



# FamilySearch Scanning

(Scanstone)

An Automated Exposure  
Method For Scanning  
Microfilm

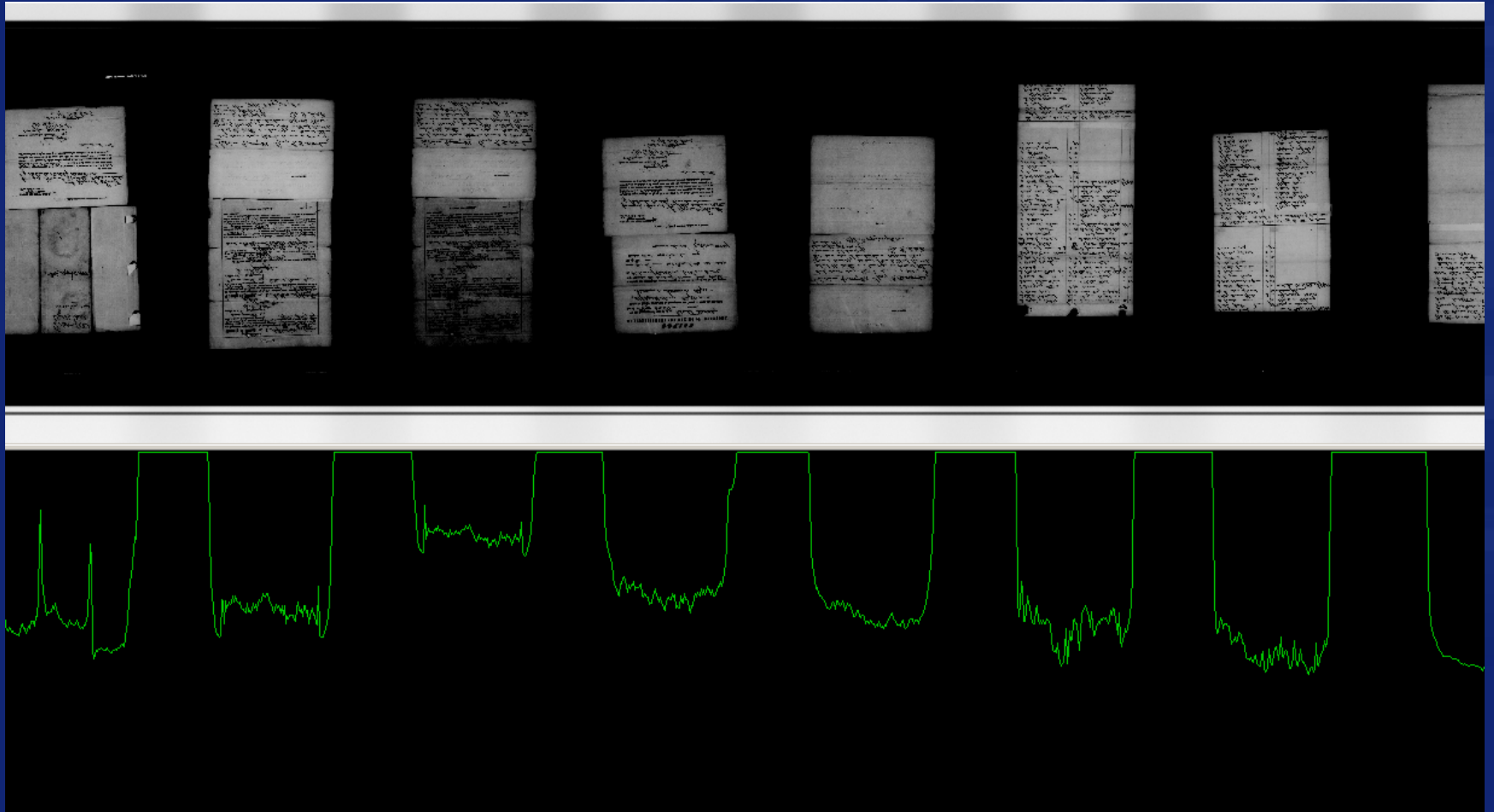
Heath Nielson  
[nielsonhe@ldschurch.org](mailto:nielsonhe@ldschurch.org)



# Digitizing Microfilm

- ❖ Scan as efficiently as possible
  - At 30 minutes a roll, it would require a single scanner 142 years of constant scanning to scan 2.5 million rolls
- ❖ Maintain consistent (good) quality
  - Adjust lamp level

