Interactive Video Lecture Summary

Hijung V.Shin

January 4, 2015

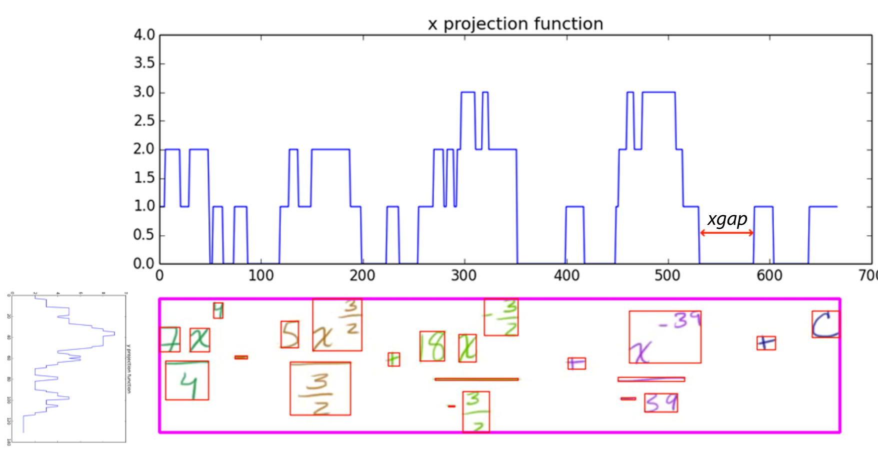
1 Segmentation

1.1 Line break algorithm

1.2 Energy function

In summary, the score for a set of lines is the sum of the following terms (multiplied by constant coefficients and raised to constant powers):

- sum of the number of horizontally aligned strokes in each line
- sum of number of strokes in each line
- weighted average of the compactness of each line
- negative of the maximum horizontal gap in each line (xgap)
- negative of the maximum vertical gap in each line
- negative of the overlapping area between lines



More formally, consider C_s the set of all strokes drawn during a lecture. Our goal is to segment the strokes into mutually exclusive lines or figures, $C_{s_1}, C_{s_2}, ...$ such that

$$\bigcup_i C_i = C_s \tag{1}$$

$$C_i \cap C_j = \emptyset \quad \forall i \neq j \tag{2}$$

To this purpose we define a scoring function which, given a set of sets of strokes C_{s_i}'s returns a scalar value indicating the *badness* of the segmentation. The scoring function depends on several factors.

1. The number of strokes in each line:

$$n_{\text{strokes}}(C_{s_i}) = \left|C_{s_i}\right| \tag{3}$$

2. The number of horizontally aligned strokes in each line: In order to identify strokes that are aligned horizontally, we define a *horizontal projection function*, $proj_h$, as

$$proj_h(y, C_{s_i}) = \left|\{s \in C_{s_i} | y_{\min}(s) \leq y \leq y_{max}(s)\}\right| \tag{4}$$

where $y_{\min}(s)$ and $y_{\max}(s)$ return the minimum and maximum y coordinate of the bounding box for stroke s.
Then given a candidate line C_{s_i}, the maximum number of horizontally aligned strokes, $n_{\text{aligned}}(C_{s_i})$, is defined as

$$n_{\text{aligned}}(C_{s_i}) = \underset{y_{\min}(C_{s_i}) \leq y \leq y_{\max}(C_{s_i})}{\text{argmax}} f(y) \tag{5}$$

3. The maximum horizontal gap in a line. $x_{\text{gap}}(C_{s_i})$, is defined using the *vertical projection function*

$$proj_v(x, C_{s_i}) = \left|\{s \in C_{s_i} | x_{\min}(s) \leq x \leq x_{\max}(s)\}\right| \tag{6}$$

$$x_{\text{gap}}(C_{s_i}) = \underset{x_i, x_{i+1} \in proj_{v:C_{s_i} \neq 0}}{\text{argmax}} (x_{i+1} - x_i) \tag{7}$$

where $proj_{v:C_{s_i} \neq 0}$ is an ordered set such that $x_i \in proj_{v:C_{s_i}}$ if $proj_v(x_i, C_{s_i}) \neq 0$, and $x_i \leq x_{i+1}$.

4. Similarly, we can define the maximum vertical gap in a line, $y_{\text{gap}}(C_{s_i})$.
5. The compactness of a line, measured as:

$$comp(C_{s_i}) = \frac{\sum\limits_{s \in C_s} area(s)}{area(C_{s_i})} \tag{8}$$

   where $area(s)$ is the area of the bounding box of stroke s.

6. Finally, given a set of lines L, the overlap penalty, $P_{overlap}$, is defined as

$$P_{overlap}(L) = \sum_{C_{s_i}, C_{s_j} \in L} \frac{area(overlap(C_{s_i}, C_{s_j}))}{min(area(C_{s_i}), area(C_{s_j}))} \tag{9}$$

The score of a segmentation L is defined as:

$$
\begin{aligned}
score(L) = &\sum_{C_{s_i} \in L} \left(n_{\text{aligned}}(C_{s_i})^{\alpha_1} + n_{\text{strokes}}(C_{s_i})^{\alpha_2} \right) && (10)\\
&- \sum_{C_{s_i} \in L} \left(x_{\text{gap}}(C_{s_i})^{\alpha_3} + y_{\text{gap}}(C_{s_i})^{\alpha_4} \right) && (11)\\
&+ \frac{\sum\limits_{C_{s_i} \in L} (npix(C_{s_i}) \times comp(C_{s_i}))}{\sum\limits_{C_{s_i} \in L} npix(C_{s_i})} && (12)\\
&- P_{overlap}(L) && (13)
\end{aligned}
$$

References

[1]