The PeRCeiVe Lab
From fine-grained pattern recognition to "big data"-driven computer vision

A. Faro, D. Giordano, S. Palazzo, and C. Spampinato

Department of Electrical, Electronics and Computer Engineering,
University of Catania - Viale Andrea Doria, 6 95125, Catania - Italy
{afaro,dgiordan,simone.palazzo,cspampin}@dieei.unict.it
http://perceive.dieei.unict.it

1 Extended Abstract

The Pattern Recognition and Computer Vision Lab - *PeRCeiVe Lab* - at the Department of Electrical, Electronics and Computer Engineering of the University of Catania is actively involved in pursuing cutting edge research on issues related to Computer Vision and Pattern Recognition and their applications to medical image analysis, ecological informatics and bioinformatics. More specifically, the PeRCeiVe Lab's research is focused on some topical aspects of:

- **Computer Vision and Image Understanding**. In particular, feature extraction, image segmentation, object recognition, motion analysis and scene understanding are topics which have been advanced by the research carried out at the PeRCeiVe Lab. Applications of the devised methods are mainly in the areas of environmental video monitoring, event detection, fine-grained visual categorisation and medical image analysis;
- **Pattern Recognition**. Pattern recognition research within the PeRCeiVe Lab has mainly investigated decision trees, deep learning methods, Hidden Markov Models and mixture models for applications in the medical imaging domain, bioinformatics, eye tracking and video content analysis. Methods for automatic generation of large scale annotated datasets and for reliable performance evaluation in pattern recognition are also under investigation;
- **Multimedia Content Analysis**. The focus of the research under this topic is on devising techniques for the analysis, processing, indexing and filtering of multimedia data. This includes a broad range of problems, from content extraction and understanding to similarity assessment mainly for information retrieval applications. Particular emphasis has been recently given to multimedia recorded for environment monitoring and to integration of heterogeneous multimedia data.

The lab's research has been recently driven by some new technology trends: the growth of computing power, digital cameras and information storage capacities in the past few decades has generated new possibilities for investigating

disparate "real-world" applications with the consequent generation of massive non-stationary visual data. Dealing with non-stationarity of big visual data is one of greatest challenges of the PeRCeiVe Lab. In particular, the PeRCeiVe Lab people have gained skills and experience in the following topics:

- *Performing automatic visual tasks in complex "real-life" environments.* Nowadays, it is becoming common practice to use digital cameras and sensors in a range of monitoring or exploratory applications, in particular for biological, geological and physical surveys. These technologies have revolutionised our ability to capture visual data on large scale for environmental monitoring and are also greatly improving our ability to effectively manage natural resources and increasing our competitiveness. For example, the Pl@ntNet project[1] has collected massive plant images for analyzing geographic distribution of plants in the Mediterranean area; the Fish4Knowledge project[2] has recorded many Terabytes of video data for monitoring Taiwanese coral reefs. While computer vision and pattern recognition approaches may run effectively in human-centered scenarios, the tides turn against them when more complex domains need to be modelled, as in the aforementioned scenarios. For example, the underwater environment is a rather complex scenario, where dynamic or multi-modal backgrounds, abrupt lighting changes, light propagation in water, radical and instant water turbidity changes greatly complicate the development of automated visual systems. The lab's research is mainly focused on computer vision and pattern recognition approaches for detecting and recognising events based on interactions between targets in noisy environments, as well as retrieving objects by integrating multiple 2D distorted views over time. The PeRCeiVe Lab people have gained a lot of experience on this topic as to be considered the pioneers on computer vision in the underwater domain, and more in general on the "ecological multimedia" field. The PeRCeiVe Lab is also involved in the organisation of several international challenges, e.g. the fish task of the ImageCLEF 2014 competition[3] or the data challenge track of the MAED 2014 workshop[4].
- *Generation of large-scale labeled datasets.* Many of the existing computer vision and pattern recognition approaches are not able to scale up to many scenarios due mainly to the lack of large labeled training sets. Large scale labeled datasets (*ground truth*) are, therefore, of key importance for the development of such methods as they, from one hand, allow training of multi-class classifiers and, from the other hand, support the algorithms' evaluation phase. The main limitation to collect large-scale and high quality ground truth is the daunting amount of time and human effort needed to generate it. To address these needs a significant part of the recent lab's research has been directed towards developing collaborative and crowdsourcing tools to annotate objects, their shape, their motion (in the case of videos) and

[1] http://www.plantnet-project.org/
[2] http://www.fish4knowledge.eu/
[3] http://www.imageclef.org/2014/lifeclef/fish
[4] http://maed2014.dieei.unict.it

the image/video containing them. Computer vision methods and scalable clustering approaches working at large-scale have been developed to support annotators in their labelling tasks. Data-driven methods for assessing object detectors and trackers performance without ground truth data are also other explored research topics. The PeRCeiVe Lab people have organised workshops (see the VIGTA workshop series) and are involved in guest-editing special issues on top-ranked journals (see, for instance, the "Large Scale Evaluation in Computer Vision" special issue of the Computer Vision and Image Understanding journal) on this topic.

