IA applications: image classification

Everardo Bárcenas

Image decomposition

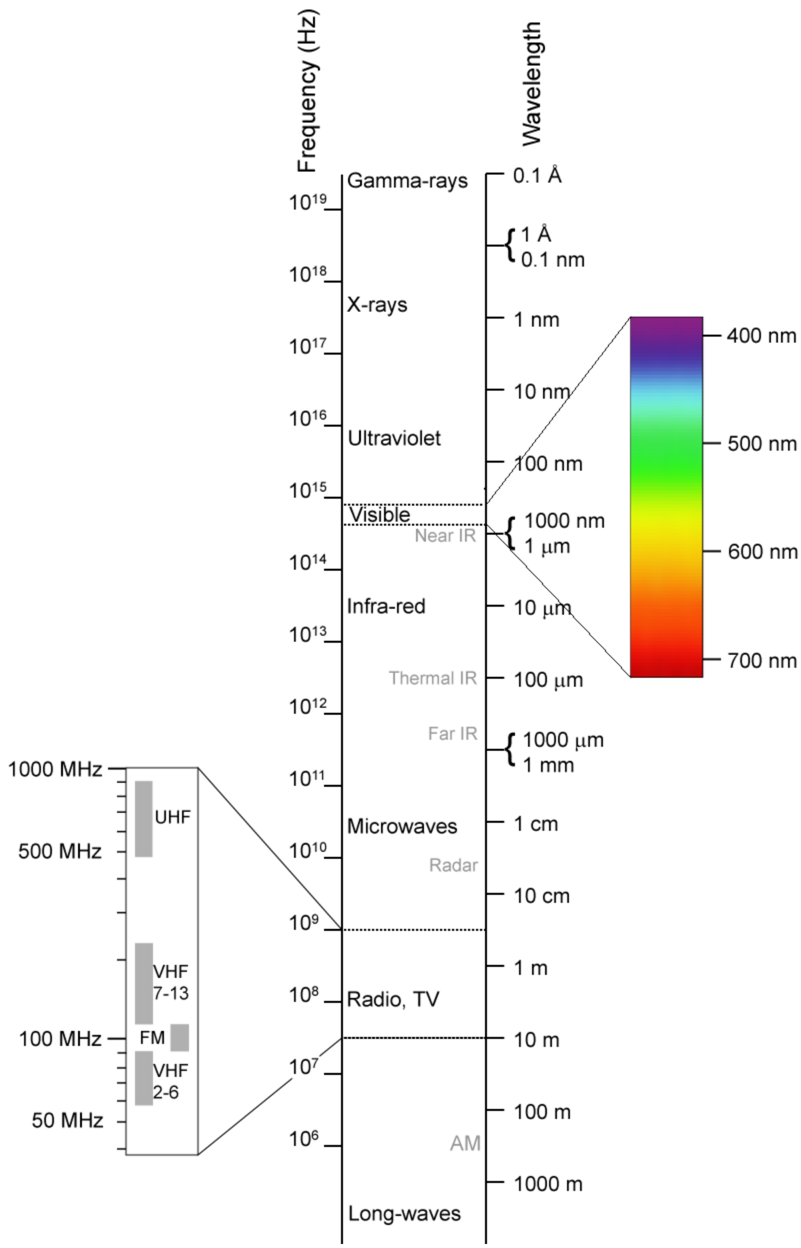


Image composition: RGB model (Discrete representation)

- $f : [0, 1] \mapsto D$
- D is commonly bounded from 8 bits to 64 bits

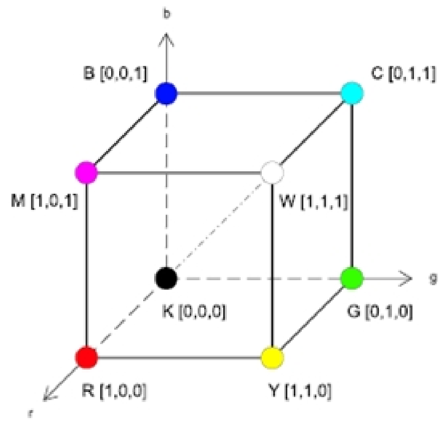
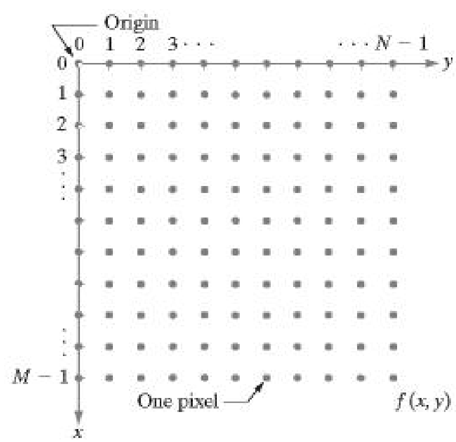