# Changing light level



Lamp level: 200



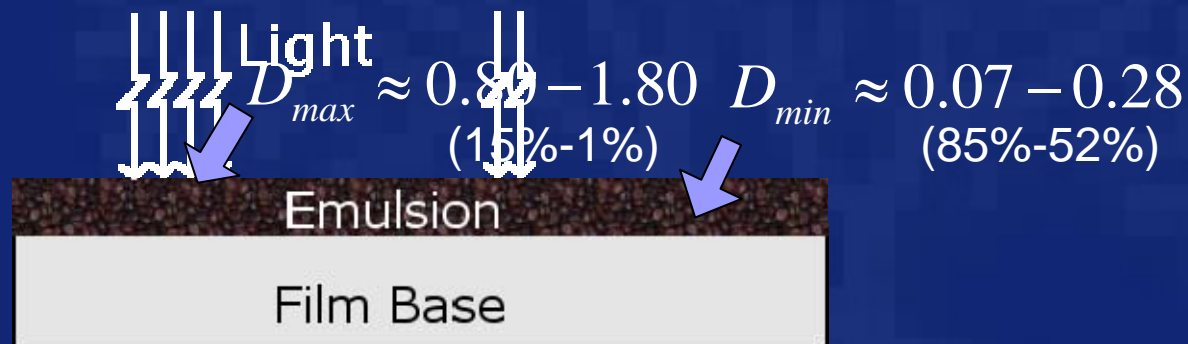
# Goal

- ❖ To maintain efficiency and quality, we need to automatically identify the optimal lamp setting for a given film
- ❖ Optimal lamp level is dependent upon:
  - Film density
  - Focal length
  - Light output

# Film Density

- ❖ Amount of incident light ( $I_i$ ) transmitted through the film ( $I_t$ ) is measured in terms of optical density  $D$

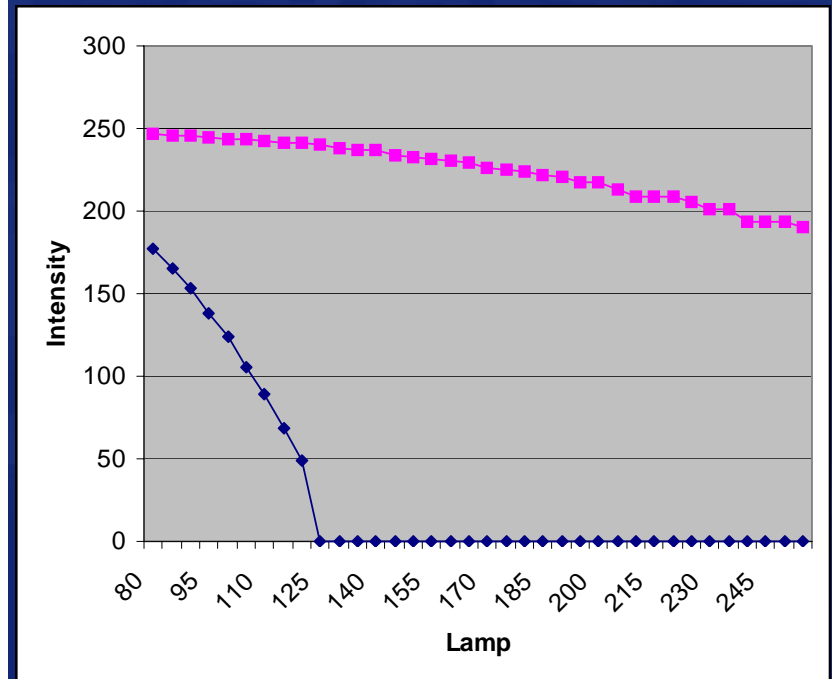
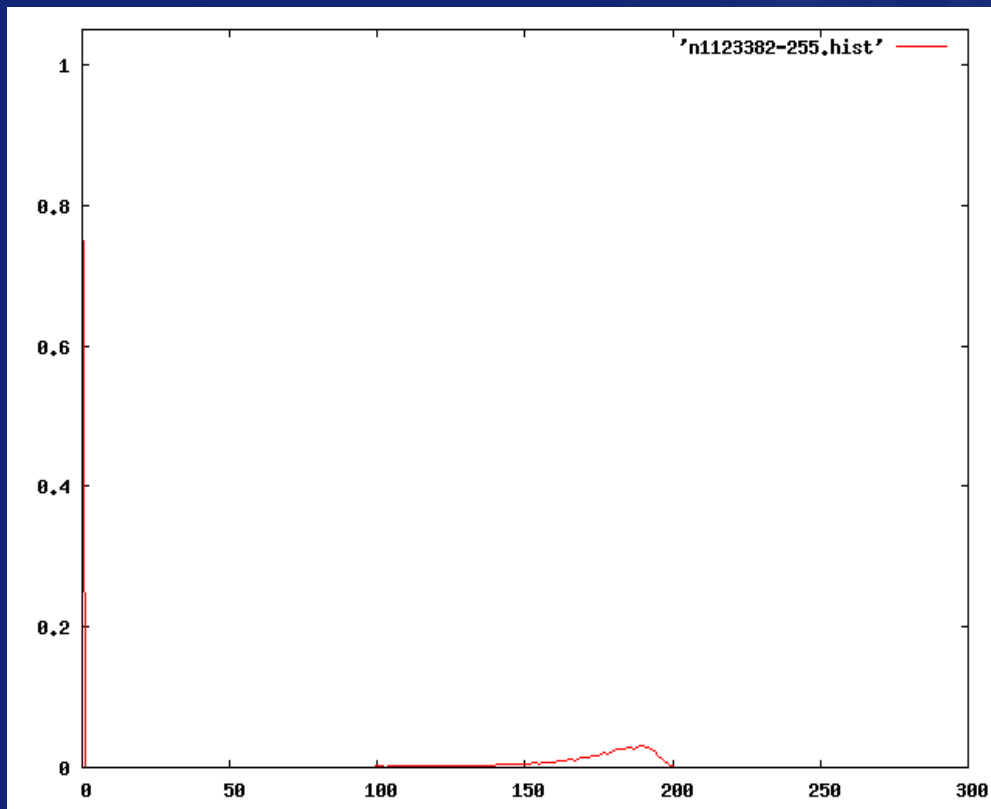
$$D = \log \left( \frac{I_i}{I_t} \right) \quad 0 \leq D \leq \infty$$



