Application for participation at the Doctoral Forum

Last name: Perez-Molina

First name: Eduardo

Institutional affiliation (university/department/link to website): CTB-Universidad Politecnica de Madrid (http://www.ctb.upm.es/)

Full address of the doctoral student including phone and fax numbers and email address:
Address: Groenewegje 166B, 2515Nc Den Haag, The Netherlands
Phone: +31.70.3894415
email: eperezmolina@icloud.com

Names of the supervisor(s) (links to their websites): Francisco Del Pozo
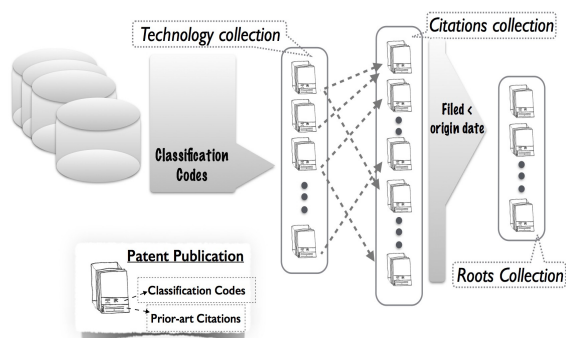(francisco.delpozo@ctb.upm.es)

(in total max 1500 words)
Description of doctoral research project (including research questions, theoretical background, planned methodology, current status)

The ideas and the Basics.- The purpose of my work is to create a new visual tool for characterizing a specific technology by its technological foundations. This tool, the TechnologyRoot-spectrum (TR-spectrum) visualizes the influence of the prior existing technologies on the specific technology under study, quantifying the connexions and distances to other technologies.

My thesis relies on the hypothesis that the documents cited in the prior-art search reports by patent examiners against patent applications in a particular – new – technology link the new emerging techniques to the conventional existing ones. The nature of this particular set of references, namely who produced the citations – the patent examiner in place of the author – and why they are cited – the evaluation of the novelty and nonobviousness -, is unique within the body of bibliographic references (Meyer, 2000) and explicitly points to temporal and conceptual proximity. These two factors seem fundamental in the study of history and technology. This is why, in our opinion, prior-art patent citations could play an unique and effective role as historical footprint.

My tool will use the computer to exploit the network formed by prior-art citations in patent publications and the classification codes assigned to them.
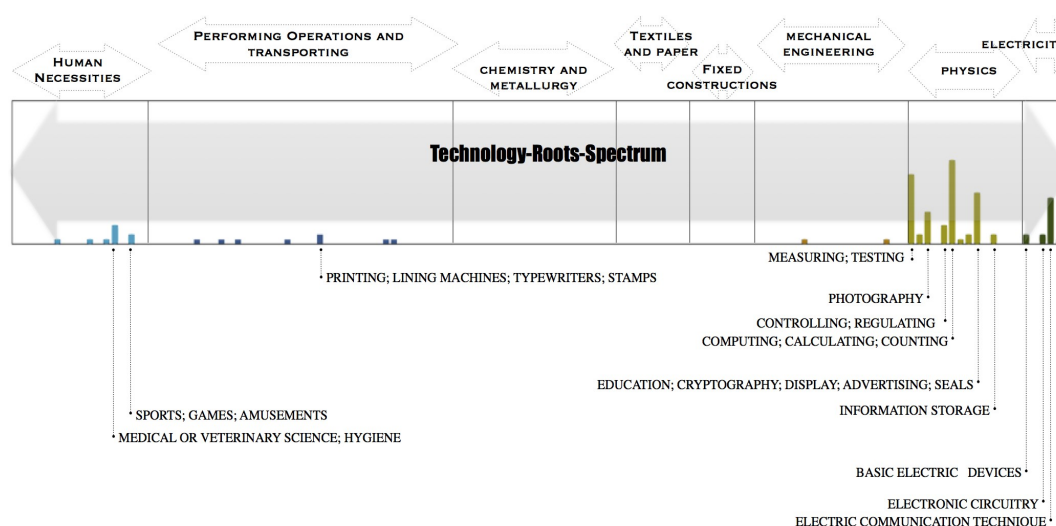
Data Collection Methodology.**-** For the characterisation of a specific technology the dataset to process is collected by the following path (see figure 1): First, we select, using the _espacenet_ public database (http://worldwide.espacenet.com/?locale=en_EP) the whole collection of patents published in the specific technology using classification codes forming the "specific technology" collection. From this set we extract all the citations from their search reports building the "citations" collection. Then, we keep the patents filed before the date of emergence of the specific technology - the origins date -, obtaining the "roots" collection.

