

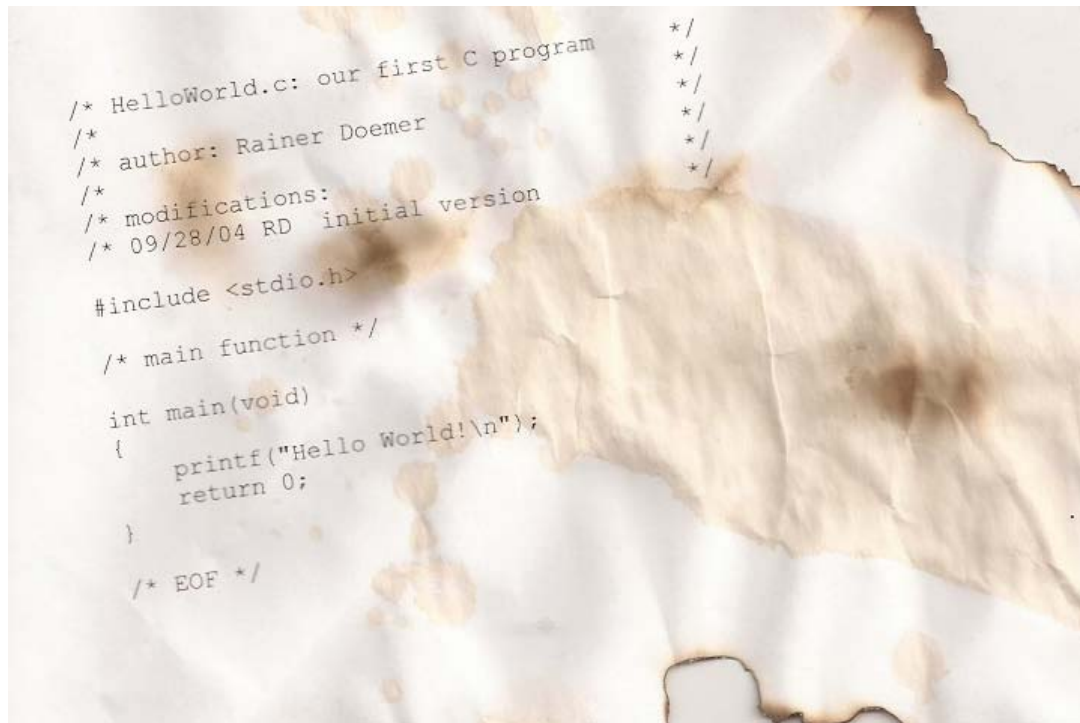


OCR **(OPTICAL CHARACTER RECOGNITION)**

Che-Wei Chang

OCR (OPTICAL CHARACTER RECOGNITION)

- This is the goal...



OCR (OPTICAL CHARACTER RECOGNITION)

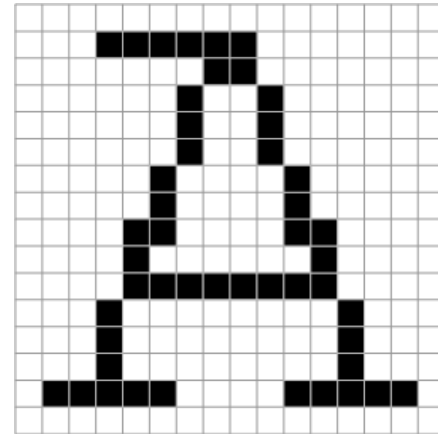
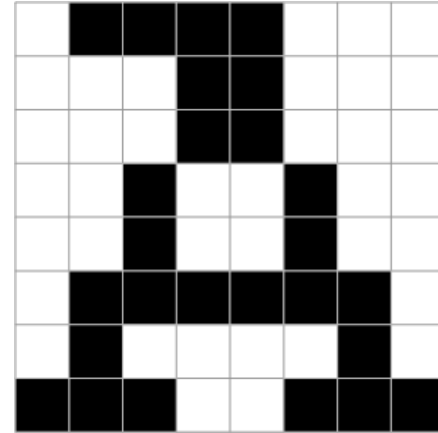
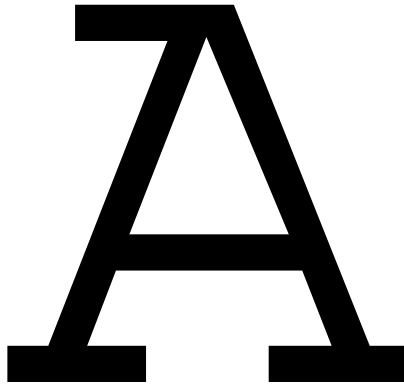
○ Desirable Features

- Multiple Formats Supported
- Multiple Fonts/Sizes Supported
- Image Rotation
- Stain Filtering
- Wrinkle Filtering
- GUI (highly recommended !!)
- Handwritten Document Supported
- Keyword Dictionary Supported



