IMAGE SEGMENTATION AND 3D MODELING TO BOOST TEXT RECOGNITION IN NATURAL SCENES

Shounak Gore

04/26/11

PROBLEM STATEMENT

- Text detection in natural scene images is an unsolved problem.

- The objective of the project is to segment images and understand the context to boost the detection of text in natural scenes.

MOTIVATION

- Text recognition in natural scenes is a question of open research.

- Robotic systems doing exact opposite thing are available.

ALGORITHM

- Segment the image.

- Select the area of interest, find the text location and apply text restoration if required.

- The text obtained in the above part is given to an OCR for text detection.
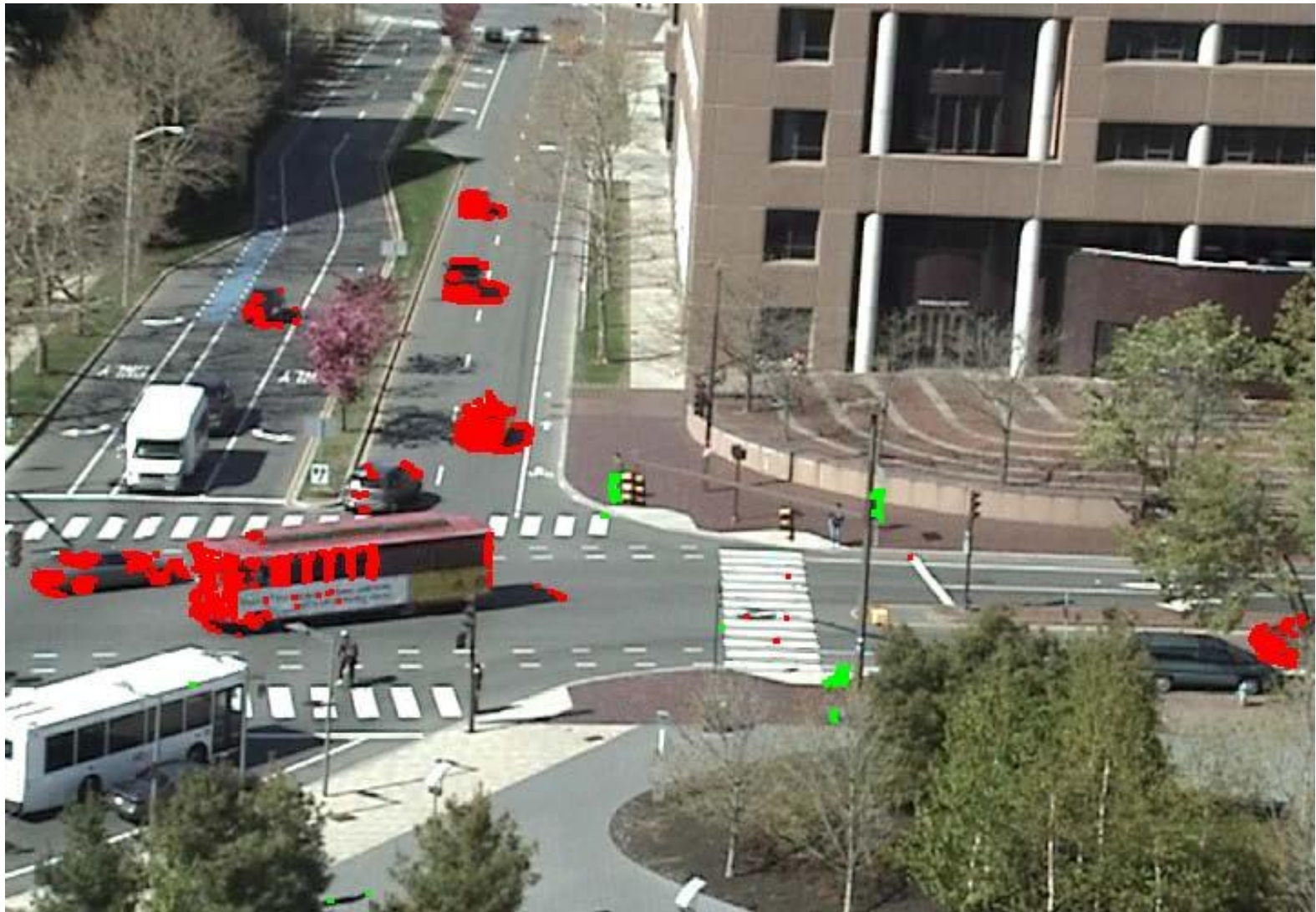
Image Segmentation