– *Heterogeneous Data Integration.* Achieving a satisfactory level of adaptation and flexibility of automated visual systems requires 1) tracking of the behaviour the developed solutions under different circumstances, 2) performing sensitivity analysis of data features on the performance of computer vision components, and 3) predictive modelling of the performance of used components. Accomplishing these tasks requires the integration of different ground-truth datasets coming from variety of sources for which it is necessary to gather provenance information, associated trust assessments and other contextual metadata needed to support the reasoning necessary to assess to what extent training and evaluation data can be repurposed for other scenarios later. A key aspect in this process is to integrate heterogeneous data regarding operators, sensors, scenarios, etc, and this is pursued within the PeRCeiVe Lab by adopting best practices in the open data and semantic web communities. In particular, data schemas and ontologies, including SKOS for taxonomic data, Dublin Core for publication data, PROV-O for provenance tracking, and Open Annotation for user-contributed metadata are deployed by the PeRCeiVe Lab people in order to model the whole research process in computer vision and pattern recognition, from dataset to annotation to performance evaluation. However, heterogeneous data integration at the PeRCeiVe Lab has not only been used for computer vision problems, but also in bioinformatics applications where textual data (e.g. scientific literature), image data (microarray and miRNA) and biological annotated - semantic - databases are integrated for knowledge discovery.
– *Big Visual Data.* With the explosion of visual data on the web, computer vision researchers have started to direct their attention towards the development of non-parametric data-driven models for solving complex visual tasks (which at the moment are far from being solved) such as object detection, recognition and scene understanding and promising preliminary results have been obtained. The PeRCeiVe Labs research under this scenario aims at using big visual data for addressing the issues of generating large scale ground truth and, consequently, to develop data-driven approaches for solving visual tasks from object classification to scene understanding.

As further activities, people of the PeRCeiVe Lab regularly organise and chair dedicated workshops and special sessions at mainstream conferences such as ACM MM, ECCV, ICIP, ICPR, etc. The PeRCeiVe Lab has been involved in several national and international projects on computer vision and pattern

recognition topics, e.g. the Fish4Knowledge project funded under the EU FP7 or the AQUACAM research program - a joint research program between the University of Catania, the University of Edinburgh and the Universities of West Indies. Furthermore, the PeRCeiVe Lab routinely collaborates with several ICT companies and promotes technology transfer activities, especially for:

- *Biomedical Image and Signal Processing* for computer-aided diagnosis (CAD) systems. The system for automatic 3D striatum reconstruction implemented at the PeRCeiVe Lab is currently under testing at the REM radioterapia[5] (a nuclear medicine imaging company). Another example is the TMS tool, a system which enables neuroscientists to design, carry out and analyze scientific studies based on TMS (Transcranial magnetic stimulation) experiments for both diagnostic and research purposes, assisting them not only in the practicalities of administering the TMS but also in each step of the entire studys workflow. This system is currently used in several neuroscience departments around Italy;
- *Environmental Monitoring*. The CARIBSAVE organisation[6] based in Caribbean is testing the automatic underwater video analysis tools developed within the PeRCeiVe Lab for a more comprehensive analysis of fish biodiversity in marine protected areas (MPA) located in Jamaica.
- *Industrial vision applications*. Several companies are using the machine vision tools developed at the PeRCeiVe Lab for automatic visual tasks. The METRA srl[7] has adopted the machine vision tool developed at our lab for calibrating automatically temperature sensors.

