STAG: Smart Tools and Apps in computer Graphics (2016)
Giovanni Pintore and Filippo Stanco (Editors)

Visual Computing at SLIPGURU

E. De Vito [1] F. Odone and F. Solari[2]

[1]DIMA - Università degli Studi di Genova, Italy
[2] DIBRIS - Università degli Studi di Genova, Italy

1. Introduction

SLIP GURU research group was funded in the late nineties with the aim of bringing together researchers of different by related fields. By acknowledging the key role played by learning in the development of intelligent systems, we investigate the foundations of learning and learning algorithms from the viewpoints of computer science, mathematics, and statistics. Starting from our long standing experience in the design of algorithms for extracting visual information from images, our efforts are directed toward computer vision applications. Leveraging on close collaborations with molecular biologists, clinicians, and radiologists we broadened our interests in the areas of medical image analysis and computational biology.

The research group consists today of 6 Faculty, 3 senior research associates and 15 PhD students in Computer Science, Bioengineering, Robotics, and Maths.

In the lab presentation we will focus on the aspects of our research that primarily address visual computing problems, referring more specifically to Signal Processing, Computational Vision, and Visual Perception. We will discuss our main research lines, describing the most significant results obtained in the recent, and summarizing our current research activities and goals.

For more information on our research activity we refer the interested user to our web site: http://slipguru.unige.it

1.1. Signal Processing

The main research theme is to develop multi-scale methods for signal and image processing and to design efficient algorithms for signal enhancement and feature detection. The focus is both on the mathematical theory in the framework of non-commutative harmonic analysis and on the applications in 2D/3D image analysis. About the theory we investigate the properties of signal representations associated with continuous frames parametrised by Lie groups, as for example the shearlet transform for 2D signals and its generalisation to higher dimensions. About the application we study, design, and develop algorithms for feature extraction and description in images (2D) and image sequences (2D+T). Figures 1 and **??** show results on feature detection and automatic scale estimation [DPODV15]. Instead, space-time feature detection with Shearlet is one of the research directions we are currently pursuing.

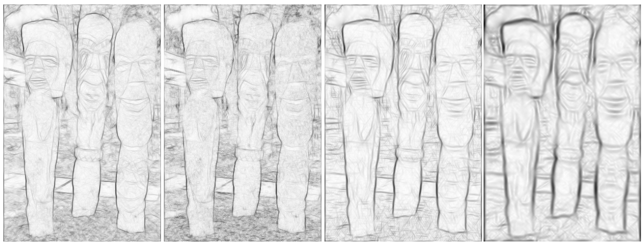

Figure 1: *A multi-scale representation of edges maps*

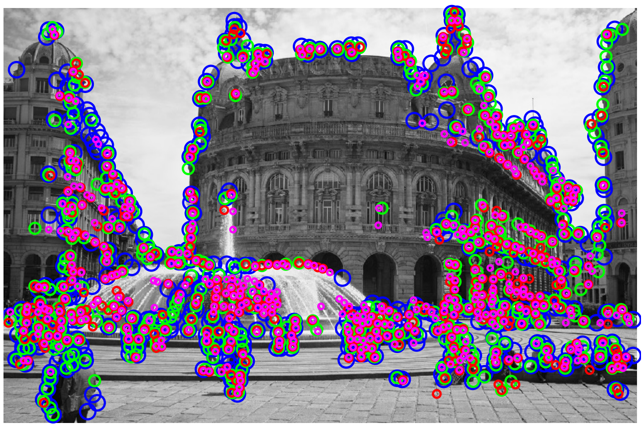

Figure 2: *Multi-scale corner detection*

Our research has also a potential applicability to the detection of meaningful points in 3D shapes.

1.2. Computational vision

With our research we explore the world of image and scene understanding combining computer vision and machine learning ingredients. Computer vision methods are used to extract information from the visual signals, while we resort to machine learning to model variability, and to gain robustness and flexibility. Within this framework we are pursuing research on object detection and object recognition as well as behavior analysis and action recogni-

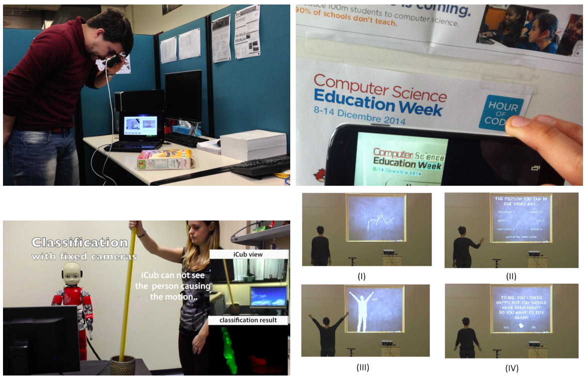

Figure 3: *A variety of computational vision systems recently developed by our research group (see text)*

tion. We apply the methods we develop in diverse contexts including human-robot interaction [FGMO13], computer vision for the visually impaired [CNO*16], video-surveillance [NO12]. Figure 3 shows an overview our very recent results; on top from the left: usability tests on a prototype system to support visually impaired users in recognizing specific object instances; a snapshot of an app running on an android phone performing text detection (notice that the entire computation is performed on the phone in real time). Bottom: testing a motion analysis method enabling a humanoid robot to attract its attention towards biological agents; a demo of a serious game prototype teaching autistic children to better express emotions through body movements.

Also, we devise biologically-inspired computational models of vision with a twofold aim: on one hand, to develop artificial vision systems that are able to face real-world action tasks, such as navigation and reaching, and on the other hand to exploit such models to better explain human perception [SCS14]. These models share similarities with the multi-scale methods addressed in signal processing, although they are inspired by a biological motivation, more than a mathematical framework.

1.3. Visual perception

In this research line, we analyze the effects of current visual technologies (e.g. 3D displays, mobile devices, virtual and augmented reality headsets) on the human perceptual system. In particular, we aim to assess the undesired effects on the users (such as the visual fatigue and the perceptual discomfort), and to evaluate the usability of such technologies in various fields of application. By exploiting the experimental evidence, the aim of our research is to develop new paradigms that allow us to mitigate such undesired effects and to develop natural human-computer interactions in virtual and augmented environments [CGR*15].

References

[CGR*15] Chessa M., Garibotti M., Rossi V., Novellino A., Solari F.: A virtual holographic display case for museum installations. *EAI Endorsed Trans. Serious Games 2*, 7 (2015), e3. 2

[CNO*16] Chessa M., Noceti N., Odone F., Solari F., Sosa-García J., Zini L.: An integrated artificial vision framework for assisting visually impaired users. *Computer Vision and Image Understanding 149* (2016), 209–228. 1

[DPODV15] Duval-Poo M. A., Odone F., De Vito E.: Edges and corners with shearlets. *Image Processing, IEEE Transactions on 24*, 11 (2015), 3768–3780. 1

[FGMO13] Fanello S. R., Gori I., Metta G., Odone F.: Keep it simple and sparse: real-time action recognition. *Journal of Machine Learning Research 14*, 1 (2013), 2617–2640. 1

[NO12] Noceti N., Odone F.: Learning common behaviors from large sets of unlabeled temporal series. *Image Vision Comput. 30*, 11 (2012), 875–895. 1

[SCS14] Solari F., Chessa M., Sabatini S. P.: An integrated neuromimetic architecture for direct motion interpretation in the log-polar domain. *Computer Vision and Image Understanding 125* (2014), 37–54. 2