- ❖ Film Contrast is dependent upon range between  $D_{min}$  and  $D_{max}$

# Density vs. Lamp Level

## Histogram View



$$D_{min} \approx 0.16$$

$$D_{max} \approx 1.40$$



# Density Test

- ❖ Created test film comprised of 17 sections
  - The density of each section was constant
  - The sections represented a range of densities from 0.03 to 1.50
- ❖ Film was scanned multiple times across a range of different lamp levels
- ❖ The mean intensity for each density was calculated

# Density Roll



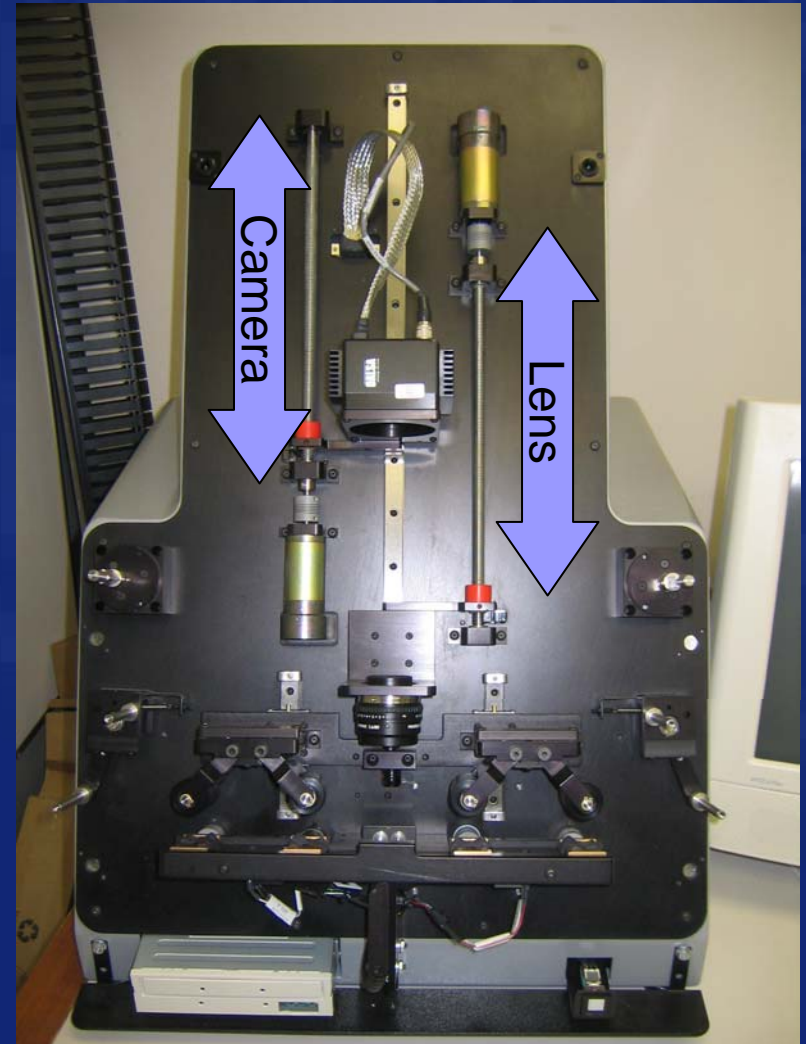


**200DPI-16x**

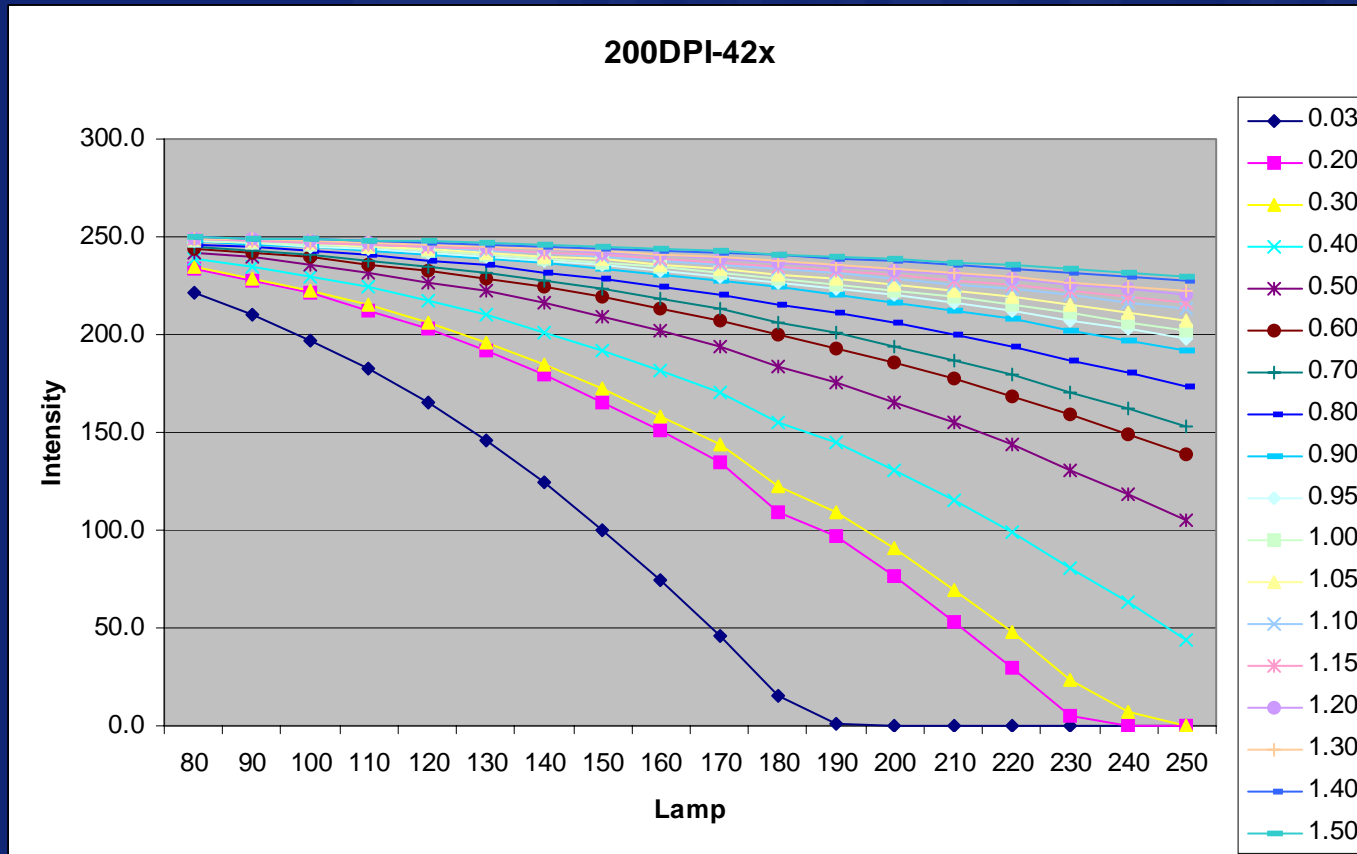
Lamp	0.03	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.30	1.40	1.50	
80	158	200	202	215	225	230	232	235	238	240	242	244	246	248	250	252	254	256	258
90	122	180	182	195	205	210	212	215	218	220	222	224	226	228	230	232	234	236	238
100	85	160	162	175	185	190	192	195	198	200	202	204	206	208	210	212	214	216	218
