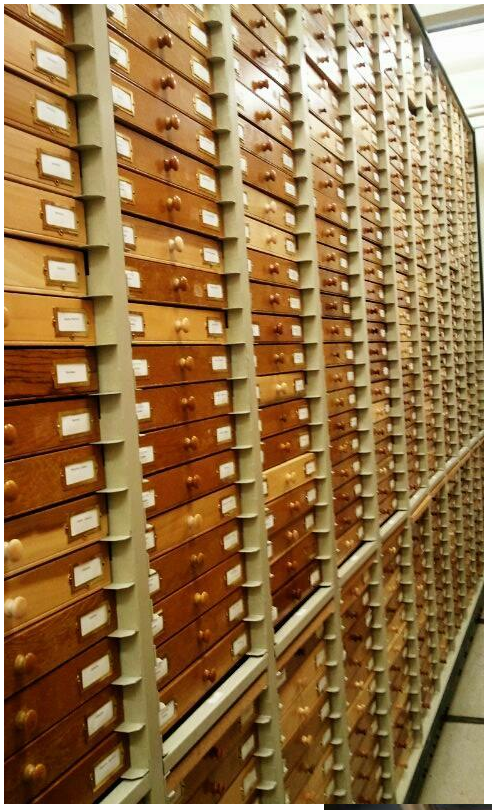
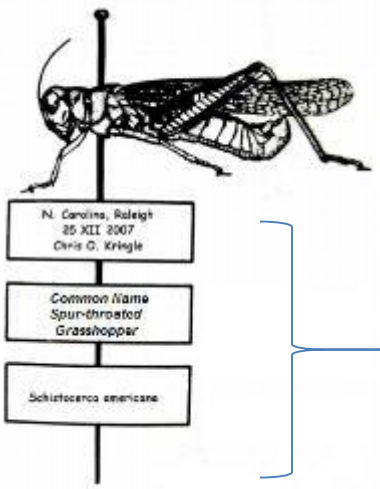
Label Image Automation

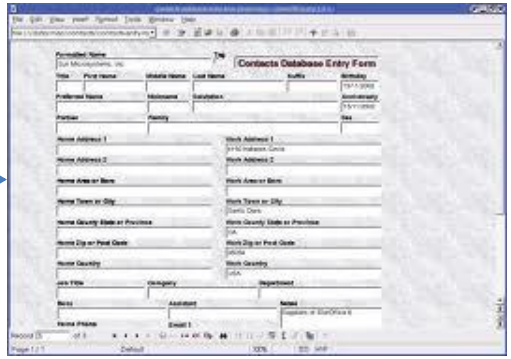
Nicola Ferrier

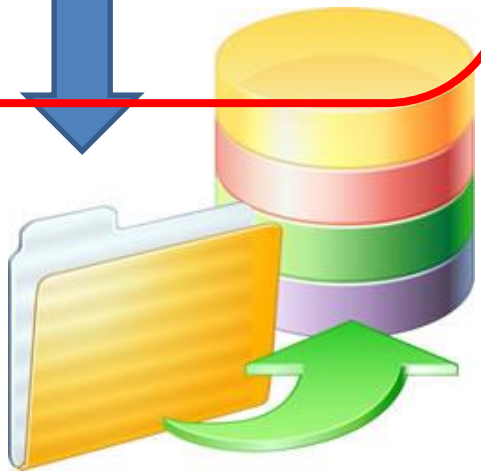
Mark Hereld

Argonne National Laboratory




N. Carolina, Raleigh
25 XII 2007
Chris O. Kringle
Common Name
Spur-throated
Grasshopper
Schistocerca americana


Contacts Database Entry Form



Effects of Image Quality, Pixel Scale, and Camera Pose on OCR Accuracy

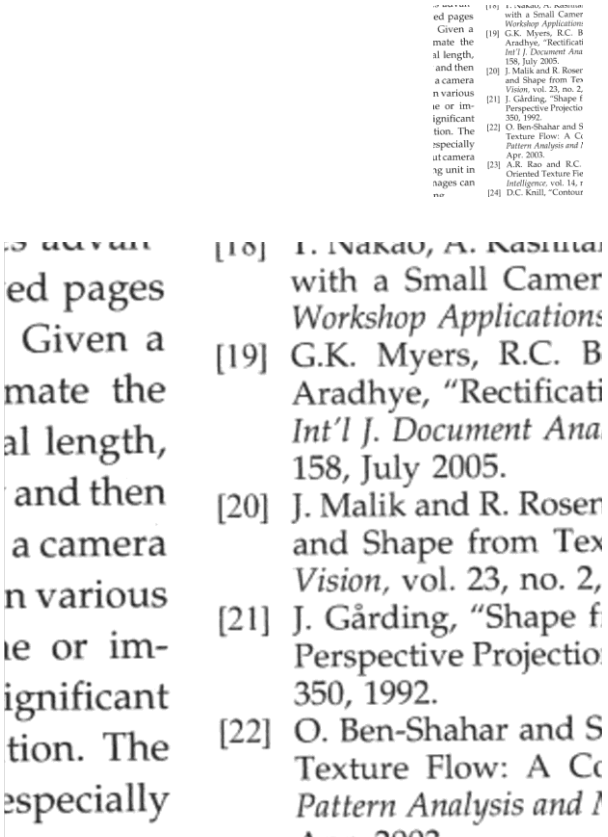
- What resolution do we require in order for OCR to be robust?
- What are the image processing steps that will help prepare imagery into an optimal form?