- *"In computer vision, **segmentation** refers to the process of partitioning a digital image into multiple segments."*

- *"Image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics."*

http://en.wikipedia.org/wiki/Segmentation_%28image_processing%29

http://people.csail.mit.edu/xgwang/HBM.html

http://people.csail.mit.edu/xgwang/HBM.html

Ways to Segment an Image

- Thresholding
- Clustering
- Compression-based
- Histogram-based
- Edge detection
- Region growing
- Split-and-merge
- Partial differential equation-based
- Level set methods
- Graph partitioning methods

FUZZY SHELL CLUSTERING

- This method combines both the edge detection and clustering approach.

- For an object like a building, simple edge detection is not sufficient. We need to cluster points together that form rectangles.
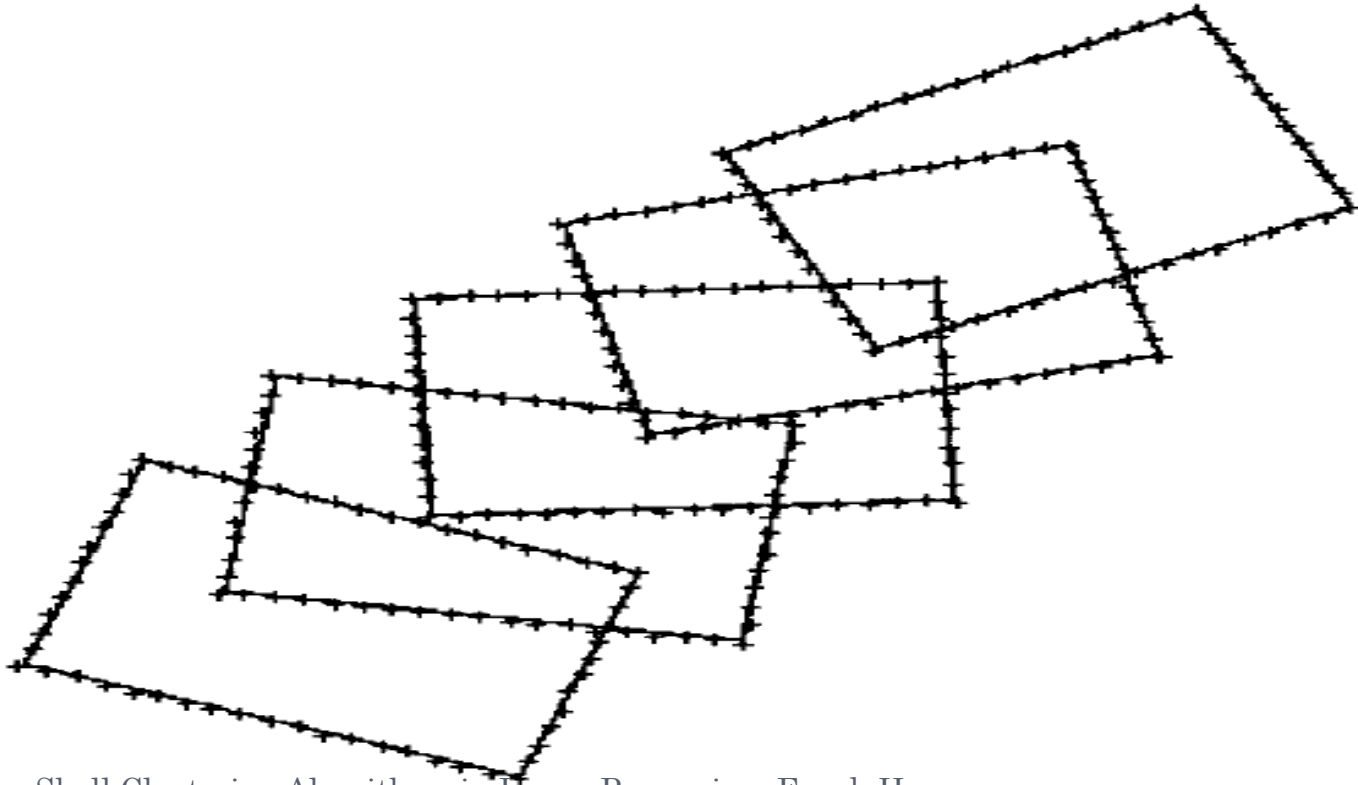
FUZZY SHELL CLUSTERING