110	38	135	138	150	160	165	168	170	172	175	178	180	182	185	188	190	192	195	198
120	0	105	110	120	130	135	138	140	142	145	148	150	152	155	158	160	162	165	168
130	0	72	80	85	95	100	105	108	110	112	115	118	120	122	125	128	130	132	135
140	0	38	52	55	65	70	75	78	80	82	85	88	90	92	95	98	100	102	105
150	0	0	15	20	30	35	40	42	45	48	50	52	55	58	60	62	65	68	70
160	0	0	0	0	10	15	20	22	25	28	30	32	35	38	40	42	45	48	50
170	0	0	0	0	0	10	15	18	20	22	25	28	30	32	35	38	40	42	45
180	0	0	0	0	0	0	10	12	15	18	20	22	25	28	30	32	35	38	40
190	0	0	0	0	0	0	0	10	12	15	18	20	22	25	28	30	32	35	38
200	0	0	0	0	0	0	0	0	10	12	15	18	20	22	25	28	30	32	35
210	0	0	0	0	0	0	0	0	0	10	12	15	18	20	22	25	28	30	32
220	0	0	0	0	0	0	0	0	0	0	10	12	15	18	20	22	25	28	30
230	0	0	0	0	0	0	0	0	0	0	0	10	12	15	18	20	22	25	28
240	0	0	0	0	0	0	0	0	0	0	0	0	10	12	15	18	20	22	25
250	0	0	0	0	0	0	0	0	0	0	0	0	0	10	12	15	18	20	22