OCR candidates being tested: Tesseract & Ocropus

- Provide command line interface that is easy to integrate with prototyping scripts
- Provide API for incorporation into performance-driven implementations of the pipeline as it matures (if necessary)
- Modular design enables range of customization paths

Effect of Pixel Scale

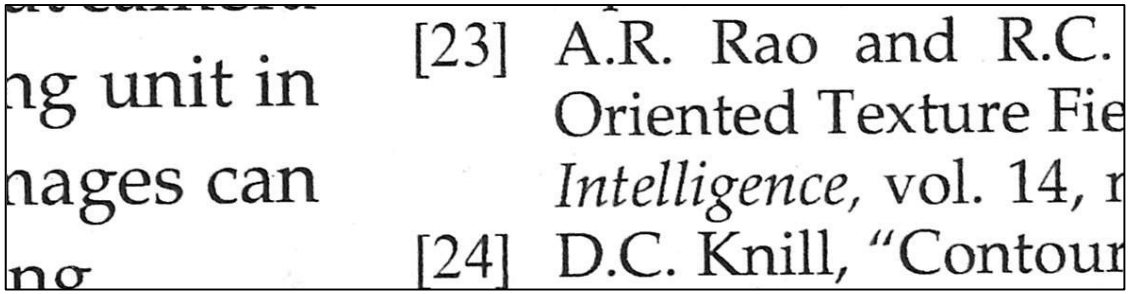
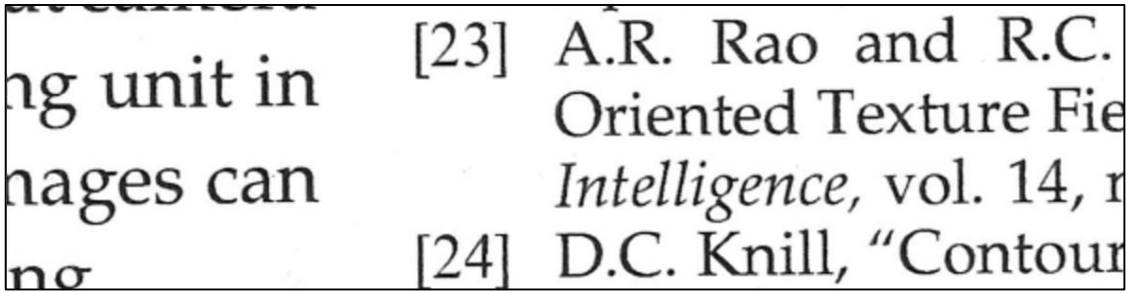
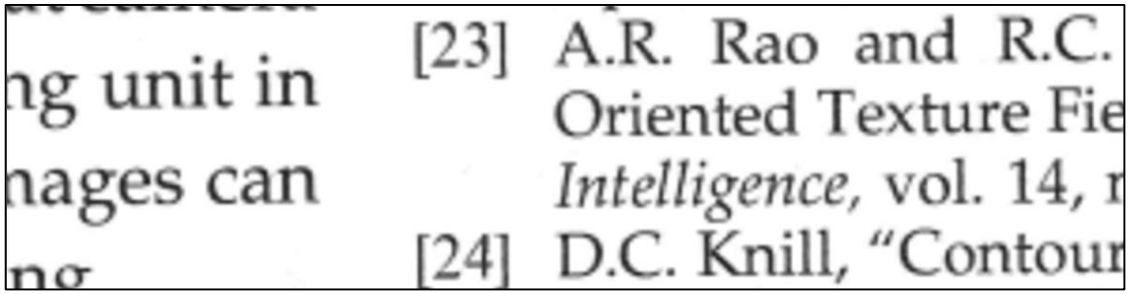
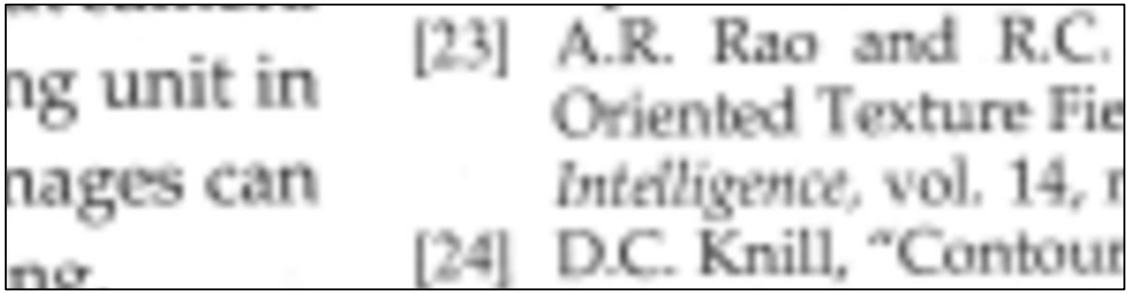
OCR ERRORS



with a Small Camer
Workshop Applications
[19] G.K. Myers, R.C. B
Aradhye, "Rectificati
Int'l J. Document Ana
158, July 2005.
[20] J. Malik and R. Roser
and Shape from Tex
Vision, vol. 23, no. 2,
[21] J. Gårding, "Shape f
Perspective Projectio
350, 1992.
[22] O. Ben-Shahar and S
Texture Flow: A Cc
Pattern Analysis and I

unreadable

1

	SCALE	ERRORS
	1x	0
	2x	0
	4x	1
	8x	11

Effect of Smoothing

ERRORS

8

1

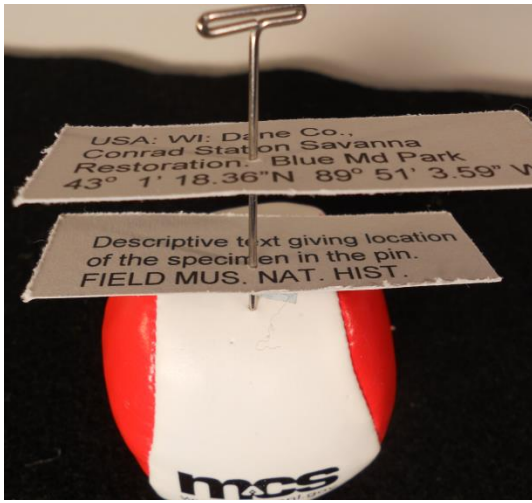
OCR

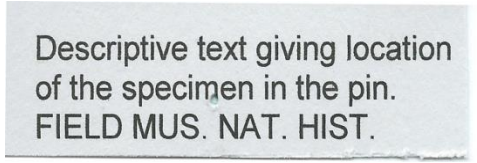
- Size matters – each OCR package seems to have a preferred size for characters
- Characters resized might need post-processing (e.g. smoothing) to successfully be read
- OCR is sensitive to the orientation of text (experiments not shown)