- The algorithm in brief :

  - Select the number of clusters that you need.

  - For every point, using the selected distance measure (Euclidean or a close approximation is usually used), decide the extent to which it belongs to every cluster.

  - Look out for parallel edges (rectangles or shape of a building will have parallel edges).
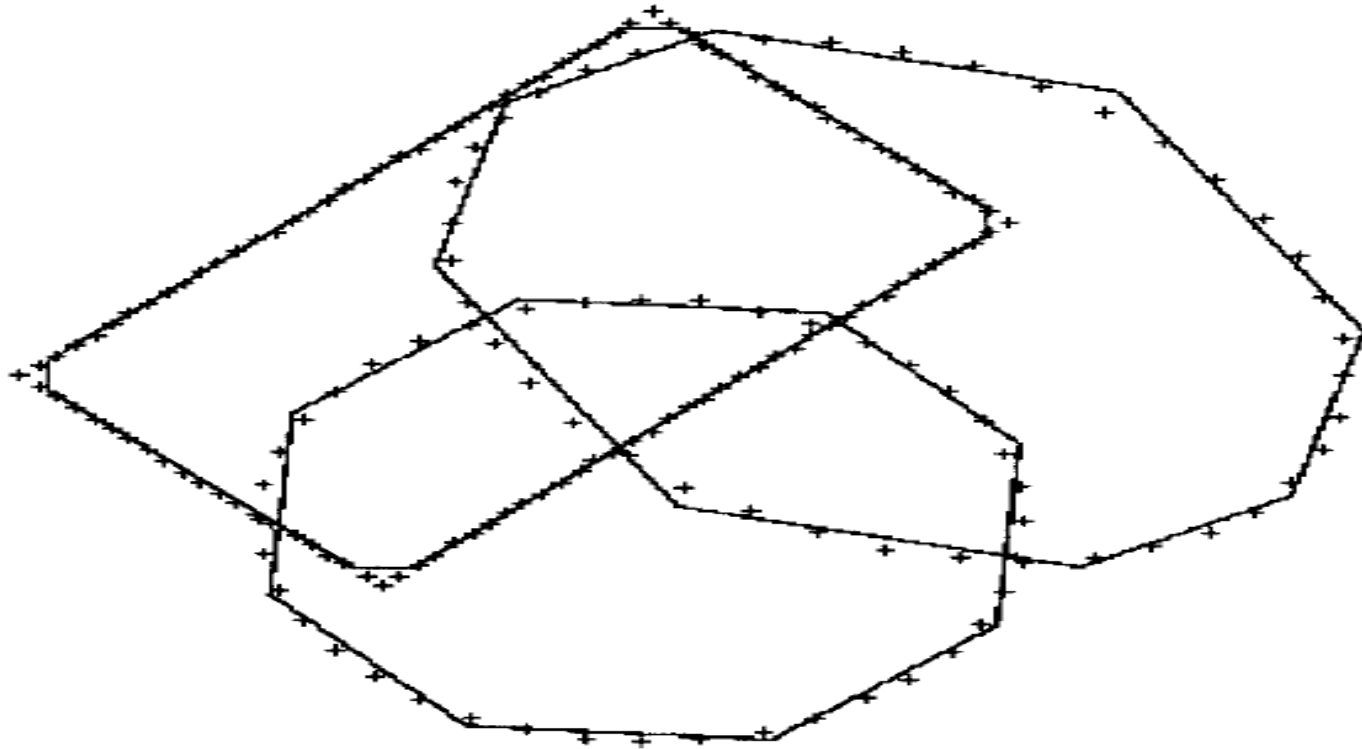
Fuzzy Shell Clustering

