POLITECNICO DI MILANO

PhD in INGEGNERIA DELL'INFORMAZIONE / INFORMATION TECHNOLOGY - 38th cycle

Research Area n. 4 - Telecommunications

PNRR_352 Research Field: GESTURE ANALYSIS AND RECOGNITION BASED ON EGOMOTION ESTIMATION AND WEARABLE INERTIAL SENSORS FOR SMART EYEWEAR

Monthly net income of PhDscholarship (max 36 months)
€ 1400.0 In case of a change of the welfare rates during the three-year period, the amount could be modified.

Context of the research activity	
Motivation and objectives of the research in this field	The aim of the PhD research is to develop and integrate novel and robust techniques for egomotion action recognition and visual odometry in the smart eyewear: in the near future eyeglasses will be capable not only of correcting vision defects and providing an object of style and protection for eye health, but also of offering new AR (augmented reality) experiences, MR (mixed reality) and VR (virtual reality) and their XR (extended reality) superset that includes the entire spectrum from "completely real" to "completely virtual" in the concept of reality-virtuality continuum, for a new way of perceiving (sensing) the surrounding world, to interact with it (human interface), to monitor one's state of health, using Artificial Intelligence (AI). The potential of smart eyeglasses, equipped with inertial sensors and micro-projectors, able to deploy Augmented Reality realistic environments, became a disrupting activity in the last few years (e.g. Microsoft Hololens®); however, the lack of an intuitive and natural Human-Machine Interaction (HMI) and a very small dictionary of recognized gestures made these systems cumbersome and mainly oriented to a professional use with trained users. The outcome of the Ph.D. research will be focused on the development of a more natural and comprehensive

	interpretation of human gestures together with an action recognition engine that will benefit from the fusion of multiple sensors (e.g. Inertial units, depth-cameras, visible cameras, LIDARs) and the context understanding of the environment. The gestures will not be limited to some hands movements to be tracked by the onboard cameras but actions themselves will be estimated and predicted from multiple time sequences obtained from different sensors using advanced Machine Learning algorithms. The system developed will be validated on the new smart eyewear prototype in real scenarios, devising test and benchmark for the accuracy and quality evaluation of the gesture/action recognition on smart eyewear.
Methods and techniques that will be developed and used to carry out the research	The objective of this PhD research is to design, develop and validate a novel approach for Human-Machine Interaction with smart glasses and eyewear. In particular, the aim is to equip the smart glasses, provided with cameras, Lidars and inertial sensors, with a set of AI tools aimed to process, analyze and recognize gestures and actions performed by the user together with the specific context/environment where they are performed. The recognition algorithms will then be based on Egomotion analysis, visual odometry and context dependent gesture/action recognition. The whole system, integrated into the smart glasses, will use information from visible cameras, inertial units and depth cameras/LIDARs, and, in case of oudoor activity, the GPS data will also be used. Egomotion estimation consists in determining one self's motion. Humans do this remarkably well by fusing several sources of information e.g. their eyes, their sense of balance, their knowledge of the world and so on. It can be argued and past research has established that combining visual and inertial sensor readings to determine egomotion yields much better results than those obtained using only one type of reading. Merging the information from these sensors and processing it with a Machine Learning approach will allow the clear definition of an ontology where the role of different inputs and their interaction is well established in a context-dependent gesture/action understanding.

	It seems that cameras and inertial measurement units (IMU) complement each other very well : for discontinuous motions where the visual feature tracker fails to make sense of incoming images, the IMU excels, whereas in slower smoother motions where the IMU's readings are drowned in noise, the visual feature tracker performs well. The spatio-temporal analysis and understanding of the person's intent will be performed on the edge computing allowing the reduction of computational efforts on the smart eyewears and delegating the continuous-learning and the real-time processing tasks to the remote computer. An important aspect that will be handled during the implementation of the whole system is the user-dependent gesture interpretation: the system will acquire, in a continuous learning paradigm, the distinctive execution of different gestures from each user and will adapt and tune the recognition engine in a personalized manner. The device will then learn from the user and the user will adapt her/his gestures to the device in order to get a joint robust and synergistic interaction. The feedback to the user will be provide in the prototype both in audio format and in an Augmented Reality manner using available Mixed Reality Toolkits.
Educational objectives	The educational objectives are the study, investigation and identification of the enabling technologies for the development of intelligent eyewears, focusing on Egocentric Action Recognition, multiple sensor fusion, complex systems ontology description and continuous learning pattern recognition algorithms. Furthermore, the PhD student will work and collaborate with other researchers and PhD students on all the areas required to develop the "smart glasses".
Job opportunities	Thanks to the acquired knowledge in the state-of-the-art techniques for Gesture Recognition, Multi Sensor Fusion and Machine Learning, the Ph.D. student can be employed in a wide number of high-tech companies working on AI and frontier research for Human Machine interaction.

