Percorso Autonomo Autorizzato

Title	**IMAGE AND SIGNAL PROCESSING FOR MULTIMEDIA APPLICATIONS AND ROBOTICS**
Coordinators	(DEIB, PoliMi): prof. Marco Marcon, Umberto Spagnolini, Stefano Tubaro
Scientific collaborations and partnerships	• (DEIB, PoliMi): prof. Fabio Rocca, prof. Claudio Prati • (DMEC, PoliMi): prof. Alfredo Cigada, D.R. Remo Sala • (DAER, PoliMi): prof.ssa Michèle Lavagna • (Dip. Elettronica, PoliTo): prof. Gabriella Olmo, prof. Enrico Magli • (Dip. Elettronica, UniBs): prof. Riccardo Leonardi • (Dip. Elettronica, UniUd): prof. Roberto Rinaldo • (Dip. Elettronica, UniSi): prof. Mauro Barni • (Queen Mary University, London): prof. Andrea Cavallaro • (Imperial College of London): prof. Patrick Naylor, prof. Pier Luigi Dragotti • (Istituto Superior Tecnico, Lisbon): prof. Fernando Pereira • (Humboldt University Berlin): prof. Peter Eisert • (EPFL, Lausanne): prof. Martin Vetterli • (Nuremberg-Erlangen University): prof. Walter Kellermann, prof. Rudolf Rabenstein • (Uni Télécom ParisTech) : Béatrice Pesquet-Popescu • (Dept. of EE-Systems, USC, L.A.): prof. Antonio Ortega • (Mississippi State University): prof. James E. Fowler • (Boston University): prof. Janusz Konrad • (New Jersey Inst. Of Tech., NJ): prof. O.Simeone, A.Haimovich • (U.C. Berkeley): prof. Kannan Ramchandran
Description and aim	Images and videos are playing an increasingly role both in interpersonal communication and in human-computer interaction. At the end of 2014 about 2 billion images were daily uploaded on social networks while on YouTube® 300 hours of video are uploaded every minute. At the same time visual information have a growing role in many applications: robotics, industrial control, autonomous vehicles, medical diagnosis. Multimedia signals acquired from a plethora of different sensors together with advanced digital signal processing techniques provide us with a very detailed description of the inputs from the surrounding environment. Pattern recognition techniques together with robust features descriptors are able to provide high-level information that allow, for example, moving a robot in *a priori* unknown contexts. The scientific and industrial communities are extremely active in the development and deploying of the multimedia acquisition and processing technologies, with a remarkable production of patents and scientific papers. On one side, tablets and smart-phones have completely replaced the traditional photo/video cameras in our daily usage. At the same time, applications and services have redesigned our life: for example YouTube® has redefined the video content acquisition and fruition paradigm, while the file-sharing applications based on peer-to-peer networks have forced a complete revision of the copyright issues for audio/video contents. On the other side, everyday tools and devices changed accordingly: e.g.. Furthermore new services become available, allowing visual search inside photographic archives (search all images similar to a reference object, search all images where the grandmother's face is present, ...) and new ethical questions arise: How can I believe the truthfulness of a photograph or a video when they can be easily manipulated in a realistic way? This revolution generated a completely new series of applications, research areas and jobs: for example, now television programs and movies routinely mix real and virtual contents in seamless way. In fact the joint use of computer vision, image processing and computer graphics techniques allows to seamless fuse contents in a realistic way. Applications like Boujou®, made by image processing experts (http://www.vicon.com/boujou), have completely changed the cinematographic postproduction.

	Another kind of application field like "motion capture", which capture the movements through the combined use of multiple cameras, built the basis of modern medical rehabilitation techniques and animation of virtual actors. The visualization of 3D data and virtual real time interaction is becoming more and more interactive thanks to modern computer graphics techniques increasingly immersive experience between the user and the 3D digital environment. In Industrial Quality Inspection, vision systems are continuously increasing their relevance assessing, with sophisticated techniques, the quality of each individual product: the analysis of images acquired along the production chain allows, for example, to discard defective pieces and to optimize robot actions. Compliance with the quality requirements of banknotes, tiles, fabrics, but also fruits or meat portions are nowadays is entrusted to vision systems and statistical image processing techniques. Automatic (or assisted) driving systems for vehicles massively use information obtained the processing of images captured by multiple cameras placed on the vehicles. Also in gaming applications, image processing and computer graphics techniques are expanding their role: the Microsoft Kinect® system, for example, offers the players a full interactive interface without any joystick or wired device. It is based on 3D cameras for automatic gesture recognition of every player and accurate avatars and virtual environments are reconstructed thanks to fast computer graphics algorithms implemented in GPUs. Satellites or subsurface images represent further application domains where microwave antennas on a satellite or seismic sensors collects data arranged in multidimensional arrays. Then image processing techniques coupled to complex wave propagation models are able to create images that describe with millimeter accuracy the earth surface or its interior. A similar evolution occurs in the audio field, where audio signal processing is deeply changing our interaction with mobile devices (smartphones, tablets). For example, a series of applications is now providing a natural language interaction with these devices (e.g. SIRI on IOS platform, or "Google Now"), and novel applications are coming out inferring the user's mood by voice and movement, or their musical preferences from the content of his/her archives. Nowadays there are also of-the-shelf devices able to localize the speaker in a 3D environment and to focus on his voice to better recognize and interpret what he is saying. A particularly promising area of audio signal processing is based on music information retrieval, where statistical techniques are able to extract descriptors at different levels of abstraction (semantics), assisting the listener to browse the huge amount of music contents available on the net. Some products are already available for generating custom audio streams (eg. StereoMood, Lastfm, Spotify). As a last remark, it must be underlined that music and video production areas are increasingly involving people with both artistic skills and deep knowledge on signal processing.
Study Plan	The programme of studies belongs to the Major in "Applied Statistics" with additional courses in image and video processing. The list of courses can be found in a separate document.
Past MSc theses	• L. Bonacina (Ing. Inf.), *3D Models Extraction For Personalized Binaural Audio Applications*, 2015 • A. Camarda (Ing. Inf.), *Scene Classification for Mountain Landscape Recognition from User Generated Images,* 2014 • M. Paracchini (Ing. Mtm.), *Localizzazione Robusta dei Punti Salienti del Volto Tramite un Approccio a Regolarizzazione Globale,* 2014 • L. Gaborini (Ing. Mtm.), *Image tampering detection and localization,* 2014 • F. Raimondi (Ing. Mtm), Signal processing for passive seismic and application to Argentière glacier, 2014 • P.F. Piazza (Ing. Mtm.), *Analisi congiunta di flussi audio-video per l'estrazione di informazioni sulla volumetria delle scene riprese,* 2013 • A. Bolognino (Ing.Tlc), *Consensus-based inference methods and cooperative wireless*