➡ Implications for labels: need carefully aligned and re-sized image of text for each label

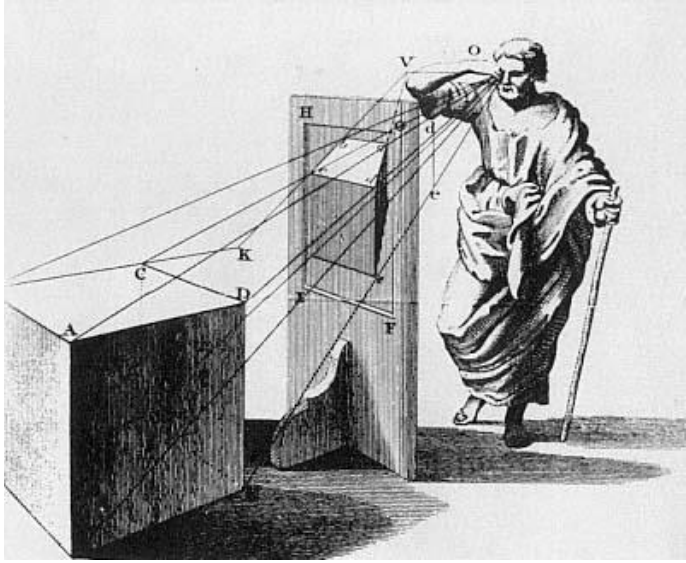
Obtaining an OCR-ready image

- Positioning of Camera(s)
  - multiple labels in one image
  - high quality image of all labels
    - minimize perspective distortion
    - characters scaled for successful OCR

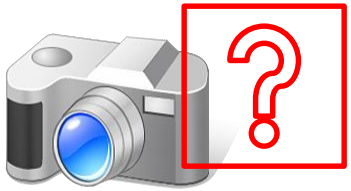


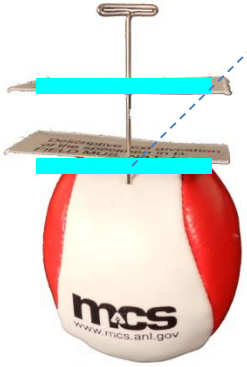


- Image processing
  - Clean up
  - Alignment
  - Composite image from multiple views



Can’t see lower label from viewpoints above the line





Can’t see upper label from viewpoints below this line

HINT

HINT

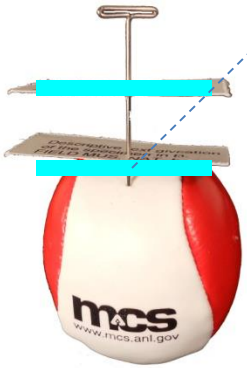
HINT

HINT

HINT

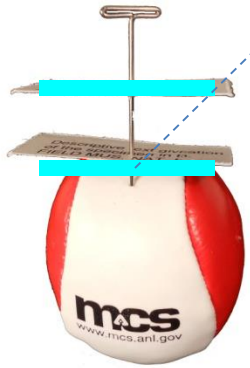
HINT

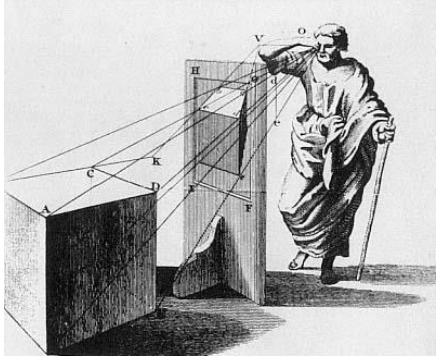
Perspective
Distortion
Increases



Character Size Decreases
(variable across the label)

HINT HINT HINT HINT





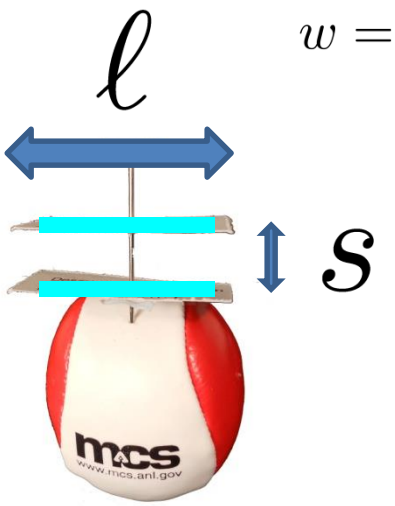
Camera Positioning :
we can mathematically model the size/distortion of the characters as a function of the camera optics (focal length, pixel scaling) and the camera pose (distance and orientation)

$w =$ width of label in image pixels — "goal"

$k_x =$ optics parameters (f, scaling)
$\ell =$ label dimension — "fixed"

