# From Sketch to Reality: Generate HTML pages with photos of drafts

Xinqi Li, Peilin Hao, Yinan Zhang

We want to create a software that help web designer quickly generate html page from draft. But scanning each of their drafts on a paper, we recognize every blocks/shapes in the draft, as well as text in each block. Then the software will directly generate an html page. Users can also manually define each part in the picture, as well as editing each block recognized by the software.

Preview:
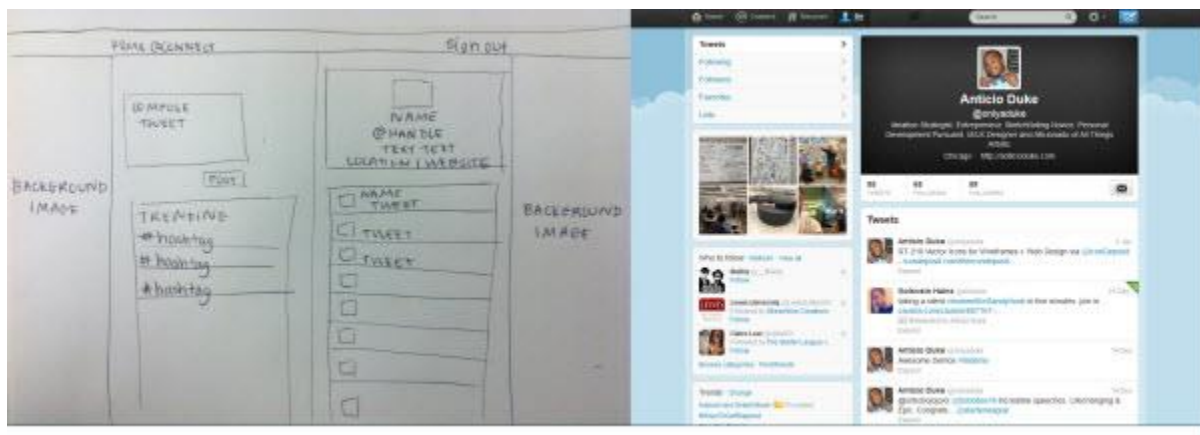
Approach:

1. Perspective Shift
   The picture users take might not be from an orthographic view. We need to change it to an overhead view of the draft. This can be done by performing a reversed projection.

2. Basic Shape recognition
   By performing edge detection and corner detection, we can recognize, at least, some basic polygons. These polygons are blocks in a web page. For some designer who wants to keep simplicity (well organized text only), he can specify virtual block in the image.

3. Image recognition

A rectangle with no text, yet not pure blank, in it is an image. Man can also define a block as an image.

4. Text Recognition
We want to use some open source OCR engine at first. For example, Tesseract started by Google.
If we have enough time, we will create our own OCR engine.

5. HTML generation engine
Given a set of rectangles or other shapes, and the text or image inside each block, HTML generation engine is responsible for generating an html page that represents exactly the same thing described in the draft.

Data Set:

We will collect about 20 simple web design sketches on line for testing purpose.

Milestone Goals:

- · Basic shapes can be recognized.
- · HTML generation engine half finished
- · OCR integrated.

Reference:

[1] PerspectiveImageCorrection: http://sourceforge.net/projects/perspectiveimg/

[2] Moon, Hankyu, Rama Chellappa, and Azriel Rosenfeld. "Optimal edge-based shape detection." *Image Processing, IEEE Transactions on* 11.11 (2002): 1209-1227.

[3] Plamondon, Réjean, and Sargur N. Srihari. "Online and off-line handwriting recognition: a comprehensive survey." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.1 (2000): 63-84.

[4] Arica, Nafiz, and Fatos T. Yarman-Vural. "An overview of character recognition focused on off-line handwriting." Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 31.2 (2001): 216-233.