○ The fuzzy clustering has many advantages :

• It can easily detect shapes at various orientations

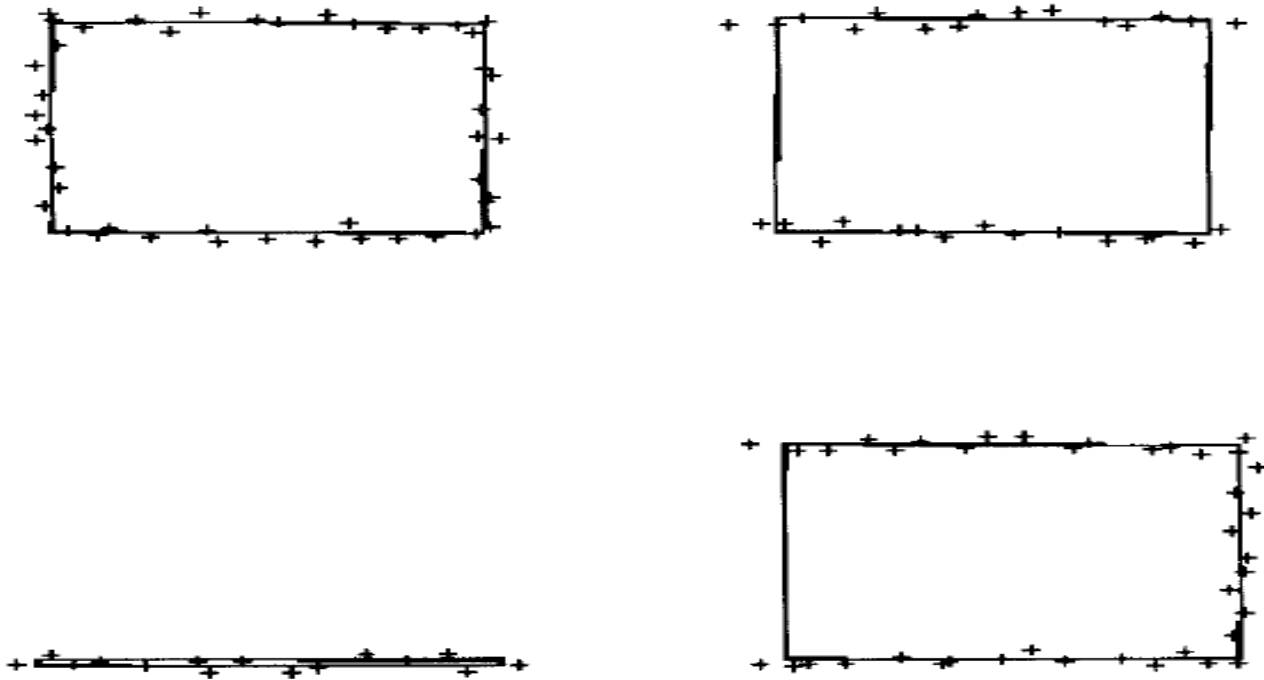Fuzzy Shell Clustering Algorithms in Image Processing, Frank Hoeppner

Fuzzy Shell Clustering

- The fuzzy clustering has many advantages :
  - It can easily detect various shapes

Fuzzy Shell Clustering Algorithms in Image Processing, Frank Hoeppner

FUZZY SHELL CLUSTERING

- The fuzzy clustering has many advantages :
  - It can compensate for missing edges

TEXT LOCATION DETECTION

- The assumption is that either the text or its background is of a uniform color.

