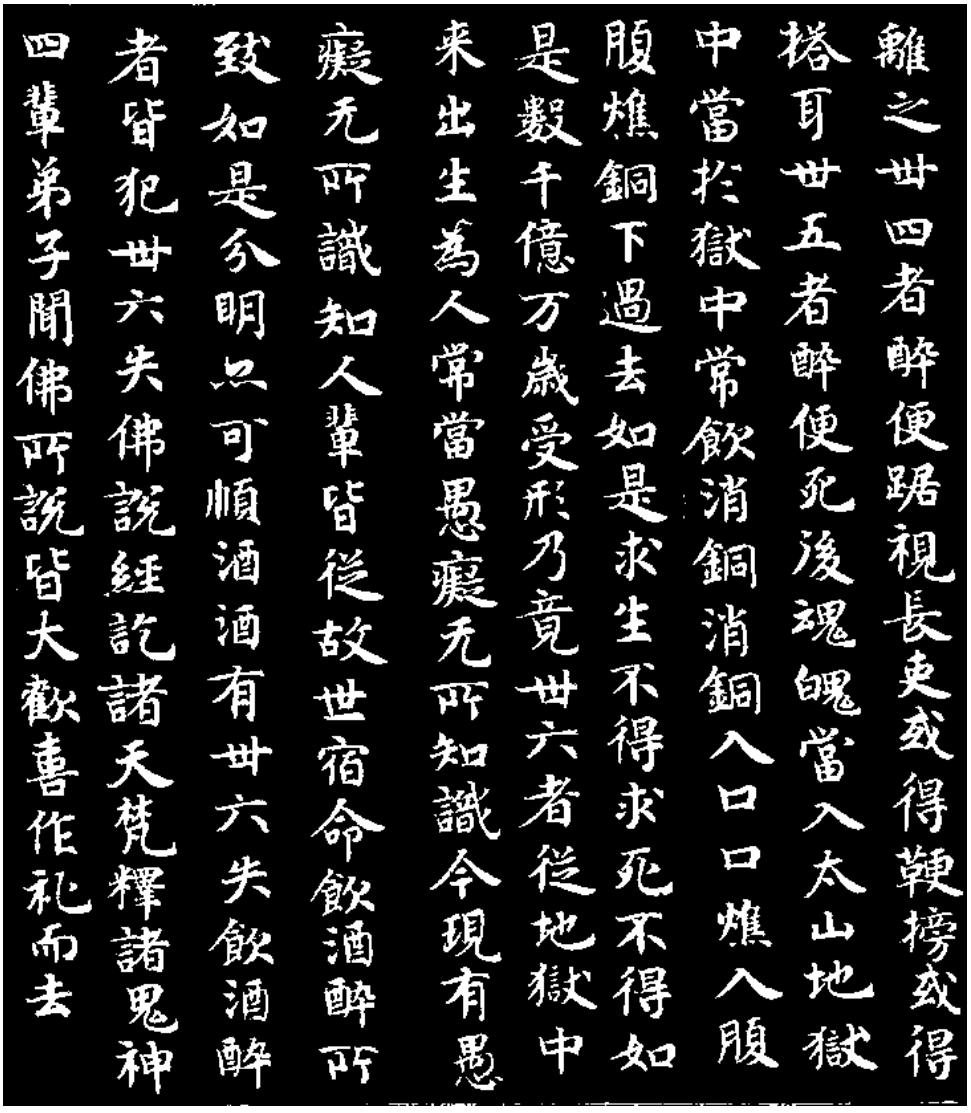
Image Processing 1 - Exercise 7 - WiSe 2012/13

Weipeng He
2he@informatik.uni-hamburg.de
6411529

January 21, 2013

1 Preprocessing

Read the color image, convert to grayscale, normalize to $[0.0, 1.0]$. Then use threshold (0.68) to make the image black and white, in order to remove the noise. The preprocessed image is shown below :