# Focal length

- ❖ Camera and lens move
  - Affects the amount of light received by the camera
  - Position is determined by the selected DPI and reduction ratio of the film prior to scanning

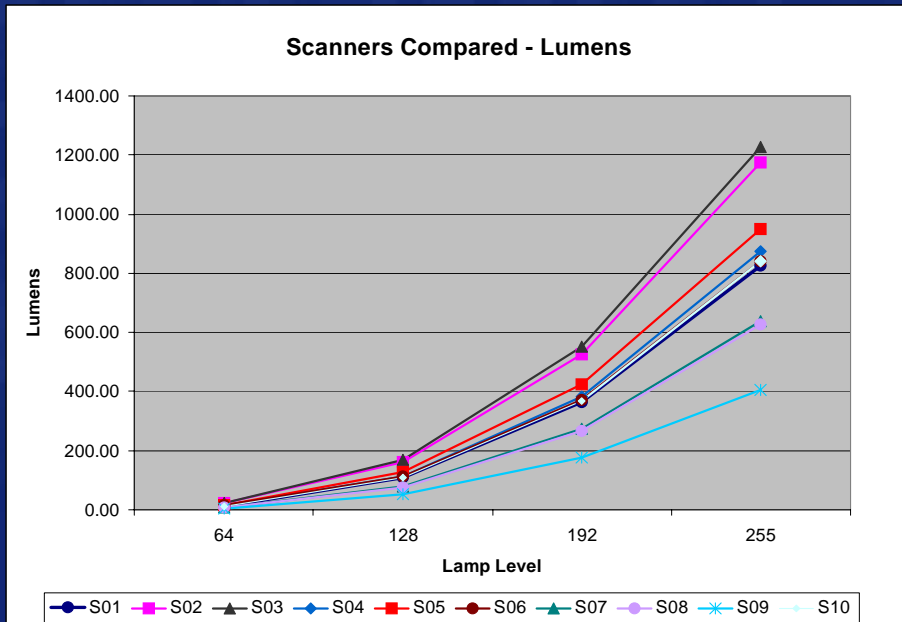


# Focal Length vs. Lamp Level



# Scanner Light Output

- ❖ Light output is inconsistent across scanners
- ❖ Mechanical differences
  - Presence of diffuser
  - Out of calibration
  - Age of lamp