Composition of the research group	2 Full Professors 1 Associated Professors 10 Assistant Professors 4 PhD Students
Name of the research directors	Prof. Marco Marcon

Contacts
marco.marcon@polimi.it +39-02-2399.3582 https://www4.ceda.polimi.it/manifesti/manifesti/controller/ricerche/RicercaPerDocentiPublic.do?evn_didattica=evento&k_doc=100349&polij_device_category=DESKTOP&__pj0=0&__pj1=b4aad5c023f1f8e39470f7d797b71ec9

Additional support - Financial aid per PhD student per year (gross amount)	
Housing - Foreign Students	--
Housing - Out-of-town residents (more than 80Km out of Milano)	--

Scholarship Increase for a period abroad	
Amount monthly	700.0 €
By number of months	6

National Operational Program for Research and Innovation	
Company where the candidate will attend the stage (name and brief description)	LUXOTTICA s.r.l. (https://www.luxottica.com/it)
By number of months at the company	6
Institution or company where the candidate will spend the period abroad (name and brief description)	Université Paris-Saclay, Prof. Giuseppe Valenzise, Laboratoire des Signaux et Systmes (L2S), Centrale Supelec
By number of months abroad	6

Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information

Attinenza PNRR

La ricerca mira a sviluppare un nuova interazione uomo-macchina tramite l'impiego di occhiali intelligenti (smart eyewear) al fine di sviluppare un esperienza sensoriale interattiva tra Persona e Mondo, mettendo in connessione ciò che l'utilizzatore può vedere e sentire attorno a sé. Tramite l'elaborazione nel Cloud/Edge in tempo reale sarà così possibile elevare il ruolo dell'occhiale da dispositivo medico e accessorio di moda ad elemento tecnologico fondamentale nel nuovo Metaverso. Grazie alle informazioni acquisite dagli occhiali intelligenti tramite telecamere/LIDAR e sensori inerziali sarà possibile ottenere una descrizione dettagliata dell'interazione tra l'utilizzatore ed il mondo circostante, permettendo, anche tramite l'elaborazione nell'edge/cloud computing, di riconoscere i gesti e le azioni compiute. Sarà quindi possibile creare le basi per un'interazione uomo-macchina in modalità Mixed Reality (ossia dove oggetti e persone reali interagiscono con

persone ed oggetti virtuali) altamente immersiva ed intuitiva.
Queste tematiche sono coerenti con i fabbisogni del Paese, in termini di figure ad alta qualificazione e orientate a soddisfare i fabbisogni di innovazione delle imprese di cui al PNRR. In particolare, le Missioni del PNRR che trarranno maggior beneficio da questo progetto sono:
1. Digitalizzazione, innovazione, competitività, cultura e turismo, e in particolare M1C2 sistema produttivo e M1C3 Turismo e Cultura 4.0
2. Inclusione e Coesione e in particolare M5C2 Infrastrutture sociali, famiglie, comunità e terzo settore
3. Salute e in particolare M6C1 Reti di prossimità, strutture e telemedicina per l'assistenza sanitaria nazionale

Impresa, presso cui si svolgerà l'attività esterna
LUXOTTICA s.r.l.
L'attività svolta presso l'azienda prevederà l'integrazione degli algoritmi sviluppati nel prototipo di occhiali intelligenti e la conduzione di una serie di test finalizzati alla calibrazione ed alla verifica di corretto funzionamento.

Università presso cui si svolgerà il periodo di studio e ricerca all'estero
Université Paris-Saclay Prof. Giuseppe Valenzise, Laboratoire des Signaux et Systmes (L2S), Centrale Supelec
Sviluppo del sistema di odometria visuale basandosi sulle telecamere e sensoristica LiDAR della quale sarà equipaggiato il prototipo di occhiali intelligenti: il sistema si focalizzerà sull'analisi e coregistrazione rapida di nuvole di punti ottenuti da un sistema avanzato di SLAM (Simultaneous Location And Mapping) che sarà implementato.

All information regarding educational activities, personal funding, regulations and obligations of Ph.D. candidates are available on the web site https://dottoratoit.deib.polimi.it/