Image composition: data structure



$$A = \begin{bmatrix} A_{0,0} & A_{0,1} & \dots & A_{0,n-1} \\ A_{1,0} & A_{1,1} & \dots & A_{1,n-1} \\ \vdots & & & \\ A_{m-2,0} & A_{m-2,1} & \dots & A_{m-2,n-1} \\ A_{m-1,0} & A_{m-1,1} & \dots & A_{m-1,n-1} \end{bmatrix}$$

Point transformations: gray level

Gray level For a range image $[0, L-1]$ in the RGB model, the gray level image is obtained by

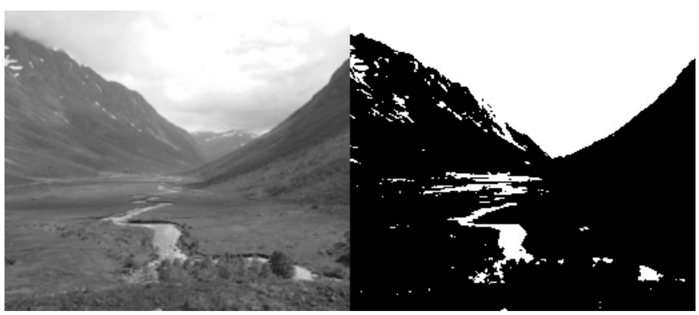
$$(r', g', b') = \begin{cases} (r, r, r) & \text{or} \\ (g, g, g) & \text{or} \\ (b, b, b) \end{cases}$$

provided that (r, g, b) is the input.

Point gray level transformations: threshold For a range image $[0, L-1]$, the negative image is obtained by

$$s = \begin{cases} 0 & \text{if } r \leq c \\ L-1 & \text{otw.} \end{cases}$$

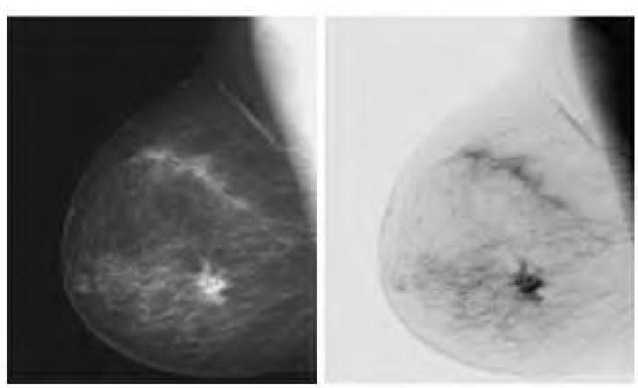
where r is the input and c is a constant.



Point gray level transformations: negatives For a range image $[0, L-1]$, the negative image is obtained by

$$s = L - 1 - r$$

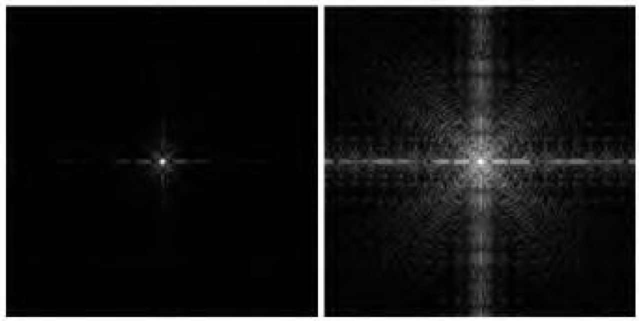
where r is the input.



Log transformations For a range image $[0, L-1]$, the logarithmic transformation is obtained by

$$s = c\log(1+r)$$

where r is the input, and c is a constant.



Power-law transformations For a range image $[0, L-1]$, the power-law transformation is obtained by

$$s = cr^{\gamma}$$

where r is the input, and γ is a constant.
Example: Images on the right: power-law transformation for $c = 1$, and $\gamma = 0.3, 0.4, 0.6$, respectively.

Point Transformations

Contrast and Brightness For a range image $[0, L-1]$, the following linear expression represent contrast (multiplication) and brigthness (addition).

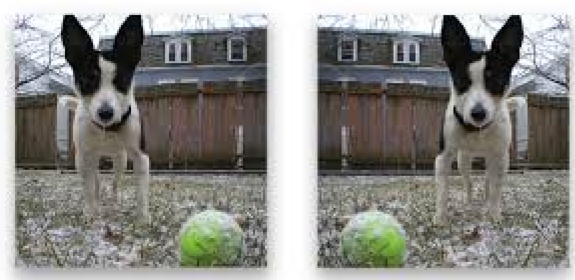
$$s = ar + b$$

where r is the input, and a and b are constants.

