

PhD Thesis Proposal in Computer Science (MAP-i)

Title: An Integrated System for Detection and Identification of Behaviors and Biometric data

Background:

Due to advances in Information Technologies, currently it is easy to collect, store and analyze data. In particular, a recent interest has been given to the use of 3D video cameras, which are becoming more affordable and more universally used. This PhD is performed within this context and assumes an interdisciplinary work, between the several Computer Science areas, such as Computer Vision, Human-Computer Interaction and Artificial Intelligence (with a particular emphasis on Data Mining). A synopsis of the related background is described below.

One interesting research and recent trend relies in identification of human activity from video (Turaga, 2008). While several temporal and spatial models have been proposed to identify these activities (Weinland et al., 2011), there are very few relevant studies that use 3D video as a biometric technique. These few studies focus more on using only one biometric aspect, such as use of facial expressions (Bowyer et al., 2006) or walking patterns (Yam et al., 2004), thus lacking the ability to integrate multiple human signals.

This PhD is associated with the R&D QREN project Co-Promoção N° 21584 (Concurso 03/SI/2011), which aims at developing a real-time system capable of analyzing 3D video, under a multi-model biometric approach (Jain, 2004). The goal is to develop a prototype system that can support the management of statistical information related with human behavior, which is vital for decision-making. For instance, such system could be set inside a commercial store window, to assess if a given human is a potential customer, neutral person or a possible threat (e.g. thief). A particular focus is given towards face and gesture recognition. Several studies have addressed this issues separately, such as: face detection - (Lee et al., 2003)(Zhao et al., 2003); and gesture recognition – (Liu et al., 2009)(Fenn, 2010). It should be noted that video gesture recognition is a very recent trend, since previous works, such as (Liu et al., 2009), are mostly based on sensors (e.g. accelerometers). Moreover, in the novelty of this PhD is that a multi-modal approach is assumed, where several biometric aspects, such as gender, face and gesture aspects will be integrated, in order to achieve novel capabilities in terms of interesting human behavior detection.

Research question: To which degree is it possible to develop a real-time computer system that is capable of detecting useful human behaviors, using as input video taken from 3D cameras and integrating multi-modal biometric aspects, such as gender, face and gesture recognition?

Goals:

To develop an integrated system that is capable of a real-time analysis of video taken from 3D cameras in order to detect useful human behaviors. From the video, several biometric aspects will be identified, such as gender, face and gestures. Using such meta-data, the proposed system will be tested in a real environment (e.g. commercial window store), to check if it can provide useful knowledge to support decision-making.

Methodology:

There are two major research methodologies that can guide this research. Given the novelty and nature of the problem to be addressed, an important part of the research will adopt an **exploratory research** approach (Jaeger and Halliday, 1998), where several Computer Vision and Artificial Intelligence techniques will be tested to check what sort of useful human behaviors can be detected within a given scenario (e.g. commercial store). On the other hand, since this PhD assumes development of a prototype, which can be seen as a the proposal of a novel artifact, this work will

follow a **Design Science** research methodology (Vaishnavi and Kuechler, 2004). Regarding the application of Data Mining techniques, these will be guided by the CRISP-DM methodology (Chapman et al., 2000).

This PhD is structured into four major tasks: 1) gender detection; 2) facial recognition and detection of human behaviors; 3) Implementation of the system prototype; and 4) testing the implemented prototype in a real-environment.

Relevant conferences and journals:

Given the interdisciplinary nature of this work, there is a huge list of potential scientific conferences and journals. We will aim at high quality events, indexed in internationally recognized scientific databases (e.g. ISI, Scopus), from the areas of Computer Vision, Human-Computer Interaction and Artificial Intelligence.

Examples of relevant conferences are:

ACM Conference on Human Factors in Computing Systems (CHI)
ACM Symposium on Engineering Interactive Computing Systems Engineering Interactive Computing Systems (EICS)
International Joint Conference on Ambient Intelligence
International Conference on Computational Science and Its Applications (ICCSA)
International Joint Conference on Artificial Intelligence (IJCAI)
International Conference on Enterprise Information Systems (ICEIS)
European Conference on Machine Learning (ECML)

Examples of relevant journals are:

IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE Transactions on Knowledge and Data Engineering
Pattern Recognition Letters
Applied Soft Computing
Data and Knowledge Engineering
Journal of Mathematical Imaging and Vision
IEEE Computational Intelligence Magazine
IEEE Intelligent Systems
Decision Support Systems
Expert Systems

References:

K. Bowyer, K. Chang and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition", *Computer Vision and Image Understanding*, vol. 101(1), pp. 1-15, 2006.

P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer and R. Wirth, "CRISP-DM 1.0 Step-by-step data mining guide", <http://www.crisp-dm.org>, 2000.

J. Fenn, "Emerging Technology Hype Cycle 2010: What's Hot and What's Not", <http://www.gartner.com/>, 2010.

R. Jaeger and T. Halliday, "On confirmatory versus exploratory research", *Herpetologica*, pp. 64-66, 1998.

A. K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, pp. 4-20, 2004.

K. Lee, J. Ho, M. Yang, D. Kriegman, "Video-Based Face Recognition Using Probabilistic Appearance Manifolds," CVPR, vol. 1, pp.313, 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '03) - Volume 1, 2003.

J. Liu, L. Zhong, J. Wickramasuriya and V. Vasudevan, "uWave: Accelerometer-based personalized gesture recognition and its applications", Pervasive and Mobile Computing, vol. 5 (6), pp. 657-675, 2009.

P. Turaga, R. Chellappa, V. Subrahmanian and O. Udrea, "Machine Recognition of Human Activities: A Survey," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 18, pp. 1473-1488, 2008.

V. Vaishnavi and W. Kuechler, "Design research in information systems", retrieved October 2, 2004, from <http://www.isworld.org/Researchdesign/drisISworld.htm>, 2004.

D. Weinland, R. Ronfard and E. Boyer, "A survey of vision-based methods for action representation, segmentation and recognition," Computer Vision and Image Understanding, vol. 115, pp. 224-241, 2011.

C. Yam, M. Nixon and J. Carter, "Automated person recognition by walking and running via model-based approaches", Pattern Recognition, vol. 37(5), pp. 1057-1072.

W. Zhao, R. Chellappa, P. Phillips and A. Rosenfeld. "Face recognition: A literature survey." ACM Computing Surveys, vol. 35, n° 4, pp 399-458, 2003.

Supervisors:



Paulo Cortez (supervisor, Associate Professor, Centro Algoritmi, Department of Information Systems, University of Minho)



Henrique Santos (co-supervisor, Associate Professor, Centro Algoritmi, Department of Information Systems, University of Minho)

Phd Student:



Palwasha Afsar (MAP-I student)

R&D Unit:

Centro Algoritmi, Univ. Minho, Portugal.