$d_x, d_z =$ camera position
$\phi =$ camera orientation — "controlled"

$$w = k_x \frac{d_x + \ell/2 \sin\phi}{d_z + \ell/2 \cos\phi}$$

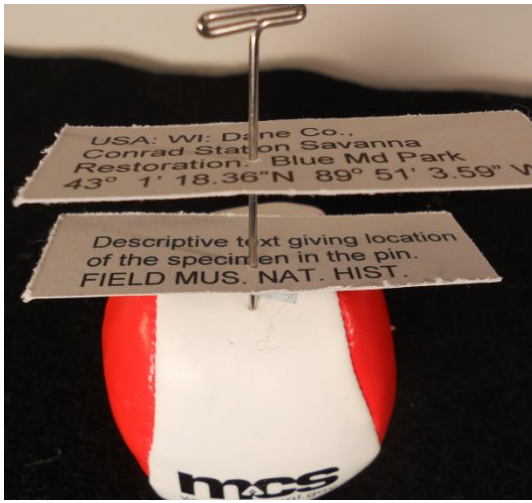


Size of individual characters

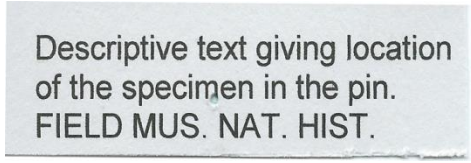
$$c = f(k_x, d_x, d_z, \phi; \text{position on label})$$

Obtaining an OCR-ready image

- Positioning of Camera(s)
  - multiple labels in one image
  - high quality image of all labels
    - minimize perspective distortion
    - characters scaled for successful OCR



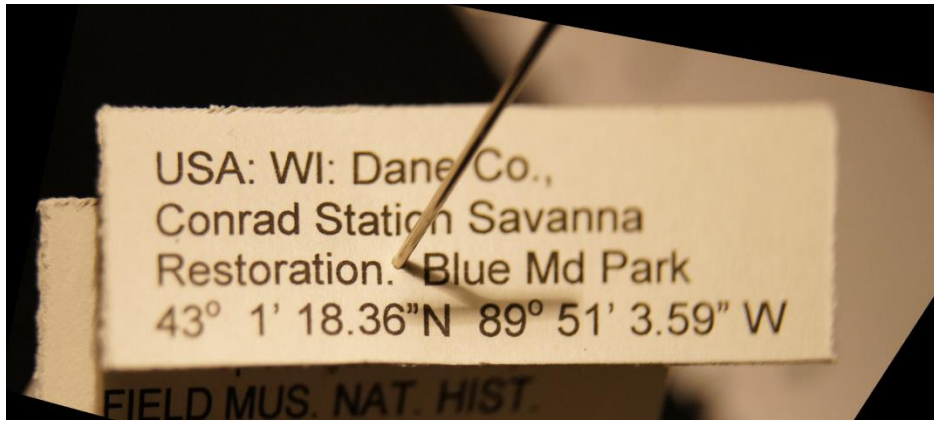
- Image processing
  - Clean up
  - Alignment
  - Composite image from multiple views

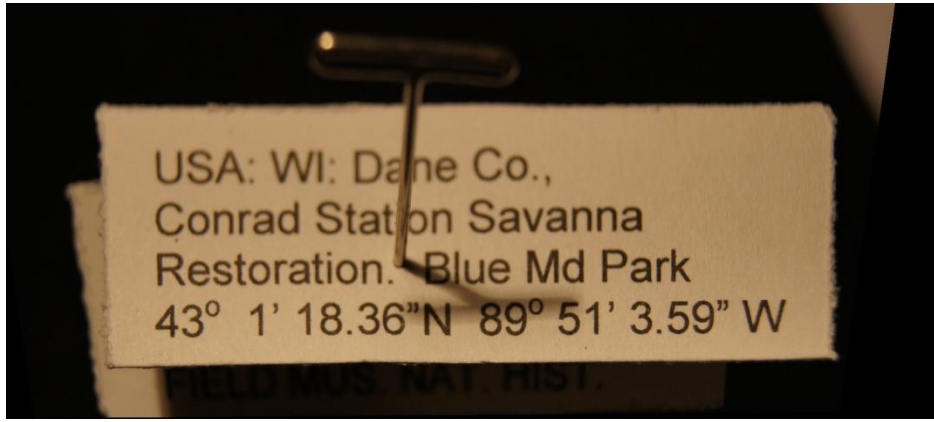


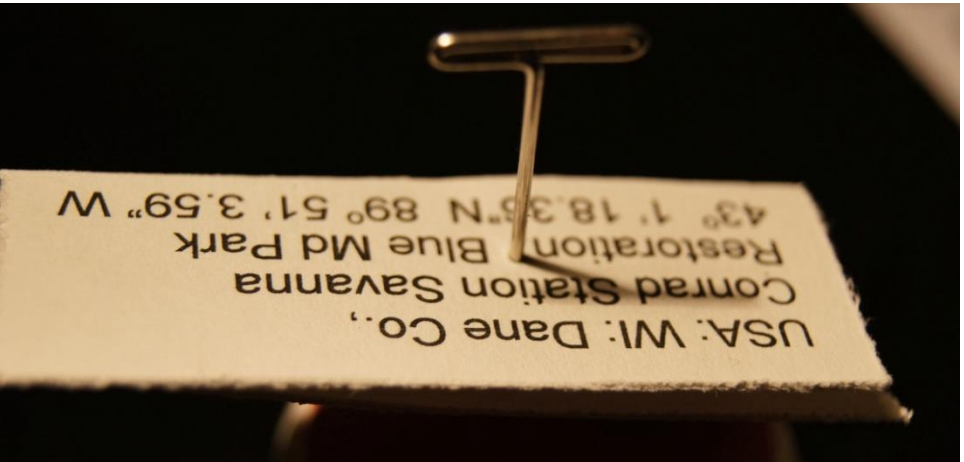
Can we re-align?