References

1. Giordano, D., Palazzo, S., Spampinato, C.: Nonparametric label propagation using mutual local similarity in nearest neighbors. To appear on Computer Vision and Image Understanding (2014)
2. Palazzo, S., Kavasidis, I., Spampinato, C.: A texton-based kernel density estimation approach for background modeling under extreme conditions. Computer Vision and Image Understanding **122** (2014) 74 – 83
3. Boom, B.J., et al.: A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage. Ecological Informatics (2013)
4. Spampinato, C., et al.: Discovering biological knowledge by integrating high-throughput data and scientific literature on the cloud. Concurrency and Computation: Practice and Experience (2013)
5. Spampinato, C., et al.: A rule-based event detection system for real-life underwater domain. Machine Vision and Applications **25** (2014) 99–117
6. Kavasidis, I., et al.: An innovative web-based collaborative platform for video annotation. Multimedia Tools and Applications (2013) 1–20

[5] http://www.grupposamed.com/en/rem
[6] http://c-fish.org/about/caribsave/
[7] http://www.metraofficine.it/

7. Giordano, D., et al.: An integrated computer-controlled system for assisting researchers in cortical excitability studies by using transcranial magnetic stimulation. Computer Methods and Programs in Biomedicine **107** (2012) 4 – 15
8. Spampinato, C., Palazzo, S., Boom, B., Ossenbruggen, J., Kavasidis, I., Salvo, R., Lin, F.P., Giordano, D., Hardman, L., Fisher, R.: Understanding fish behavior during typhoon events in real-life underwater environments. Multimedia Tools and Applications (2012) 1–38
9. Faro, A., Giordano, D., Spampinato, C.: Adaptive background modeling integrated with luminosity sensors and occlusion processing for reliable vehicle detection. IEEE Transactions on Intelligent Transportation Systems **12** (2011) 1398–1412
10. Faro, A., Giordano, D., Spampinato, C., Ullo, S., Di Stefano, A.: Basal ganglia activity measurement by automatic 3-d striatum segmentation in spect images. IEEE Transactions on Instrumentation and Measurement **60** (2011) 3269–3280
11. Giordano, D., Spampinato, C., Scarciofalo, G., Leonardi, R.: An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysial/metaphysial bones. IEEE Transactions on Instrumentation and Measurement **59** (2010) 2539–2553
12. Faro, A., Giordano, D., Spampinato, C.: Evaluation of the traffic parameters in a metropolitan area by fusing visual perceptions and cnn processing of webcam images. IEEE Transactions on Neural Networks **19** (2008) 1108–1129
13. Giordano, D., Palazzo, S., Spampinato, C.: Kernel density estimation ysing joint spatial-color-depth data for background modeling. In: IEEE ICPR 2014. (2014)
14. Palazzo, S., Kavasidis, I., Spampinato, C.: Covariance based modeling of underwater scenes for fish detection. In: IEEE ICIP 2013. (2013) 1481–1485
15. Kavasidis, I., Spampinato, C., Giordano, D.: Generation of ground truth for object detection while playing an online game: Productive gaming or recreational working? In: 2013 IEEE CVPRW. (2013) 694–699
16. Spampinato, C., Palazzo, S.: Enhancing object detection performance by integrating motion objectness and perceptual organization. In: IEEE ICPR 2012. (2012) 3640–3643
17. Spampinato, C., Palazzo, S., Giordano, D.: Evaluation of tracking algorithm performance without ground-truth data. In: IEEE ICIP 2012. (2012) 1345–1348
18. Spampinato, C., Palazzo, S.: Hidden markov models for detecting anomalous fish trajectories in underwater footage. In: 2012 IEEE MLSP. (2012) 1–6
19. Giordano, D., Kavasidis, I., Pino, C., Spampinato, C.: Content based recommender system by using eye gaze data. In: ACM ETRA 2012. (2012) 369–372
20. Spampinato, C., et al.: Covariance based fish tracking in real-life underwater environment. In: VISAPP 2012. (2012) 409–414
21. Nadarajan, G., Chen-Burger, Y., Fisher, R., Spampinato, C.: A flexible system for automated composition of intelligent video analysis. In: 7th IEEE ISPA. (2011) 259–264
22. Giordano, D., Kavasidis, I., Pino, C., Spampinato, C.: A semantic-based and adaptive architecture for automatic multimedia retrieval composition. In: CBMI 2011. (2011) 181–186
23. Spampinato, C., et al.: Automatic fish classification for underwater species behavior understanding. In: ACM ARTEMIS 2010. (2010) 45–50
24. Spampinato, C.: Adaptive objects tracking by using statistical features shape modeling and histogram analysis. In: ICAPR '09. (2009) 270–273
25. Spampinato, C., Chen-Burger, Y.H., Nadarajan, G., Fisher, R. In: Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos. Volume 2. (2008) 514–519