Towards an elementary percept

Written by Ségolène Tarte
Monday, 22 September 2008 16:28 - Last Updated Friday, 25 September 2009 21:38

[Meeting with Mike; 22nd September 2008]

 We know from various perceptive psychology studies that experts mobilise a very large and varied amount of knowledge when transcribing a text (see Melissa's book). Most literature on the subject is plausible but, according to Mike, probably wrong. The corpus of knowledge is definitely vast -- and certainly too rich for us to be able to represent it all. So where can we start? [...]

What we want to achieve is the transition from signal (the image of the text) to meaning. So far, with signal processing techniques, we can (or will soon be able to !) offer help as to strokes detection. The next logical step is to identify putative characters. Plausibly, the transition from strokes to letters would mobilise a smaller amount of knowledge and thus come into reach. The aim here is to drift away from the black box approach where an MDL algorithm proposes transcriptions; we want something more transparent that would make the percepts and thus the interpretations explicit.

Mike proposes that an elementary percept could be a region of an image that contains a grapheme. Here is how. Imagine the text was originally printed, and associate to each letter a movable printing block. Each individual area on the document that corresponds to a printing block would be an elementary percept. The set of these areas form a tessellation of the image; and this tessellation isn't unique as it depends on how its cells (tiles) are defined. Each tessellation would then be the root of an interpretation tree. In a given tessellation, each cell would be expected to contain a letter (sometimes more than a letter). So now the task is to actually identify the contents of each cell. The knowledge we can draw on is: the woodgrain, the identified strokes and the corpus of letter shapes. Likely, we would now have some possible letters for each cell. A crosswords solver type of approach could then be taken to fill in the blanks!!... This tessellation-based approach would also allow alternative parsings, depending on which tessellation is used. The historian, from a certain point on, would then commit to a certain tessellation to develop an interpretation. The alternate tessellations (interpretation trees) are then made dormant, but can be resurrected at any time if needed. (Typically for a palimpsest text, there would be at least two very different tessellations, one for each text).

The whole crux here is about tessellations!!! In a first step, we would need to define them manually. Automating this task would be an image processing job (although certainly not a trivial one). A start would be to identify the written lines. Then the cells can be defined. Refining this approach, cells could even be defined as the central area of a grapheme, making ascender and descender areas into fuzzy cells (and draw on K. Rayner's research [1] on reading and eye

movements).

[1]   K. Rayner. "Eye movements in reading and information processing: 20 years of research". *Psycho logical Bulletin*
, 124(3):372-422, 1998.