



Master's Thesis Proposal

General Information

Master's Thesis Title: **Human layout analysis.**
Publication Date: 14/11/2010
Expiry Date: 14/11/2011
Modality: technological project
 research work
M.Sc. Th. Advisor: Jordi Vitrià
M.Sc. Th. Advisor's
Dept. & University: MAIA, Universitat de Barcelona
M.Sc. Th. Advisor e-mail: jordi.vitria@ub.edu
Observations:
Student's Name:
(if already known)

M.Sc. Thesis Description

Brief Description: (a 4-5 lines long paragraph)

This work will develop several software modules to detect and analyze humans in unconstrained static images. More specifically, the objective is to predict the bounding box, label and presence of each part of a person (head, hands, feet). This work will be included in a larger international project: the PASCAL VOC Challenge. (<http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2010/index.html>).

Detailed Description: (a half-page or one-page description, may include a brief task planning)

The Pascal Visual Object Classes (VOC) challenge is a benchmark in visual object category recognition and detection, providing the vision and machine learning communities with a standard dataset of images and annotation, and standard evaluation procedures. Organised annually from 2005 to present, the challenge and its associated dataset has become accepted as the benchmark for object detection.

The most challenging task is the human layout analysis, which is focused on predicting the bounding box and label of each part of a person (head, hands, feet).



For the person layout competition, a subset of “person” objects in each of the main datasets was annotated with information about the 2-D pose or “layout” of the person. For each person, three types of “part” were annotated with bounding boxes: the head, hands, and feet. These parts were chosen to give a good approximation of the overall pose of a person, and because they can be annotated with relative speed and accuracy compared to e.g. annotation of a “skeleton” structure where uncertainty in the position of the limbs and joints is hard to avoid. Annotators selected images to annotate which were of sufficient size such that there was no uncertainty in the position of the parts, and where the head and at least one other part were

visible—no other criteria were used to “filter” suitable images. In total, the training/validation set contained 439 annotated people in 322 images, and the test set 441 annotated people in 441 images.

This thesis will be focused on the development of fast, robust, algorithms for detecting hands and feet in non-restricted images. Depending on the results, it will be presented to the international challenge.

The main tasks to be developed are: (i) A review of the state of the art; (ii) The selection of a feature extraction method to represent the appearance of legs and arms; (iii) The development of a probabilistic model for recognition; (iv) the evaluation of the methodology.

References: (bibliographical references, relevant web links, etc.)

1. Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J. and Zisserman, A. The PASCAL Visual Object Classes (VOC) Challenge, *International Journal of Computer Vision*, 88(2), 303-338, 2010.
2. *Karlinsky Leonid, Dinerstein Michael, Daniel Harari, and Ullman Shimon*. The chains model for detecting parts by their context, *CVPR 2010*.

Minimal Requirements & Previous Knowledge: (optional)

The research profile of the candidates should cover as much as possible:

- Programming experience in C or C++ and Matlab is compulsory.
- Written and verbal communication skills in English.

Additional comments: (optional)

Location and Date: Barcelona, 9/11/2010

To the Academic Commission of the Master in Artificial Intelligence (CAIMIA)