- The text is located by grouping text pixels based on their intensity and region layout analysis.
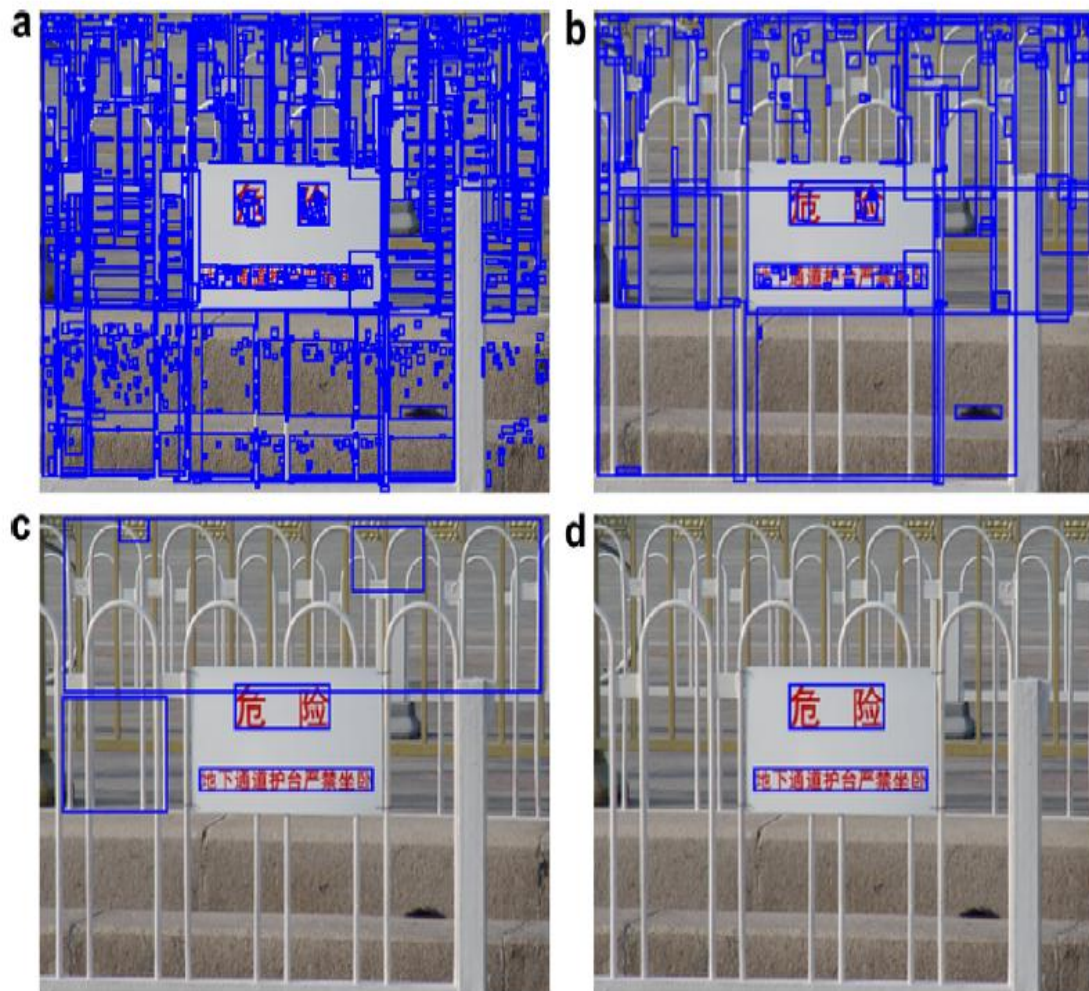
Fig. 4. Illustration of text detection process. (a) Outline rectangles of segmentation regions. (b) Spatial layout analysis result of color connected regions. (c) The classification result. (d) Final result after OCR feedback.