Technology collection · Citations collection · Classification Codes · Patent Publication · Classification Codes · Prior-art Citations · Filed < origin date · Roots Collection

The TR-spectrum.- The set of selected patents - the "Roots" collection - is formed by patent publications disclosing technology methods, concepts, devices or systems deeply related with different aspects of the specific technology under study and filed before its origin date, therefore the "Roots" collection is composed only by patents being invented necessarily before the specific technology existed. Analysing in turn the codes assigned to the "Roots" collection publications will provide us indications of the technological foundations of the technology under study. This is why we use the expression: TechnologyRoots.

Moreover, every patent publication in the "Roots collection" is classified in a code representing a technology chosen between all possible existing technologies, this is why we use the term: Spectrum.

In consequence, the TR-spectrum is built aggregating the classification codes allocated to each document within the "roots" collection, and ordering this dataset in a sequence in accordance with the IPC scheme at a certain level of granularity –section, class, sub-class, group or sub-group – (WIPO, 2014). Changing the level of granularity we zoom- in or zoom-out in the techniques to have different conceptual resolution and in consequence we can identify more technical details or we can have global views of technical fields.

Research questions:
1) The TR-spectrum as a tool for the history of technology.
The direct analysis of the main components of the spectrum provides us with a visual and numerical indication about the technological foundations of a specific technology.
2) The TR-spectrum as a tool for technology metrics.
I have defined a TR-space formed by every spectral bin, namely the IPC codes at a certain level of granularity. Computing the TR-spectrum for different technologies we can assign a position in the space for each technology. I have applied a multi-dimensional scaling method (Wickelmaier, 2000) to this space for computing and visualising the distances between the origins of technologies.

Current Status.- Until now I have tested the TR-spectrum for four technologies - Computer Graphics, Graphical User Interfaces, Computerised Tomography and Airbags -, I have studied the results in deepness for Computer graphics (Perez-Molina, 2014) and I have applied multi-dimensional scaling to these four technologies using a euclidean distance. I will submit a poster to ISSI2015 with my results.

Motivation for student participation at the Doctoral Forum and the issues you wish to receive feedback on from the senior researchers.

My aim for this application is to get comments from the scientometrics community about the interest of the characterization of a specific technology from its origins, as well as to get a critical view about my tool.

More concrete questions are:

- What is the best definition of a metric between groups of documents in a TR-space - a classification space-? Distance computation in a classification space?
- There are other works or projects processing the network of patent citations and classification codes for characterizing a technology?

References

Meyer, M. (2000). What is special about patent citations? Differences between patent and scientific citations. *Scientometrics*, 49 (1). pp. 93-123.

WIPO (2014). International Patent Classification - Guide. Retrieved February 05 2015 from: http://www.wipo.int/export/sites/www/classificat ions/ipc/en/guide/guide_ipc.pdf.

Perez-Molina, E. (2014). The Technological Roots of Computer Graphics. *IEEE Annals of the History of Computing*, Vol.36,Issue 3, 30-41.

Wickelmaier, F. (2003). An Introduction to MDS. *Sound Quality Research Unit*, Aalborg University, Denmark. Retrieved on November 25, 2014 at http://steep.inrialpes.fr/~Arnaud/indexation/mds03.pdf