Interactive Adaptive Text Localization and Distinction

Samuel James Pinson
Dr. William Barrett
Computer Science Department
Brigham Young University
Text Localization and Distinction

- **Text localization**
  - Automatic data extraction
  - Automatic image indexing

- **Text distinction**
  - Improved automatic data extraction
  - Automatic image indexing
  - Document segmentation
  - Decision making
Related Work

Text localization
- Bottom-up
  - Connected component analysis
  - Texture analysis
- Top-down
  - Run-length smoothing
  - Recursive projection profile cuts

Text distinction
- Connected component analysis
- Empirically determined heuristics
- Text distinction
Our Approach

- Threshold image
- Extract features
- Accumulate training examples
- Induce decision tree
- Contextual cleanup

Aspect Ratio
Mean
Variance
Skew
Kurtosis
Compactness
Orientation
Number of Holes
Contextual Cleanup

- Horizontally Adjacent
- Similar

You check "Legal Alien Not Allowed" on the group information form if there is a government agency requiring your last name, Social Security number and that you meet all the requirements except for the number. A State or local government

If you check "Other," you need to provide a benefit for which a Social Security number is required.

Providing race/ethnic information is voluntary, it helps us prepare statistically correct people. We do not reveal the information to the month, day and full (4 digit) year.
Future Work

- Analyze features
- Improve Contextual Cleanup
  - Context
  - Class prototypes
  - Classification confidence
  - Misclassification costs
Conclusion

- Adaptivity
- Interactivity
- Multiple Applications
Questions