# Developing an Exposure Strategy

## ❖ Input

- Unknown film density
- Variable focal length
- Inconsistent light output across scanners

## ❖ Output

- Inter-frame background and handwriting with an intensity  $\sim 0$
- Frame background with an intensity  $\sim 200$  or greater

## ❖ How should the lamp be adjusted?

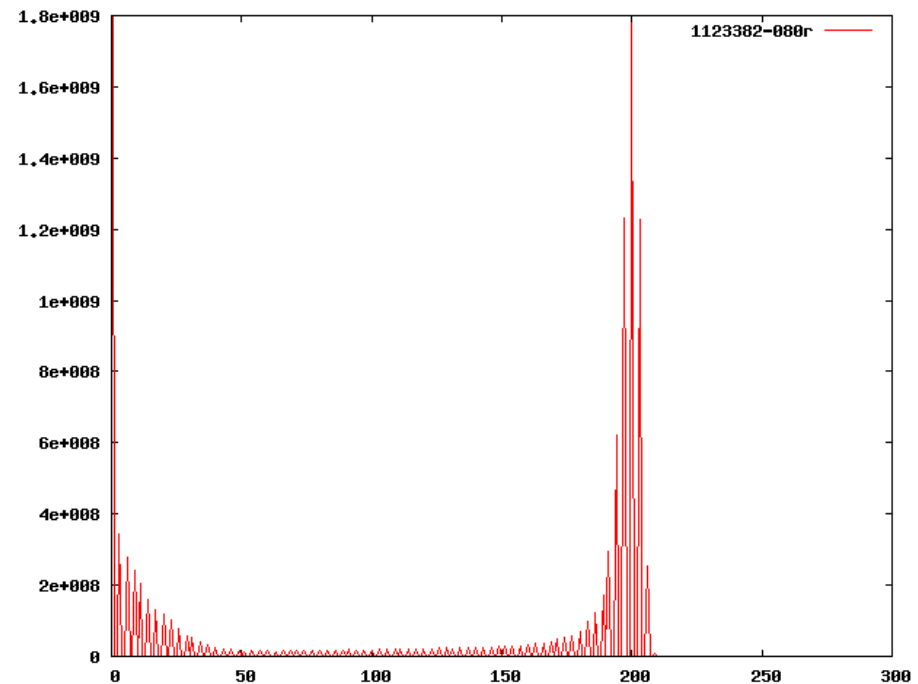
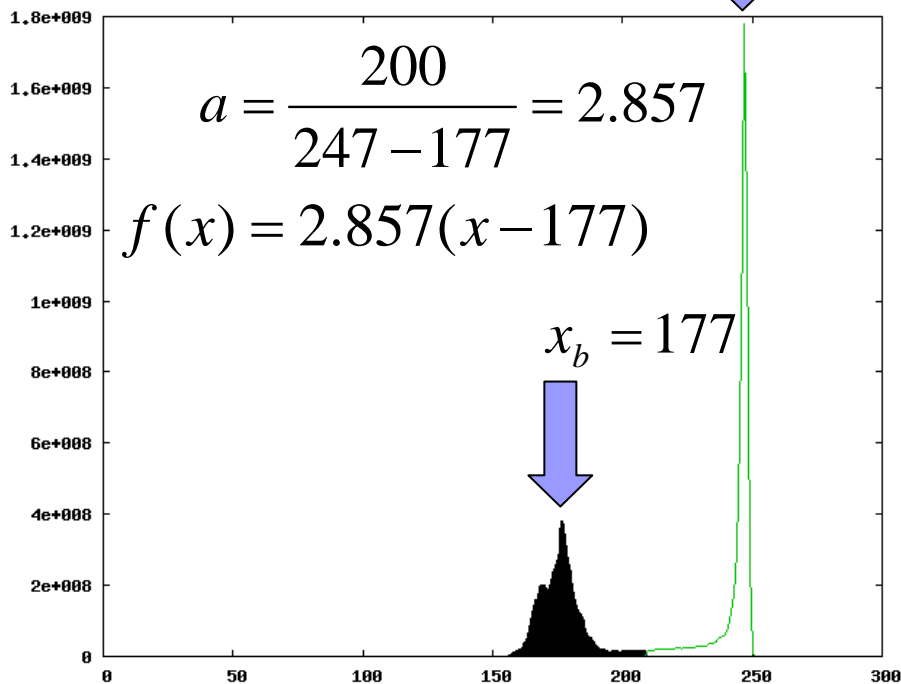
# Histogram Stretching