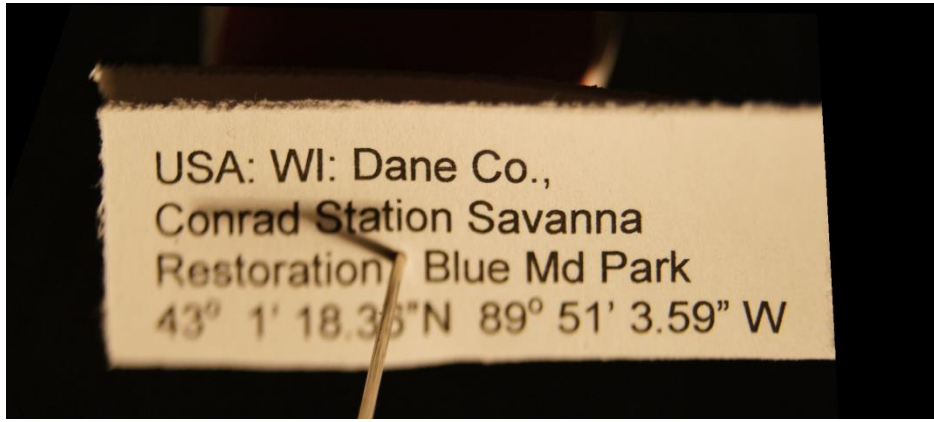
Camera Images

Re-Aligned

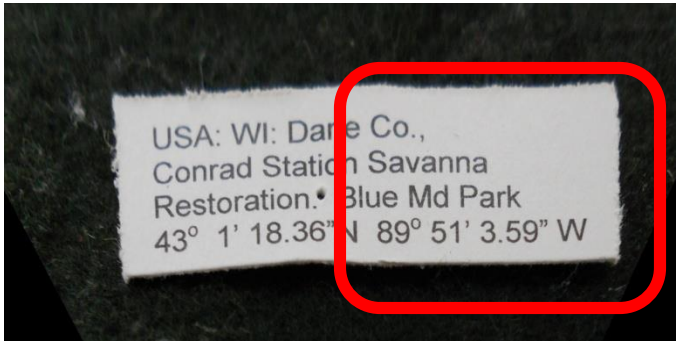


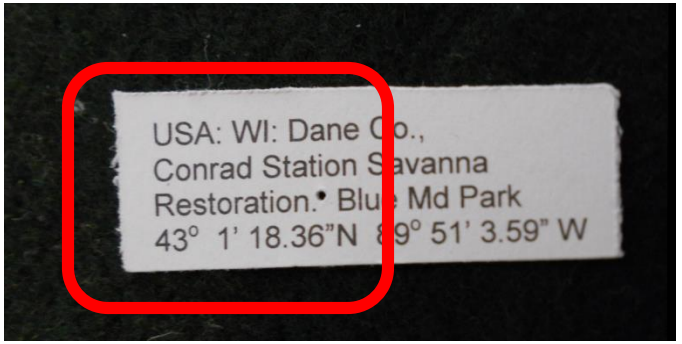






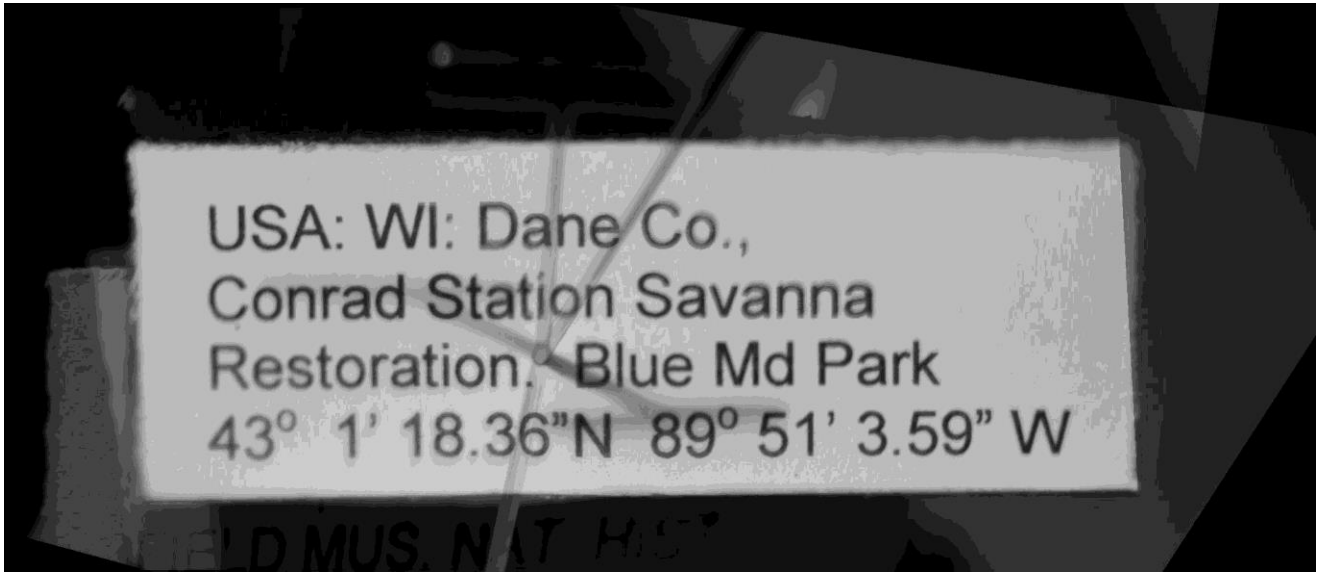
Can we make a composite or use just pieces of each?