Horizontal Reflection For an image A of size $n \times m$, its horizontal reflection is given in another image A' of the same size, such that

$$A'_{i,j} = A_{i,n-j}$$

where $i = 0, \ldots, m-1$ and $j = 0, \ldots, n-1$.



Original Horizontal reflect

Vertical Reflection For an image A of size $n \times m$, its vertical reflection is given in another image A' of the same size, such that

$$A'_{i,j} = A_{m-i,n}$$

where $i = 0, \ldots, m-1$ and $j = 0, \ldots, n-1$.

Original Vertical reflect

Exercises

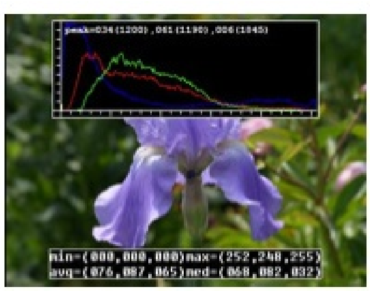
- Zooming
- Shrinking
- Rotation

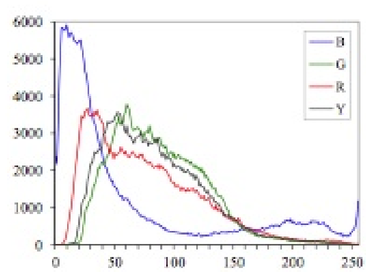
Without Zoom (Original Size) With Zoom

Histogram Equalization

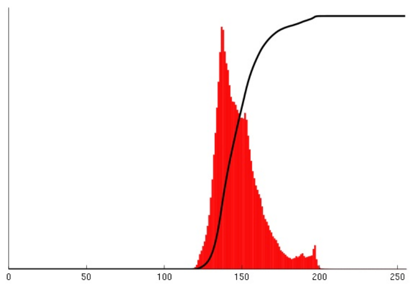
Histogram For a range image $[0, L-1]$, consider cumuluative the occurrence function of a pixel of level i in the source image

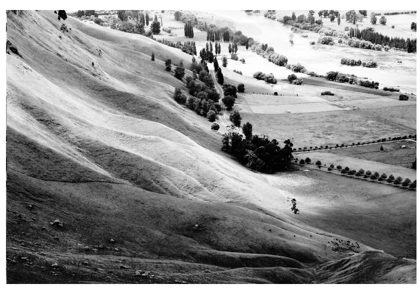
$$f(i)$$

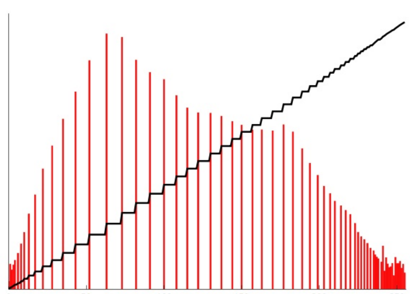




Histogram Equalization







Unequalized image

$$\begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 55 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix}$$

Equalized image

$$\begin{bmatrix} 0 & 12 & 53 & 93 & 146 & 53 & 73 & 166 \\ 65 & 32 & 12 & 215 & 235 & 202 & 130 & 158 \\ 57 & 32 & 117 & 239 & 251 & 227 & 93 & 166 \\ 65 & 20 & 154 & 243 & 255 & 231 & 146 & 130 \\ 97 & 53 & 117 & 227 & 247 & 210 & 117 & 146 \\ 190 & 85 & 36 & 146 & 178 & 117 & 20 & 170 \\ 202 & 154 & 73 & 32 & 12 & 53 & 85 & 194 \\ 206 & 190 & 130 & 117 & 85 & 174 & 182 & 219 \end{bmatrix}$$

Equalization We now define the histogram equalization as follows

$$h(v) = round\left((L-1)\,\frac{f(v) - f_{\min}}{n - f_{\min}}\right)$$

where n is thenumber of pixels in the image and $f_{\min}$ is the minum non-zero value of the cumulative distribution function f.

Linear regression as a classiffier

- Consider a training set of n images $I_1, I_2, \ldots, I_n$, such that each image belongs to a class $\{1, 2, \ldots, m\}$.
- Let p_k the probability of an image to be in class k. Then, dependent variable in the training set is defined by $p_k * 255$
- Consider images as vectors of t pixels.
- Also consider the following linear regression model:

  $$R(\overline{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_t x_t$$

  where x_j is variable over the image pixels.

We say an image $\overline{x}$ is in class k when

$$\min |R(\overline{x}) - p_k * 255|$$