- ❖ Set the lamp to an “average” value for the selected focal length
- ❖ Scan the film
- ❖ Find the mode of the black and white distributions ( $x_b, x_w$ )
- ❖ Create a stretching function to remap the scanned pixel data

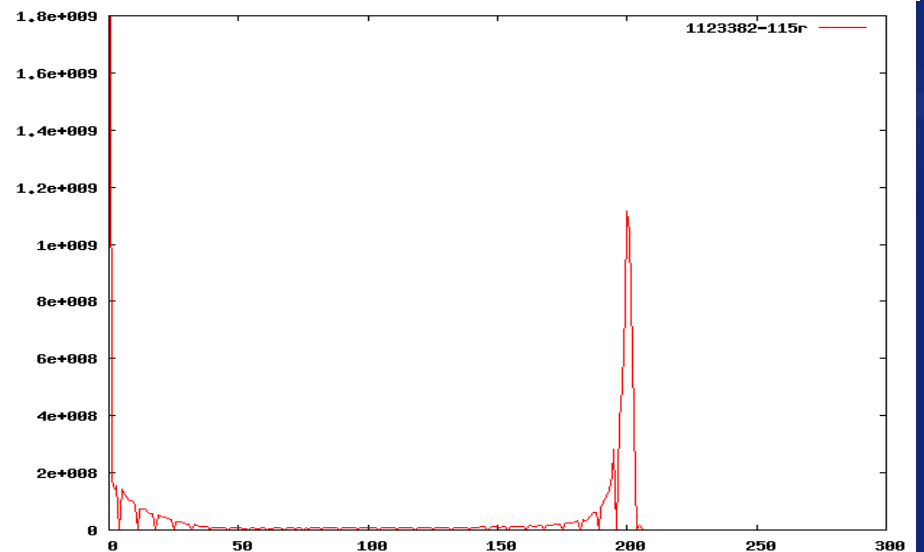
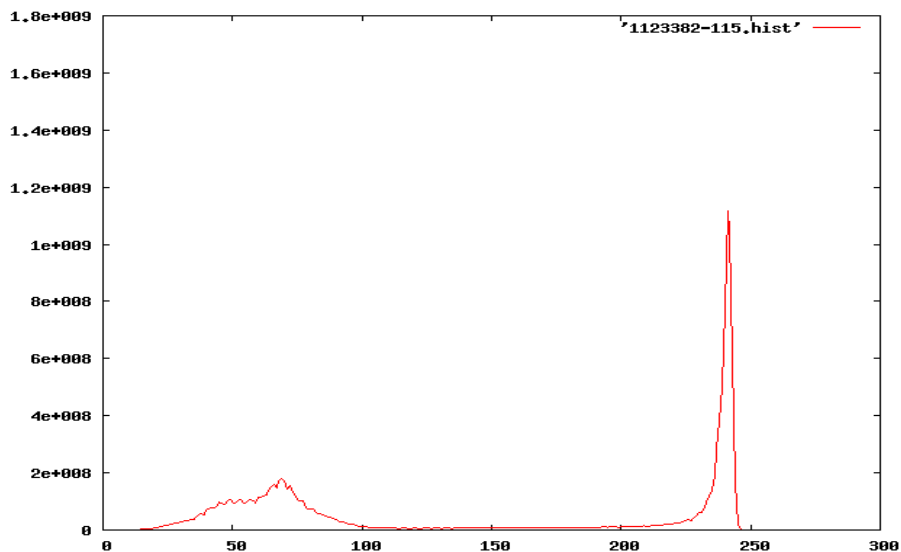
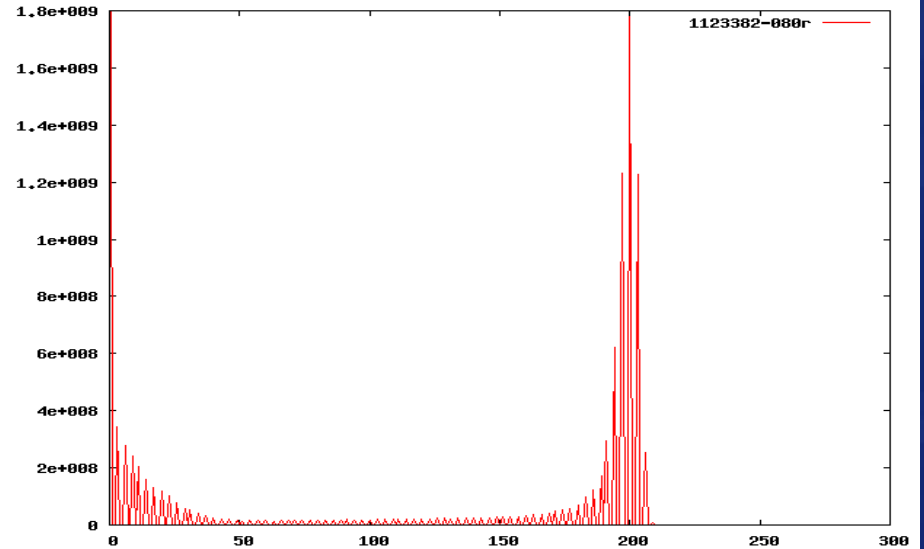
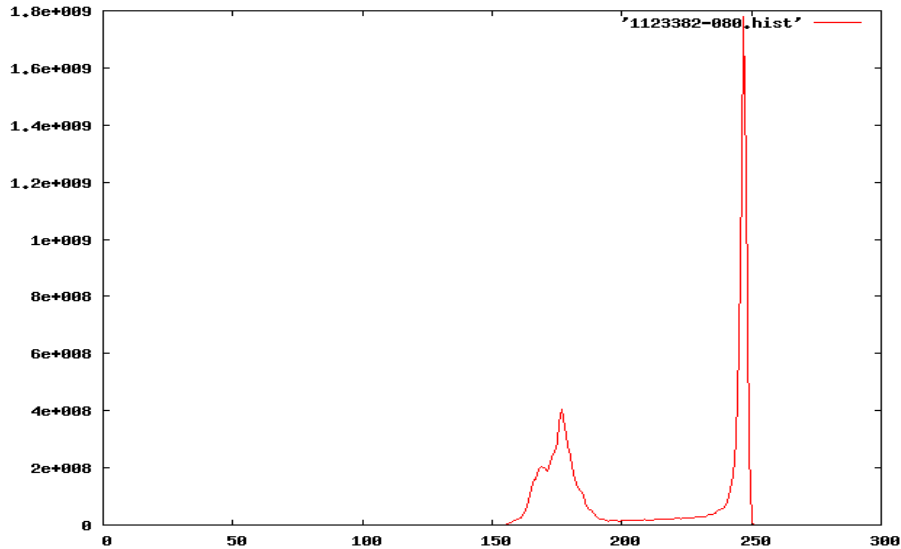
$$y = a(x - x_b) \quad a = \frac{200}{x_w - x_b}$$