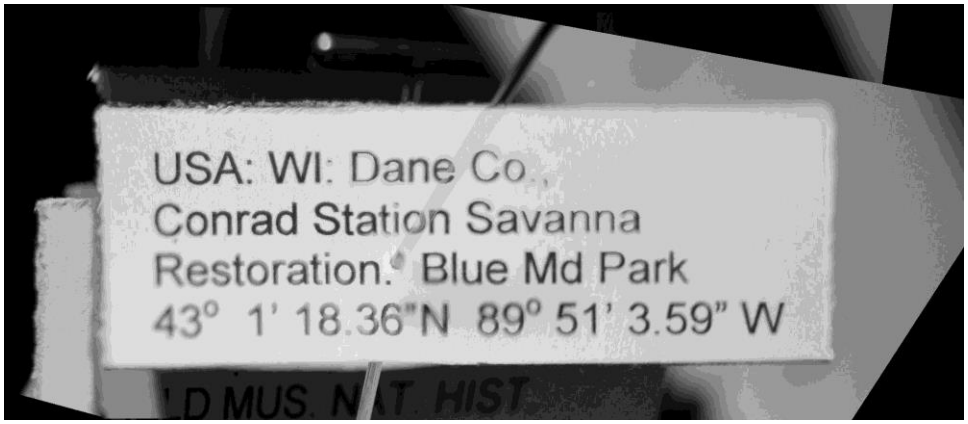
Composite Image
via registration & averaging

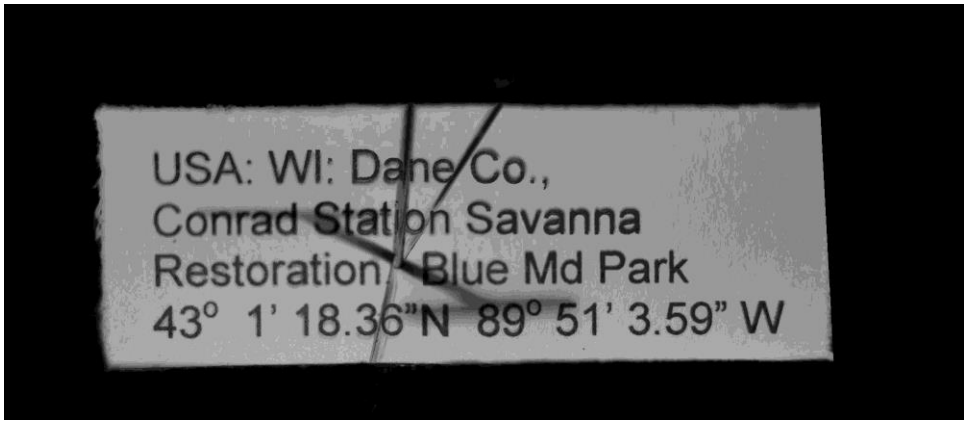
- Used “best in class” registration to automatically align images
- Average pixels from all aligned images

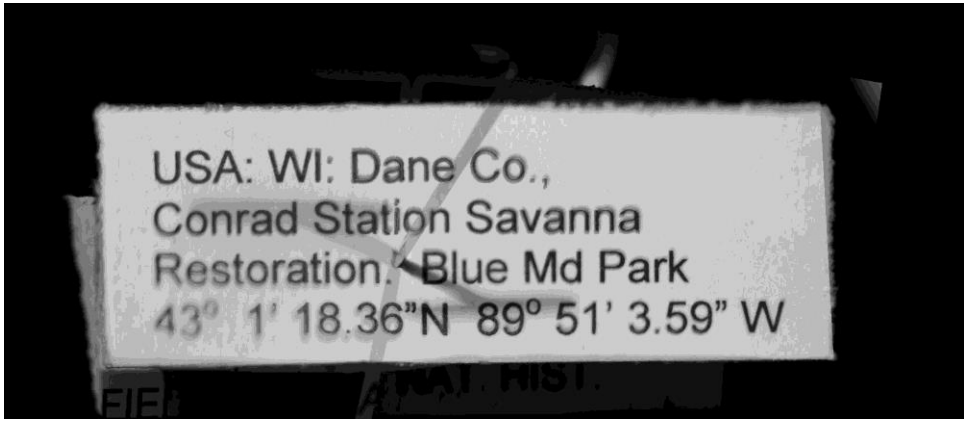


Does not adjust for:

- Poor quality (i.e. out of focus pixels)
- Obstructions (e.g. pin)







