and recognition which is able to continuously update a metadata base in an iterative process where new logos are learned as they are detected and classified.

4 Video Copy Detection/Retrieval

Video copy detection carries out the comparison of a query video sequence with a target video sequence in order to establish if a copy of the query video is present in the target one. This application is a useful tool for both digital content management and protection of intellectual copyrights (IPR). Thus, on the one hand, the localization of video copies, e.g. an advertisement, can be used to catalog broadcasted material in a multimedia database, to check whether the material was broadcasted at the suitable time or if its duration was correct. On the other hand, the video copy can be compared with original video to check if any violation of IPR has happened. In this sense, video copy detection is an alternative to watermarking techniques when no marks can be inserted in the video copies.

Signature selection is a key point to develop a specific approach to video comparison. Spatial and/or temporal information can be used to generate the
video signature based on global or local descriptors [25]. Global descriptors are, in general, efficient to compute, compact in storage, but insufficiently accurate in terms of retrieval quality. Local descriptors present more invariance to aggressive transformations, such as PiP (Picture in Picture), cropping, insertions of patterns, change of gamma, etc.

In addition, signature content can be determined for every video frame or it can be given for a summary in the form of a set of key frames which is calculated through a temporal clustering process. These methods require less computational resources during the comparison process and generate shorter signatures. Kim and Park present in [26] an approach to video sequence matching by calculating the similarity between sets of key frames. Key frames are extracted using the cumulative directed divergence. Then, the similarity between videos is calculated by employing the modified Hausdorff distance between sets of key frames. Cheung and Zakhor [27] propose a measure for video similarity calculation where the percentage of clusters of similar frames shared between two video sequences is obtained. Previously, each video is summarized by selecting frames that are similar to a set of predefined feature vectors common to all video sequences. Guil et al. [28] divide the query video in clusters and extract a representative key frame for each cluster. Features from these key frames constitute the signature of the video. Then, they perform a dense comparison between the signature of the query video and every frame of the target video using relaxed distance constraints to speed-up the search process.

Comparison between query and target video sequences is carried out using some kind of similarity search. If the database of target videos is large, some efficient indexing process needs to be performed. Some research has been carried out on databases of more than ten thousand hours of videos, where the indexes have to be stored in hard disks. In this context, a trade-off between computing time and robustness must be considered. In this research line, Joly et al. have showed that trading quality for time during the search is highly profitable when statistical filtering is implemented, even when the size of the DB becomes very large [29]. When databases are not so large other authors [30] [31] address the problem of searching for repeated subsequences by hash tables. However, if both strong temporal and spatial transformation are applied to copied videos, more robust indexing techniques are needed. Thus, Douze et al., [32], use two refinements in order to carry out the query. Firstly, the feature descriptor is used as a quantization index which is augmented with a binary structure. Then, partial geometry consistency is calculated between the matching frames. This method obtains excellent results for the TRECVID 2008 copy detection task [33].

5 Coding Faces in Video Sequences

A facial recognition system is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source. Moghaddam and Pentland introduced the use of Principal Component Analysis (PCA) for face recognition from still images [34] and further explores
The main idea of using PCA for face recognition is to express the large 1-D vector of pixels constructed from 2-D facial image into the compact principal components of the feature space. This can be called eigenspace projection. Eigenspace is calculated by identifying the eigenvectors of the covariance matrix derived from a set of facial images (vectors) [37].

In applications like video-mobile the information to transmit is a sequence of the same face (maybe in different poses). To encode this kind of video sequences, Piqué and Torres [38] represent the frames using PCA and adapt the eigenspace taking into account the different poses, expressions and lighting conditions of the faces. The idea is to predict the actual frame by calculating their projections onto the eigenspace calculated from previous frames. The coefficients are therefore coded and transmitted. Full frames are only coded when a poor representation is obtained and, in this case, the eigenspace is updated. The quality of the recovered image is obtained by using a metric based on the peak signal to noise rate but it is also possible to include visual information. In [39], it has been introduced the idea of predicting the faces in two directions which allows us to define three kinds of frames (I, P, and B) similar to those used in MPEG.

Acknowledgements

This work was supported by Xunta de Galicia (grants numbers 10TIC105003PR and 09TIC008105PR), Junta de Andalucía (grants number P07-TIC-02800) and Ministerio de Ciencia e Innovación of Spain (grants numbers TEC2010-19545-C04-01 and CSD2008-00010).

References