# Example

$$x_w = 247$$



# Invariant to minor lamp changes







# Potential problems

- ❖ Still some dependence upon lamp level
  - Extreme values cause camera saturation
- ❖ Assumes a single black or white distribution
  - What value produces acceptable results for a multi-modal distribution?
- ❖ Fewer gray levels by stretching the image data post-scan



# Manual Audit

- ❖ A random sample of 469 films were selected from the 2.5 million rolls
- ❖ Each film was scanned and audited for quality
- ❖ If quality was deemed unacceptable, the contrast could be overridden



# Results

- ❖ 185 (39%) films were manually adjusted
  - As the test proceeded, the number of films adjusted decreased
  - In 155 films (84%), the contrast was increased
- ❖ Average number of gray levels
  - 187 before audit
  - 169 after audit

# Future Directions

- ❖ Better metrics for determining a 'good' quality image
- ❖ Localized histogram stretching
  - Per region
  - Per frame
- ❖ 12-bit scanning
  - Decreases further, dependence upon lamp
  - Increases likelihood of capturing full dynamic range between  $D_{min}$  and  $D_{max}$



# Conclusion

- ❖ We have shown an exposure algorithm which produces consistent output in the presence of differing amounts of light and is automated, maintaining the efficiency necessary to scan large numbers of film