**Question that needs to be answered:**

1/ The outer line of [the character](https://drive.google.com/drive/folders/0B3RpLOFbPBk-fldlQ3ZDSi1zeEtWRDlkSUxRTGJmQ2RjUXk2anVSRlRrZkdXSzhwbUd1YlE/0B3RpLOFbPBk-MkE1aHBNcTh1d2c/0B3RpLOFbPBk-fjZzc2FUS1V0b0pQZ1F0YnhTeGM1SVN0eXdKck44MmVuUVFaT0VnXzZ5YXc) is not smooth. I believe this is because we took these characters directly from the pdf of the papers. Is the outer line of the character supposed to not be smooth, or should we generate new character with smooth outer line while maintaining the ratio of the different strokes in the characters ? Is there anyway to obtain the original dataset (characters) that the author of the paper used ?

2/ GT email 12/14/14: Hi Noris and Quan, The size if the PNG files that Noris originally computed are not correlated at all with human complexity judgments. However, when they are converted to PDF files there is good correlation between the size of the PDF files and human complexity. I just saved the PNG files as JPEG files with best quality selection, and the correlation with human complexity is 0.82. Incredible! We should find out how JPEG works since it looks like it may incorporate symmetries in order to do so well, and much better than Papentin compression.

3/ Ask Professor GT for the paper containing the Chinese characters.

**To Do:**

1/ Recheck code to remove any hard code written for analysis of 6\*6 patterns

2/ Go through code to ensure there is no bug

3/ Upload code onto github for ease of maintenance and sharing

4/ Calculate the scores of each image for the [categories](https://drive.google.com/drive/folders/0B3RpLOFbPBk-fldlQ3ZDSi1zeEtWRDlkSUxRTGJmQ2RjUXk2anVSRlRrZkdXSzhwbUd1YlE) that were used on the 6\*6 patterns

5/ Retrieve human’s scores for each image in each category from the paper

6/ Compare the calculated scores of each image in each category against human’s scores and output into an excel file

7/ Run [Spearman Rank Order Correlation](http://vassarstats.net/corr_rank.html) for each category to compare calculated scores and human scores.