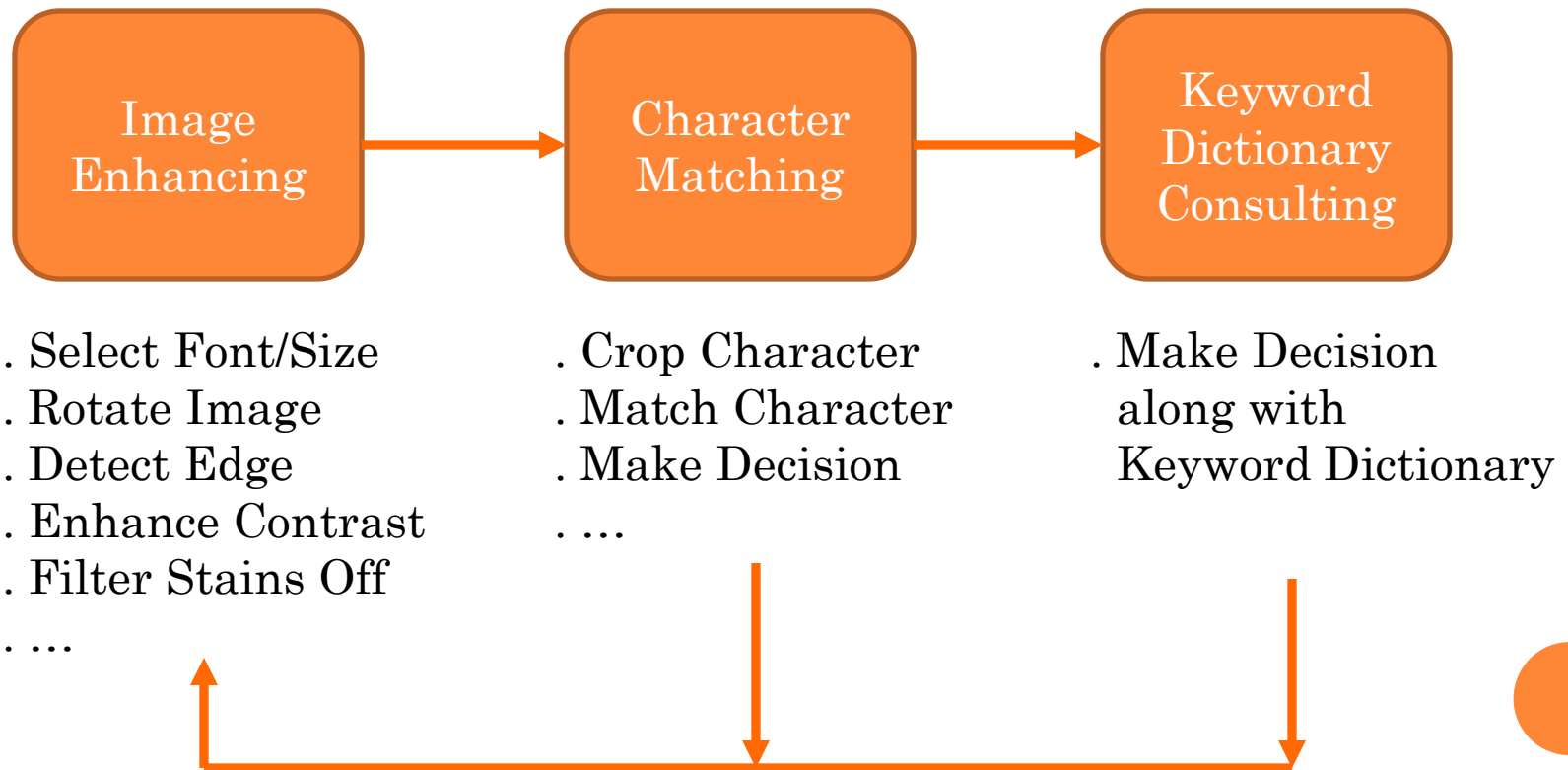
MY NAÏVE OCR PROCEDURE

- Character Matching



MY NAÏVE OCR PROCEDURE

- If considering the process of OCR as image matching...



OBJECTS FOR TODAY

- Consensus about what features you want to cover in the OCR programs
- Initial idea about your software architecture
- Working CVS repository for every team



SAMPLE CODE – JPEG DISPLAYER

○ JPEG Displayer

- Read a JPEG file into an array.
- Display (part of) the image with GTK.

○ Needed Libraries

- libjpeg
- GTK

○ References

- libjpeg <http://www.cim.mcgill.ca/~junaed/libjpeg.php>
- GTK <http://www.cs.dartmouth.edu/~xy/cs23/gtk.html>