This and previous image from : Text detection and restoration in natural scene images, Qixiang Ye , Jianbin Jiao, Jun Huang, Hua Yu

TEXT RESTORATION

- Once the location has been determined, we need to restore the orientation of the text.

- This usually needs the camera parameters are necessary.

- But as these parameters are not available, perspective projection matrix which relates the image coordinate to the world coordinates is used.

TEXT RESTORATION

- We use the following equation for the purpose :

$$\begin{bmatrix} x' \\ y' \\ h \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{h} \begin{bmatrix} x' \\ y' \end{bmatrix},$$

Text detection and restoration in natural scene images, Qixiang Ye , Jianbin Jiao, Jun Huang, Hua Yu

TEXT RESTORATION

- But as we have a 2D view after the transformation Z =1 and the equation transforms to :

$$\begin{bmatrix} X^1 & Y^1 & 1 & 0 & 0 & 0 & -X^1x^1 & -Y^1x^1 \\ 0 & 0 & 0 & X^1 & Y^1 & 1 & -X^1x^1 & -Y^1x^1 \\ X^2 & Y^2 & 1 & 0 & 0 & 0 & -X^2x^2 & -Y^2x^2 \\ 0 & 0 & 0 & X^2 & Y^2 & 1 & -X^2x^2 & -Y^2x^2 \\ X^3 & Y^3 & 1 & 0 & 0 & 0 & -X^3x^3 & -Y^3x^3 \\ 0 & 0 & 0 & X^3 & Y^3 & 1 & -X^3x^3 & -Y^3x^3 \\ X^4 & Y^4 & 1 & 0 & 0 & 0 & -X^4x^4 & -Y^4x^4 \\ 0 & 0 & 0 & X^4 & Y^4 & 1 & -X^4x^4 & -Y^4x^4 \end{bmatrix} \begin{bmatrix} H_{11} \\ H_{12} \\ H_{14} \\ H_{21} \\ H_{22} \\ H_{24} \\ H_{31} \\ H_{32} \end{bmatrix} = \begin{bmatrix} x^1 \\ y^1 \\ x^2 \\ y^2 \\ x^3 \\ y^3 \\ x^4 \\ y^4 \end{bmatrix}.$$

Text detection and restoration in natural scene images, Qixiang Ye , Jianbin Jiao, Jun Huang, Hua Yu

TEXT OCR

- The text after the preprocessing is then given to a text OCR.

- For this project an HMM based OCR will be used.

- The HTK toolkit was used for the generation of the OCR.

Minimum Goals

- The text localization, restoration and text OCR are already completed.

- Get the Fuzzy cluster code working and make it compatible for all possible cases that can occur in natural scenes.

Testing

- The metric for testing is the result of the OCR.

- ICDAR 2003 dataset is used for training and testing the proposed algorithm.

- The segmented, localized and restored text will be tested for its accuracy as compared to a text obtained with just the localization segment of the algorithm.

ADDITIONAL GOALS

- Try to segment a varied variety of images.

- Find more time efficient algorithms.

- Train the OCR based on the contextual information and test the accuracy.

REFERENCES

- http://en.wikipedia.org/wiki/Segmentation_image_processing

- http://www.ee.surrey.ac.uk/ccsr/research/ilab/hci/segmentation

- http://people.csail.mit.edu/xgwang/HBM.html

- Fuzzy Shell Clustering Algorithms in Image Processing: Fuzzy C-Rectangular and 2-Rectangular Shells, Frank Hoeppner

- Text detection and restoration in natural scene images, Qixiang Ye , Jianbin Jiao, Jun Huang, Hua Yu

- Thank You