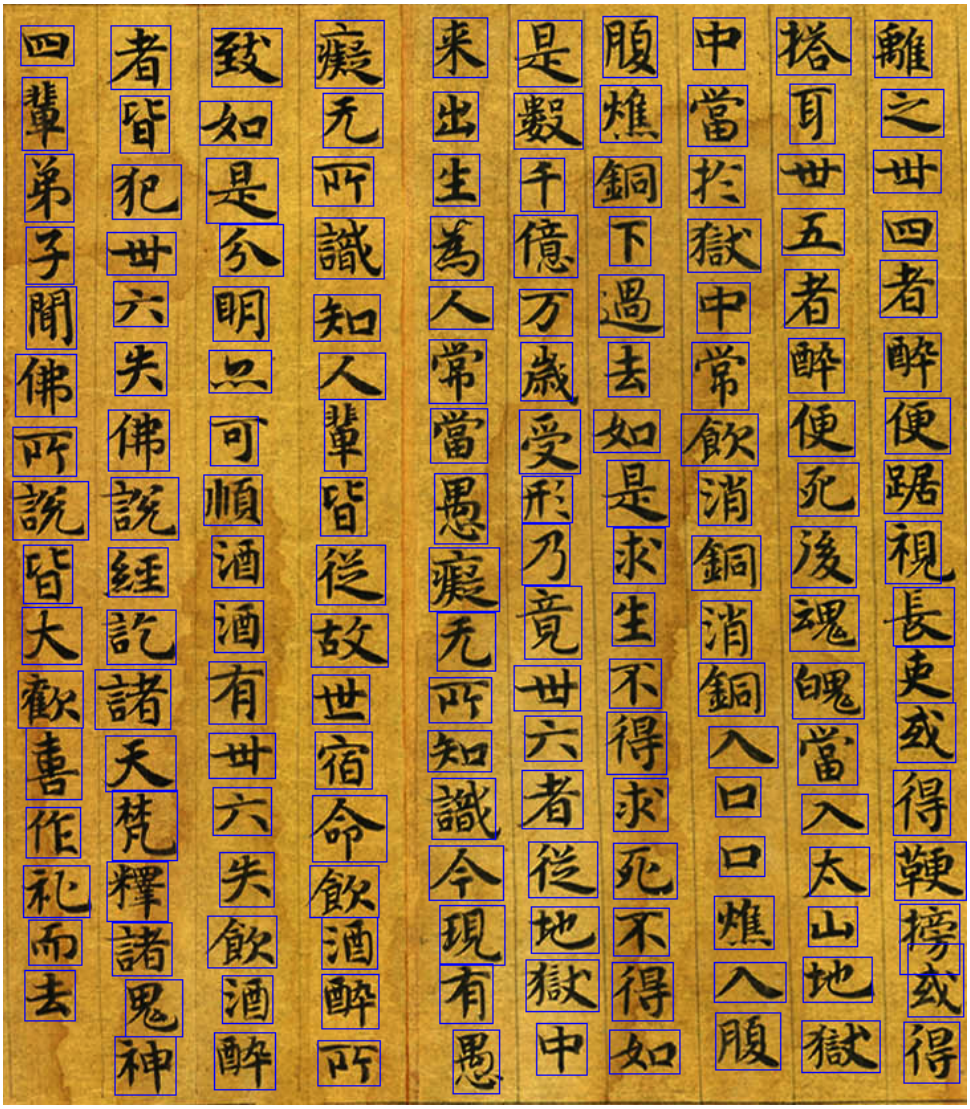
2 Finding candidates characters

Filter the image using a convolution mask shown below:

$$\begin{pmatrix}
-2 & -2 & -2 & -2 & \cdots & -2 & \cdots & -2 & \cdots & -2 \\
-2 & -1 & -1 & -1 & \cdots & -1 & \cdots & -1 & \cdots & -2 \\
-2 & -1 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & -2 \\
-2 & -1 & 0 & 1 & \cdots & 1 & \cdots & 1 & \cdots & -2 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & -2 \\
-2 & -1 & 0 & 1 & \cdots & 5 & \cdots & 5 & \cdots & -2 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & -2 \\
-2 & -1 & 0 & 1 & \cdots & 5 & \cdots & 5 & \cdots & -2 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & -2 \\
-2 & -2 & -2 & -2 & \cdots & -2 & -2 & -2 & \cdots & -2
\end{pmatrix}$$

The size of the mask is about the same with an average character (51×51), the value of the center is high (maximum of 5) and the margin is negative. The mask can detect where are likely to be a character.