	networks, 2013 • G. Soatti (Ing.Tlc), *Distributed estimation of multi-link channel in dense cooperative networks*, 2013 • G. Sandrini (Ing. Mtm.), *Acoustic Imaging in the Ray Space: Application to Environment Inference*, 2012 • S. Battaglia, (Ing. Inf.), *Individuazione della moto-interpolazione in sequenze video*, 2012 • A. Panichella (Ing. Inf.), *Identificazione del cambio di bitrate attraverso modelli di stima della qualità in modalità no-reference*, 2012 • D. Totaro (Ing. Inf.), *Example-base definition of high-level descriptors of musical excerpts,* 2012 • D. Rodriguez Salgado (Ing. Inf.), *Music recommendation system based on audio segmentation and feature evolution*, 2012 • L. Chiarandini (Ing. Tlc.), *Automatic audio compositing system based on music information retrieval*, 2012
Available subjects for an MSc thesis	• *Compressed sensing and sparse recovery* • *Statistical signal processing over sparse geometric graphs* • *Video tampering detection through compression history analysis, estrazione di parametri di compressione dal bitstream per cercare tracce di compressione multipla* • *Video tampering detection exploiting steganalitic methods, applicazione di metodi di steganalisi per la detection di modifiche video* • *Video splicing detection and localization using no reference quality metrics, uso di statistico di metriche no-reference nell'analisi video per riconoscere manipolazioni* • *Who's best? Image/video quality comparison, ordinare sequenze video o immagini in base alla loro qualità (blurring, logo insertion, crop, etc.)* • *Odometria visuale e ricostruzione di scena 3D da singola telecamera in moto* • *Stima diretta delle mappe di visibilità acustica basata su analisi di segnali audio acquisiti con schiere di microfoni* • *Co-registration and calibration of a low-res 3D ToF camera with a Hi-Res camera* • *Graph analysis for deformable meshes in space* • *Feature-based on-the-fly genre classification of audio streams* • *Non-invasive affective browsing of audio content* • *Ottimizzazione delle sensitività spettrali in sensori CMOS multi-canale per imaging multi-banda* • *Sensori d'immagine CMOS filterless in tecnologia backside* • *Sensori RGB/IR per rivelazione di immagini 3D* • *Sensori riconfigurabili ad elevato range dinamico per acquisizione di immagini a colori in condizioni di illuminazione fortemente variabili* • *Simulazione microelettronica di sensori sensibili al colore basati sulla diffusione* • *Relazione tra sharpening dei filtri e progetto delle microlenti in sensori sensibili al colore* • *Studio ed ottimizzazione della intersezione tra lo spazio cromatico visibile e quello di un sensore di acquisizione di immagini con filtri accordabili*
Industrial internships	On students' demand it is possible to activate internships with Italian or foreign companies.
Work opportunities	The proposed study plan offers a specific and highly qualified education in multimedia contents acquisition, processing and broadcasting. After achieving the MSc many working opportunities will be available both in Italy and abroad. In particular, possible employing contexts range from large enterprises for electronic devices, like ST Microelectronics (where a significant part of their business focuses on embedded cameras for mobile devices and on-chip systems for image and video processing, restoration and enhancement) or companies related to development and deployment of *ad-hoc* systems for post-production in cinematographic and television movies production. Furthermore, video surveillance and security are other applicative contexts where image processing has a continuously growing role and where students following this plan can find a job.

	Earth monitoring and remote sensing are further domains where image and signal processing cover a relevant role: e.g. from the research in alternative energies to reduction of harmful emissions and environmental protection. Concerning audio signals processing possible working opportunities range from audio post-production, active and passive noise control, music instruments synthesis and audio footprint design for objects and consumer products (e. g. the noise from an exhaust pipe or from a car door closure can greatly influence our evaluation of the car itself!)