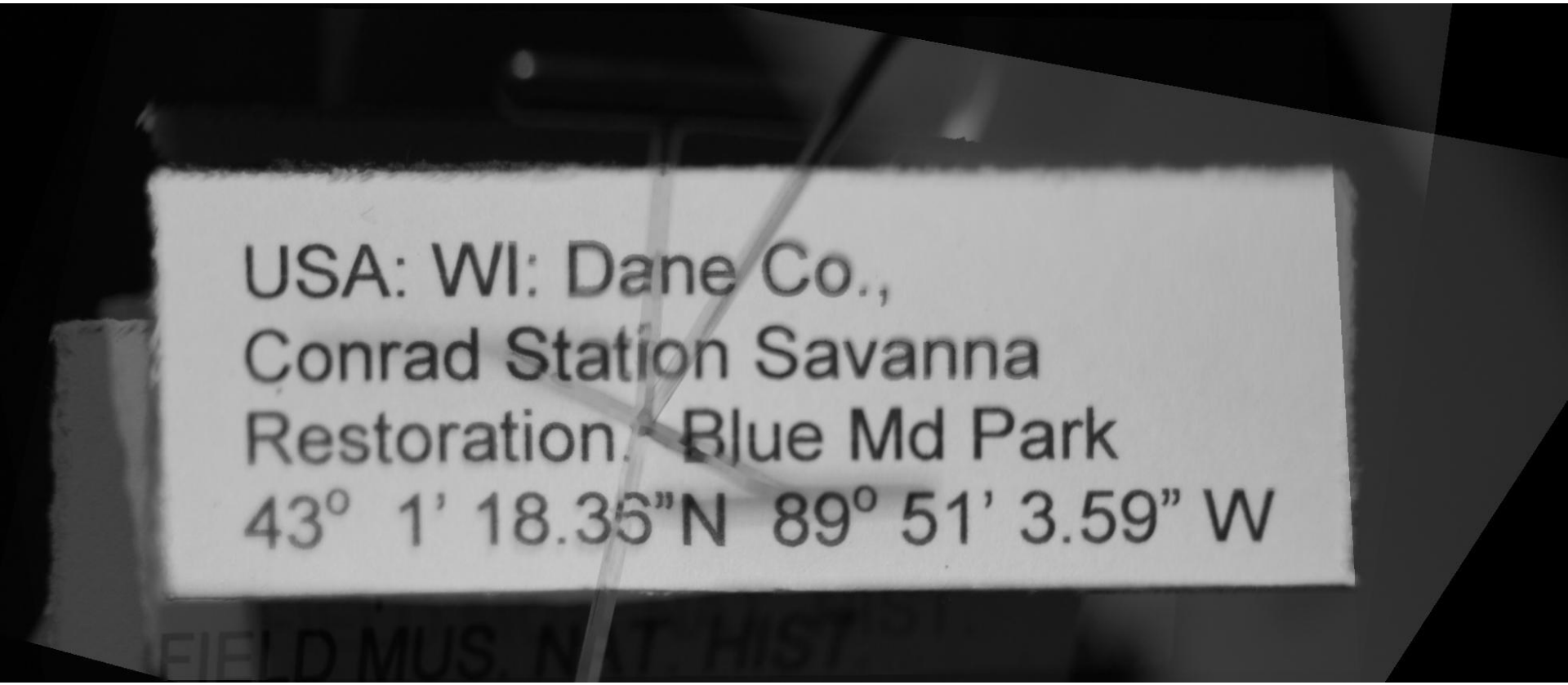
Winner take all

One pixel selected from each image based on:

1) Max value
2) Min value
3) median value

Future: Need to consider other methods – e.g. Pixel "quality" measure?

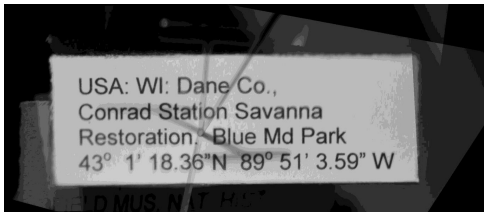
Try in-painting?



- Need a decent mask
- Doesn't seem to be an improvement over simpler methods

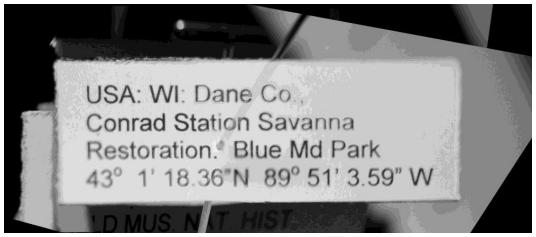
OCR on composite images

- Average



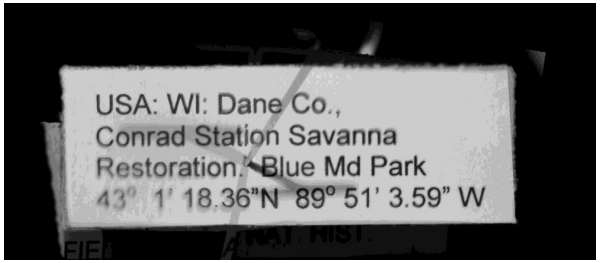
USA: WI: Dane Co.,
Conrad Station Savanna
Restoratior[ib]N3lue Md Park
43^{0} 1' 18.36"N 89° 51' 3.59" W

- Max



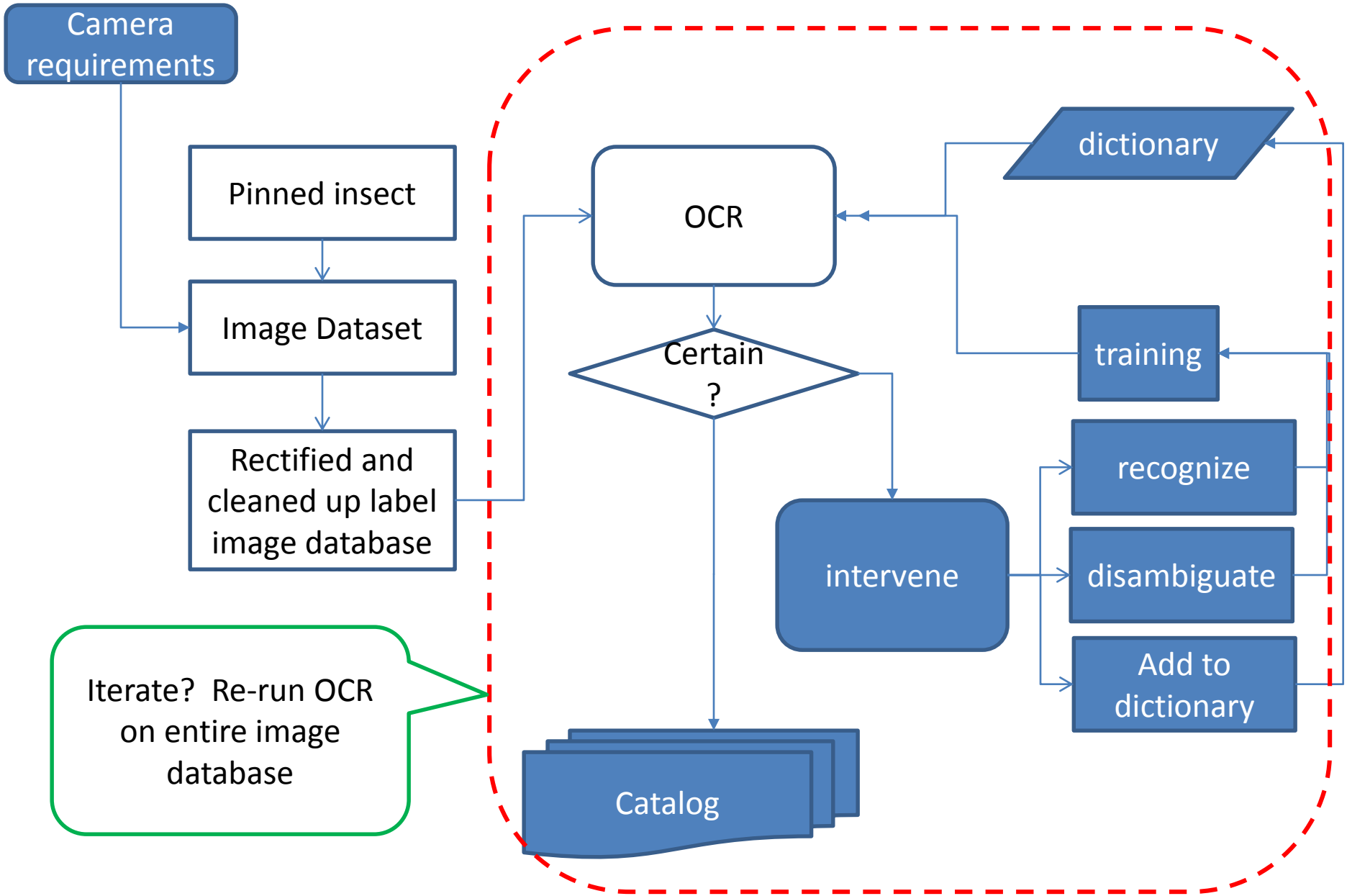
USA: WI: Dane Co●
Conrad Station Savanna
Restoration! Blue Md Park
43^{0} **1' 18_36"N 89" 51' 3.59" W**