Afterwards, use maximum filter to find out local maximum which are not less than all neighbor pixels(distance less than 30). From these points, I find the best fit box which make the density of 3-pixel margin to be low (0.05). Then, the candidate characters are found. Results are shown below:



離之世四者醉便踞視長吏或得鞭撲或得
搭耳世五者醉便死後魂魄當入太山地獄
中當於獄中常飲消銅消銅入口口燋入腹
腹燋銅下過去如是求生不得求死不得如
是數千億万歲受形乃竟世六者從地獄中
来出生為人常當愚癡无所知識今現有愚
癡无所識知人輩皆從故世宿命飲酒醉所
致如是分明亦可順酒酒有世六失飲酒醉
者皆犯世六失佛說經記諸天梵釋諸鬼神
四輩弟子聞佛所說皆大歡喜作礼而去

3 Finding Points of Interest

I calculate the points of interest using standard SIFT algorithm, which is caculating the DoG pyramid.

4 Generating SIFT-Features

I calculate the feature descriptor using standard SIFT algorithm.

5 Retrieval

The retrieval of matching character is done by comparing the template and the candidate character (subpicture of the original) one by one. A quality score is given for classifying if these two charater match.

First, I use brute force matching to find out nearest neighbor of each key points in the template. Then, calculate the score of each match by :

$$Score = \frac{1.0}{.1 + Distance^2}$$

where the $Distance$ is the matching distance. Divide scores by the maximum score, so that the maximum score is 1.0. Filter out the matches with score less than $0.4 \times MaxScore$.

For each match, I can calulate the shifts in x and y direction (denote by array $ShiftX$ and $ShiftY$). The quality score is the weighted standard deviation (using score as weight) of the shifts.

$$Quality = \sqrt{\frac{\sum_{i=1}^{n} Score_i((ShiftX_i - \bar{ShiftX})^2 + (ShiftY_i - \bar{ShiftX})^2)}{\sum_{i=1}^{n} Score_i - 1.0}}$$

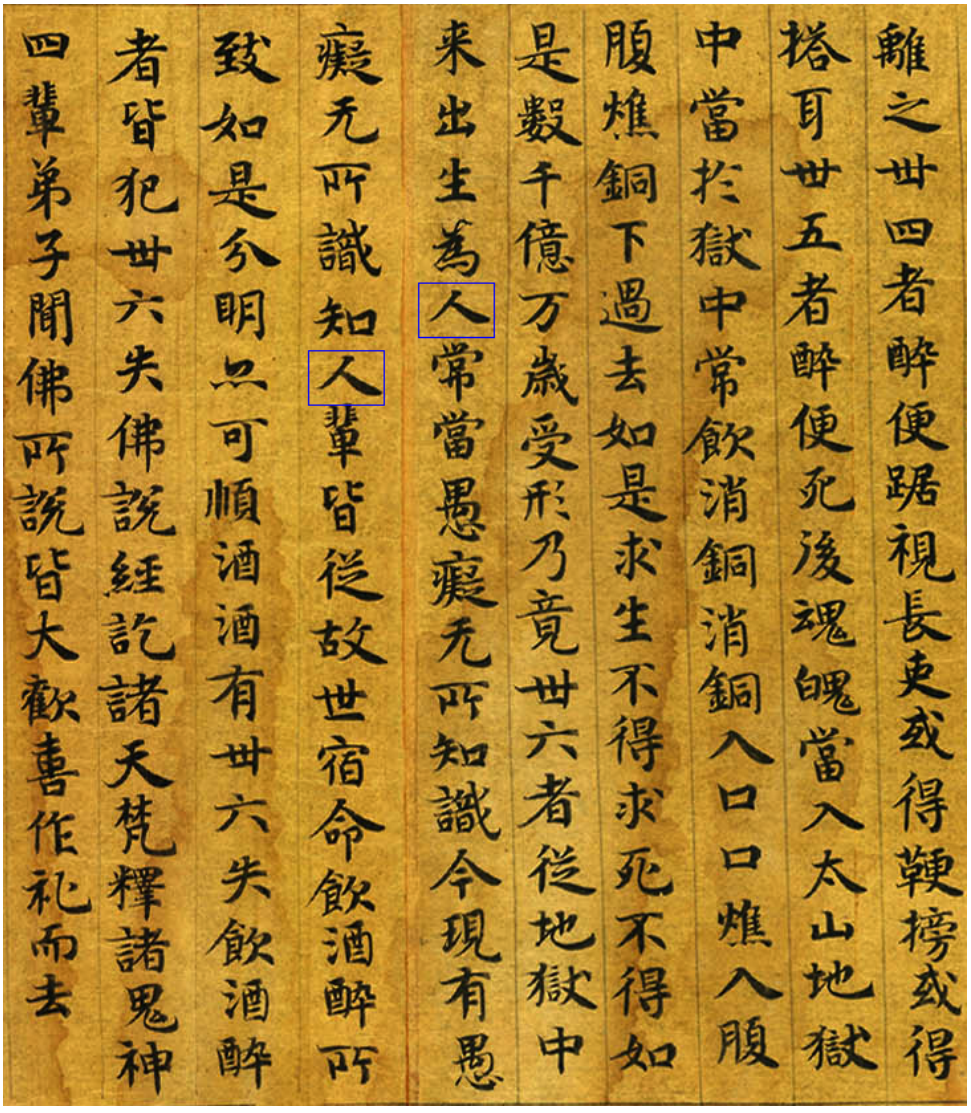
Those pair of candidate and template with quality score less than 2.0 is considered to be matched. And, the mean of shifts in x and y can be used to calculate the match position.

