Identifying Computer Graphics Using Second-order Statistics

March 1, 2017

Motivation: Difference in Imaging Mechanism

- Natural Images: Inevitably influenced by both environment and equipment, and the imaging process, is very complex
- Computer Graphs: Can only primitively simulate the generating process of real images

Selection of Color Model:HSV

- The HSV model: Composed of hue(H),saturation(S), and brightness(V).

$$H = \begin{cases} 60(\frac{G-B}{\delta}) & \text{if MAX = R;} \\ 60(\frac{B-R}{\delta}+2) & \text{if MAX = G;} \\ 60(\frac{R-G}{\delta}+4) & \text{if MAX = B;} \\ notdefined & \text{if MAX = 0.} \end{cases}$$

$$S = \begin{cases} (\frac{\delta}{MAX}) & \text{if } MAX \neq 0; \\ 0 & \text{if MAX = 0.} \end{cases}$$

$$V = MAX$$

where $\delta = (MAX - MIN)$,$MAX = max(R, G, B)$, and $MIN = min(R, G, B)$

Second Order Difference Signals

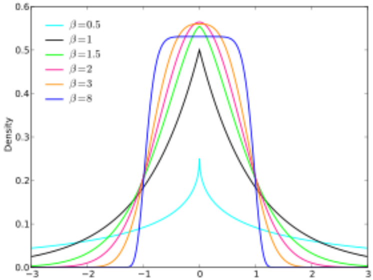
In order to capture the correlation of adjacent pixels, the second-order difference signals are used here. Let $I(i,j)$ represent the color intensity(H,S or V) at position (i,j)

- Horizontal Direction: $D_h = 2I(i,j) - I(i,j-1), I(i,j+1)$
- Vertical Direction: $D_d = 2I(i,j) - I(i-1,j), I(i+1,j)$
- Diagonal Direction: $D_d = 2I(i,j) - I(i-1,j-1), I(i+1,j+1)$
- Minor Diagonal Direction: $D_m = 2I(i,j) - I(i-1,j+1), I(i+1,j-1)$

Statistical Model: Generalized Gaussian Distribution

$$f(x;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} exp(-|\frac{x}{\alpha}|^{\beta})$$

$$\alpha = \sigma\sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)}}\ \sigma \geq 0$$



Hypothesis Test

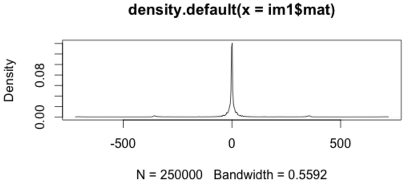
$$log \frac{P(X|H_1)}{P(X|H_0)} = \sum_{0}^{N} (log \beta_1 - log \beta_2 - log(2\alpha_1 \Gamma(1/beta_1)$$

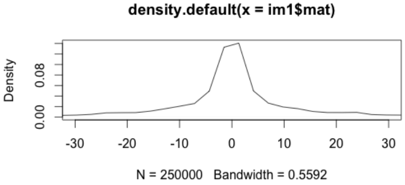
$$-log(2\alpha_2 \Gamma(1/beta_2)) - |\frac{x_i}{\alpha_1}|^{\beta_1})$$

- Need to estimate shape parameter β and α

Shape Parameter Estimation

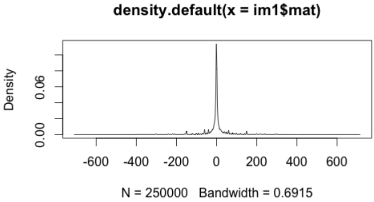
Shape Parameter Estimation:Histogram of Leaf.jpg

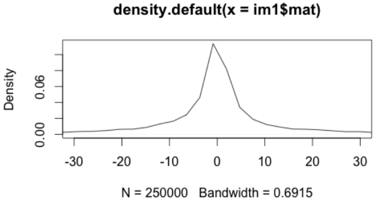




Shape Parameter Estimation

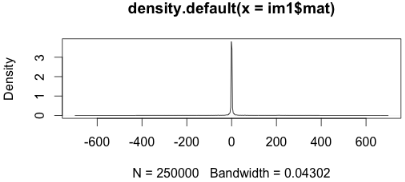
Shape Parameter Estimation:Histogram of LakeBryan.png

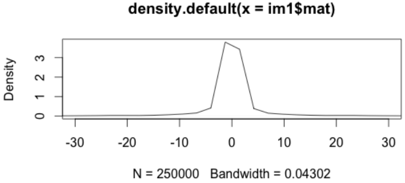




Shape Parameter Estimation

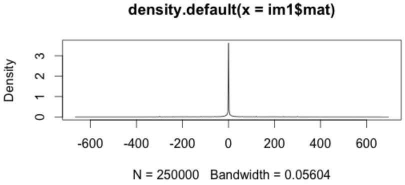
Shape Parameter Estimation:Histogram of PrinceTennis.jpg

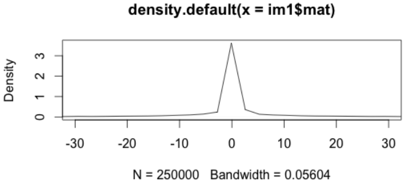




Shape Parameter Estimation

Shape Parameter Estimation:Histogram of tree.jpg





Decision Rule

- Shape Parameter β: Computer Graphs $\geq$ Real Photos
- Variance σ: Computer Graphs $\leq$ Real Photos
- By training data, the distributions for two hypothesis are:
  - H0: $\beta = 1$,$\alpha 1 = 2.12$
  - H1: $\beta 2 = 2$,$\alpha 2 = 4.24$
- Empirically speaking, we found 30,000 can be a good threshold τ
- We can get the decision rule by solving the following problem

$$log(\beta_1) - log(\beta_2) - log(2\alpha_2\Gamma(1/\beta_1)) + log(2\alpha_1\Gamma(1/\beta_2))$$
$$-|\frac{x_i}{\alpha_1}|^{\beta_1} + |\frac{x_i}{\alpha_2}|^{\beta_2}$$

- H1: $x \geq 700$
- H0: $x \leq 700$

Tests on 10 Images

Image	True	Test
L1	Real	Real
L2	Real	Synthetic
L3	Real	Real
L4	Real	Real
L5	Real	Real
L6	Real	Real
L7	Synthetic	Synthetic
L8	Synthetic	Real
L9	Synthetic	Synthetic
L10	Synthetic	Synthetic

Conclusions and Discussion

- How to enhance the performance
  - Using more training dataset for Parameter Estimation
  - It's possible for synthetic images; there's some other more suitable distribution(other than GGD)