- Median



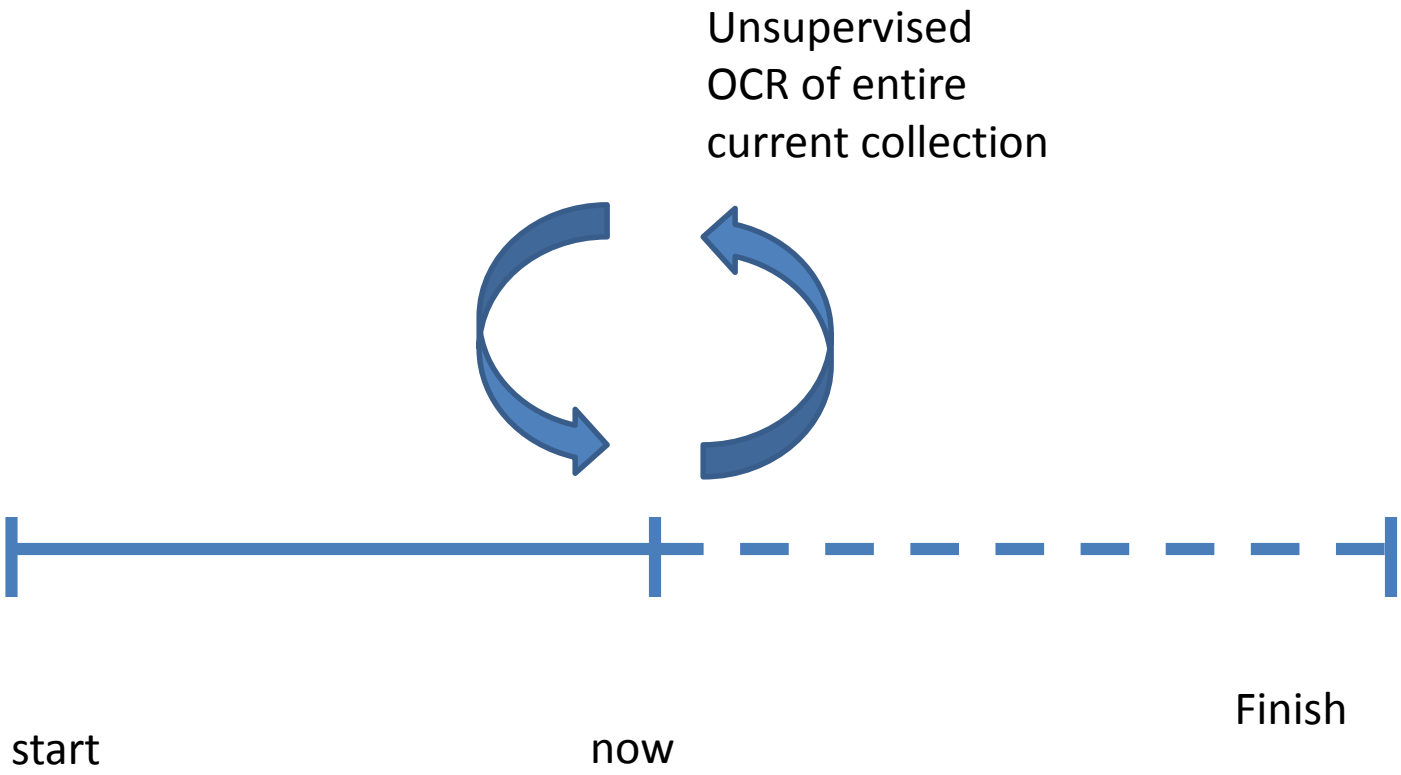
USA: WI: Dane Ca,
Conrad Station Savanna
Restoration.431ue Md Park
43^{0} 1' 18.36"N 89° 51' 3.59“W

A Label Automation Flowchart



Tesseract OCR timing test: 1000 characters per second

Data storage: in TBs range



Label Automation

- OCR may be feasible solution
  - Train & improve performance
- Need to collect high quality images
  - Enough resolution for successful OCR
  - Handle occlusion, poor focus, distortion
- Computation (OCR, possibly improved image processing) is not a bottleneck