6 Results

a

Template:

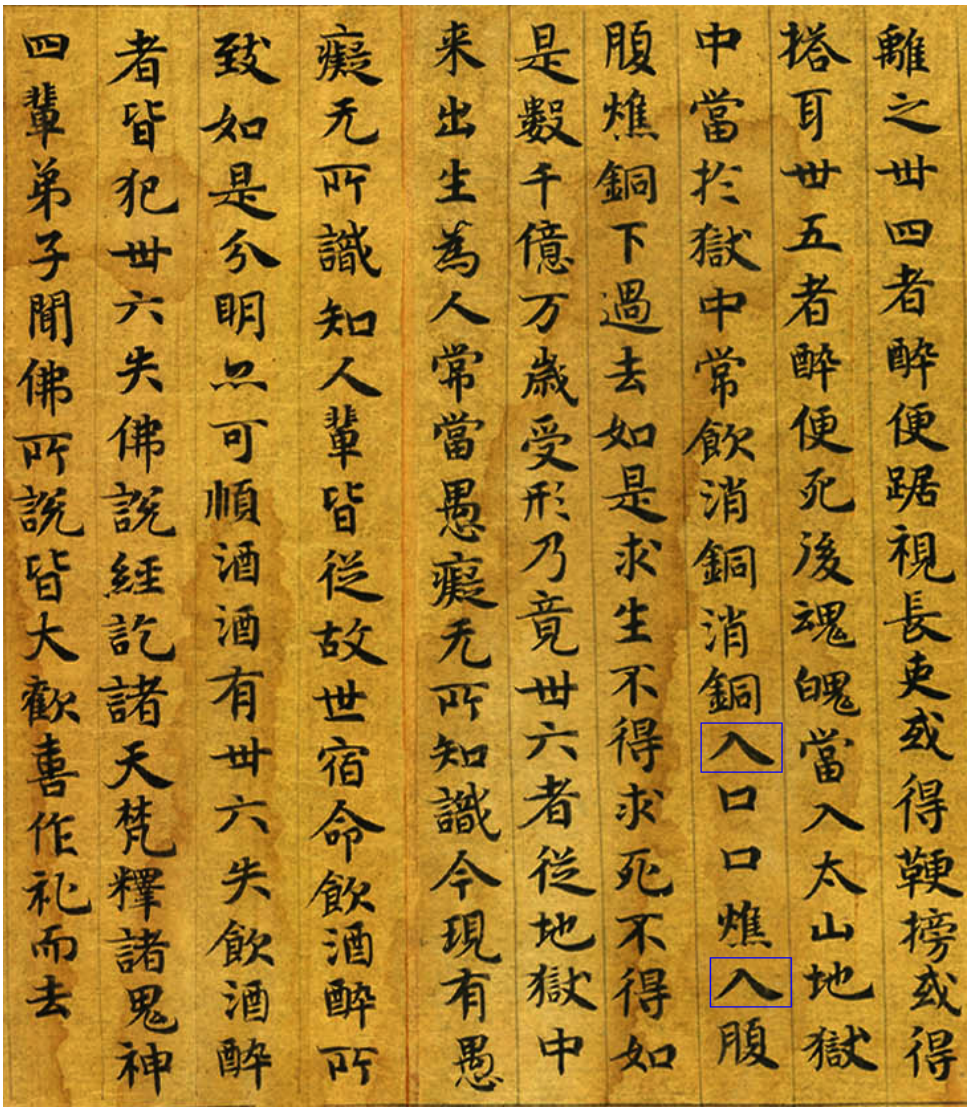
Result:



b

Template:

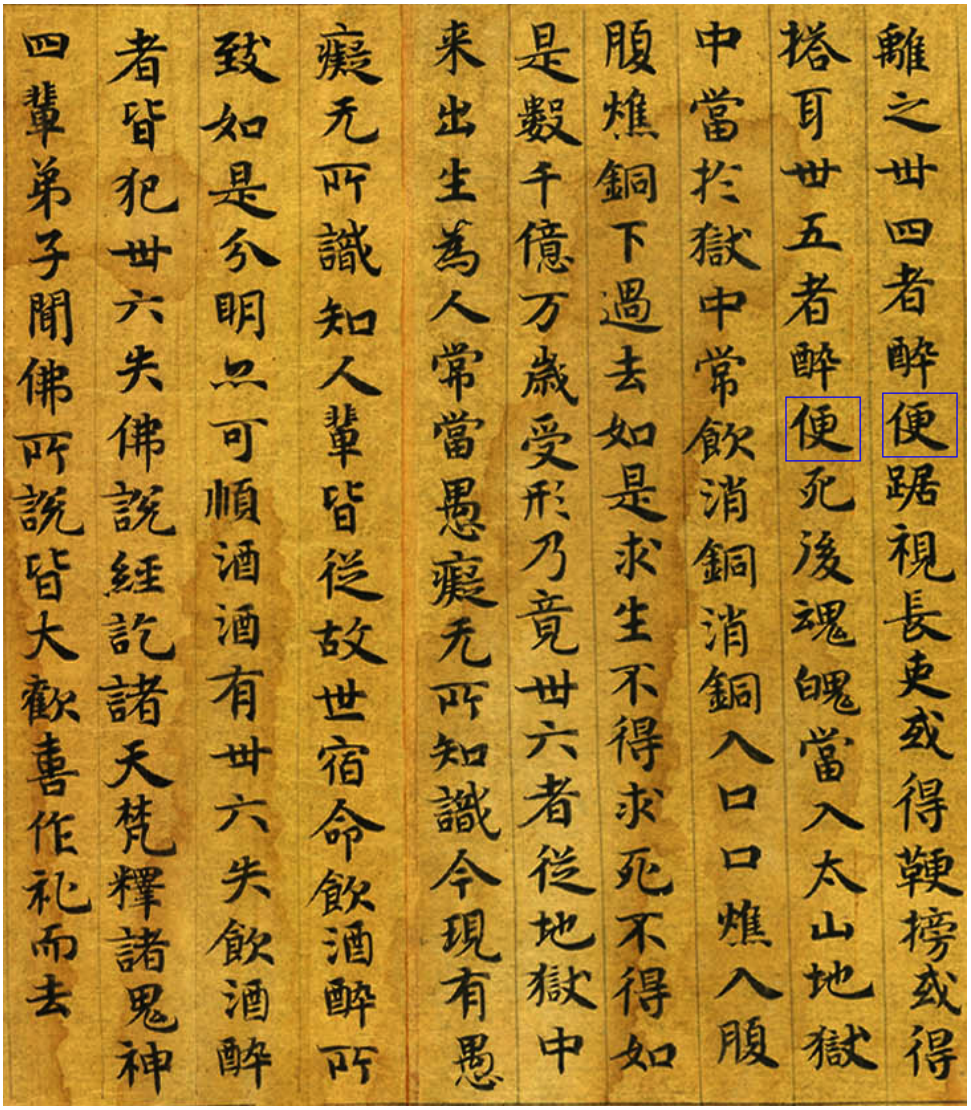
Result:



c

Template:

Result:



d

Template:

Result: