Power law in Natural Images: sources and applications

April 15, 2016

Weibo Gong

University of Massachusetts, Amherst

Motivations and outline

- Natural images exhibit power law cluster size distributions and self-similarity;
- A possible cause could be that the vision systems evolved to a critical point in order to take advantage of the self-similarity for pattern searching;
- Percolation theory could provide mathematical analysis tools for the phenomena;
- Image data could provide a huge resource for rare event data to facilitate the development of tools for the challenges in the NRC report on Frontiers in Massive Data Analysis (2013).

Source 1: NRC report on Frontiers in Massive Data Analysis (2013)

Committee on the Analysis of Massive Data
Committee on Applied and Theoretical Statistics
Board on Mathematical Sciences and Their Applications
Division on Engineering and Physical Sciences

National Research Council, 2013

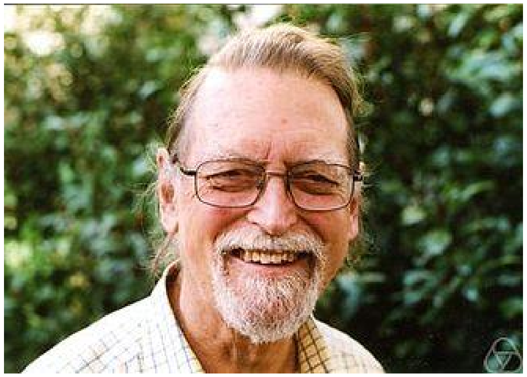
- There are many sources of potential error in massive data analysis, many of which are due to the interest in "long tails" that often accompany the collection of massive data. Events in the "long tail" may be vanishingly rare even in a massive data set.

- For example, in consumer-facing information technology, where the goal is increasingly that of providing fine-grained, personalized services, there may be little data available for many individuals even in very large data sets.

- **In science, the goal is often that of finding unusual or rare phenomena, and evidence for such phenomena may be weak, particularly when one considers the increase in error rates associated with searching over large classes of hypotheses**

- In general, data analysis is based on assumptions, and the assumptions underlying many classical data analysis methods are likely to be broken in massive data sets.

Source 2: David Mumford (1974 Fields Medalist): Self-similarity of image statistics and image models (2010)



- One of the earliest discoveries about the statistics of images was that their power spectra tend to obey power laws. This has a very provocative interpretation: this power law is implied by self-similarity!

- The hypothesis that natural images of the world, treated as a single large database, have renormalization invariant statistics has received remarkable confirmation from many quite distinct tests.

Source 3: Ken Wilson - fluctuations were happening on all scales at once

- One difficult problem was phase transitions, the passage from water to steam or atoms lining up to make a magnet. At the critical point — the temperature at which the change happens — orderly behavior breaks down, but theorists had few clues to how to calculate what was happening.

- Dr. Wilson realized that the key to the problem was that fluctuations were happening on all scales at once — from the jostling and zooming of individual atoms to the oscillations of the entire system — something conventional theory could not handle.

Sources 4: Origins of Scaling in Natural Images Daniel Ruderman,1997.

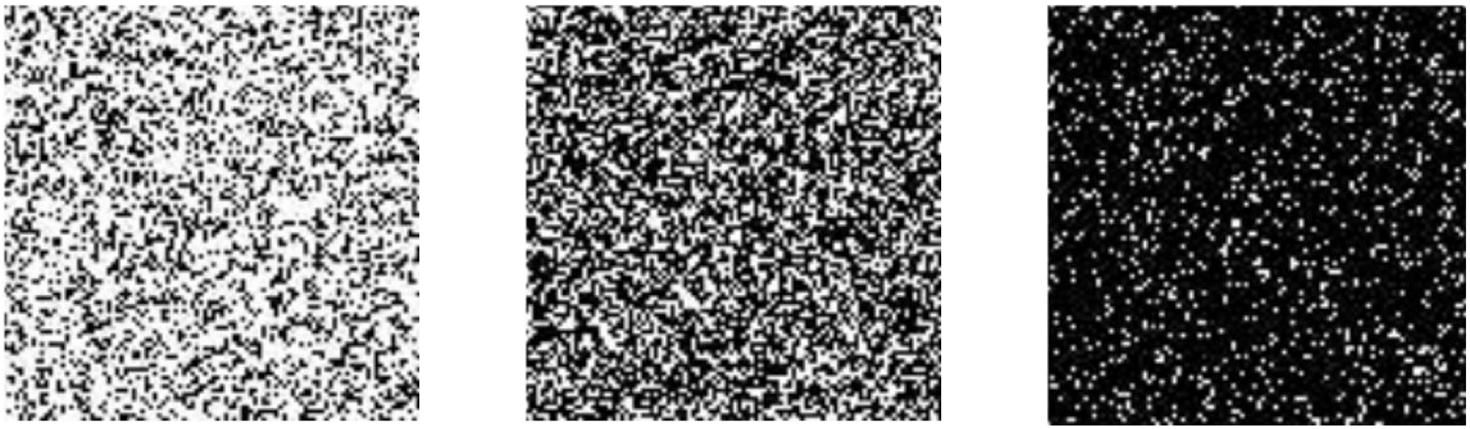
One of the most robust qualities of our visual world is the scale invariance of natural images. Not only has scaling been found in different visual environments, but the phenomenon also appears to be calibration-independent. This paper proposes a simple property of natural images which explains this robustness: they are collages of regions corresponding to statistically independent "objects". Evidence is provided for these objects having a power-law distribution of sizes within images, from which follows scaling in natural images.

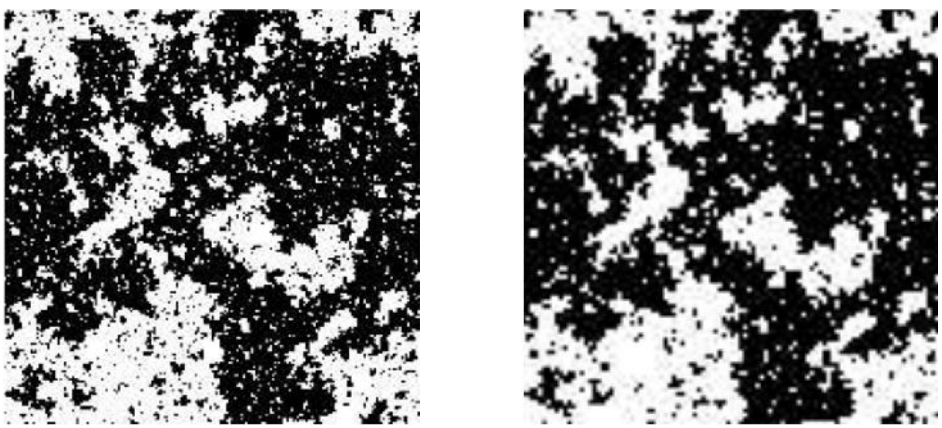
Hope: The causes behind the rare data may not be rare

- Generative models could be useful. For example when confronting 2D tails we might base our analysis of the observation data on asymptotic independence;
- Other uses of the generative models…;
- Natural images provide a data base for such investigations;
- Self-similarity of patterns in visual images seen in blurring transformations is useful in developing generative models;
- Percolation theory could help understanding the self-similarity/power law cluster size distribution in images.

Effect of blurring an image

Effect of coarse graining in percolation theory



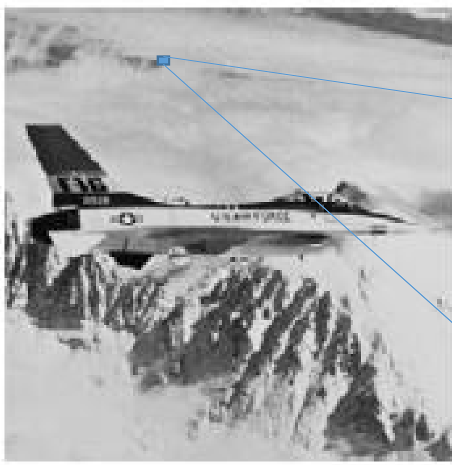


The blurring transformation preserves most of the large-scale structure of the configuration, although a lot of the small detail is lost.

“Zipf’s Law” in natural image

Zipf’s Law: Word frequency in text

? frequency in image



Matrix [g(x,y)]

(Grey level from 0~255)

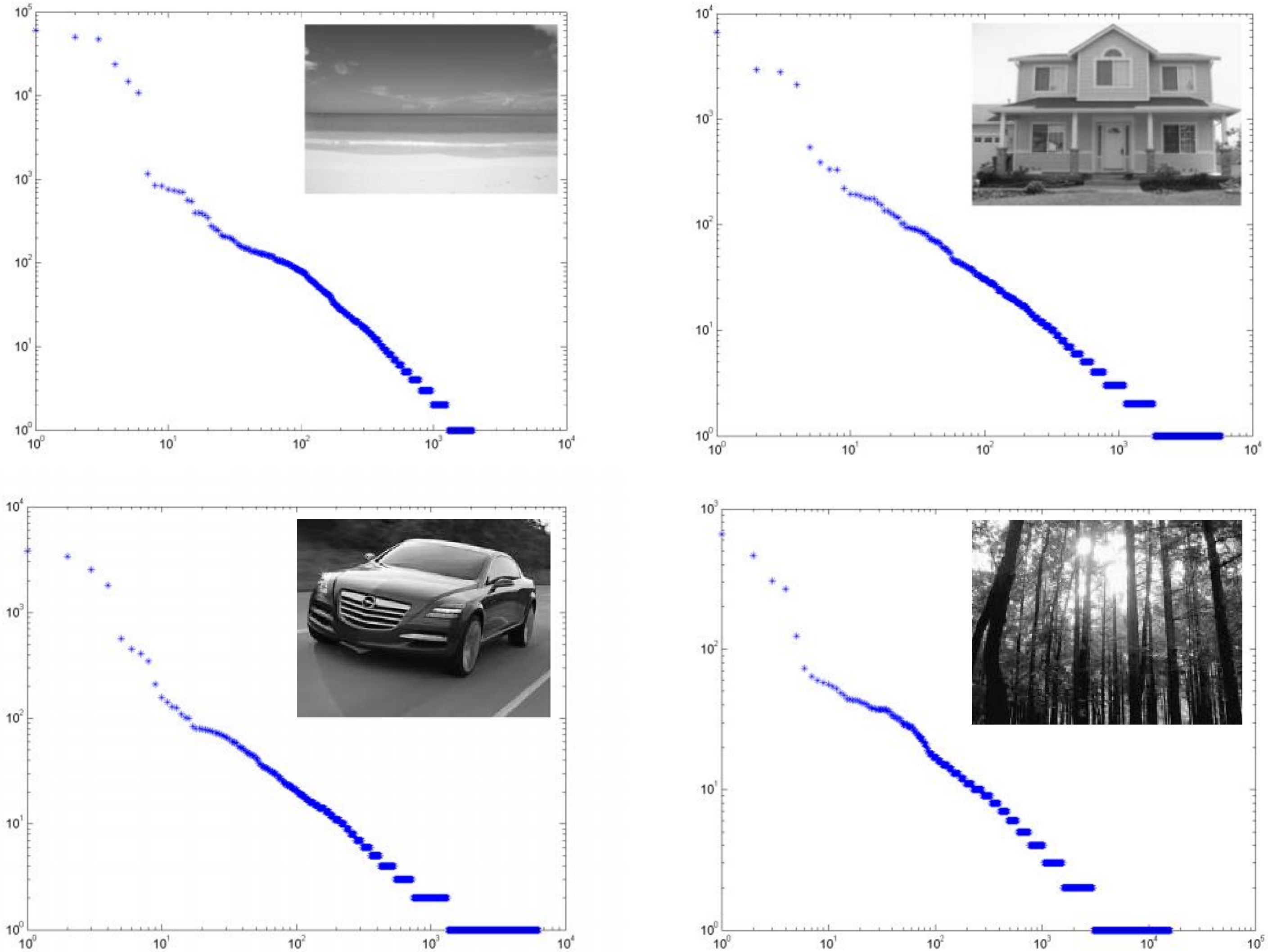
“Word” in Image

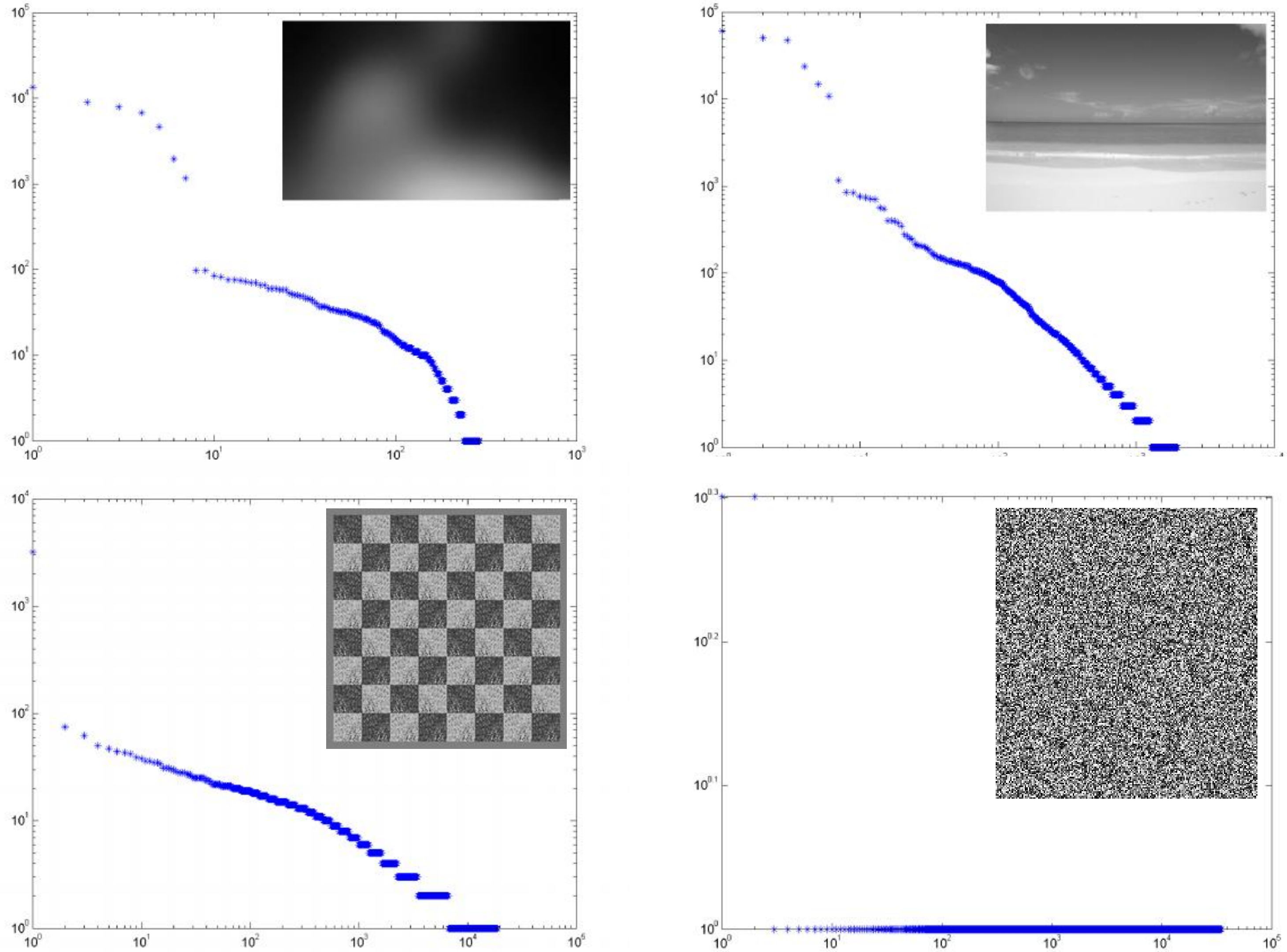
255	210	210
25	2	34
40	2	40

8	7	7
0	0	1
1	0	1

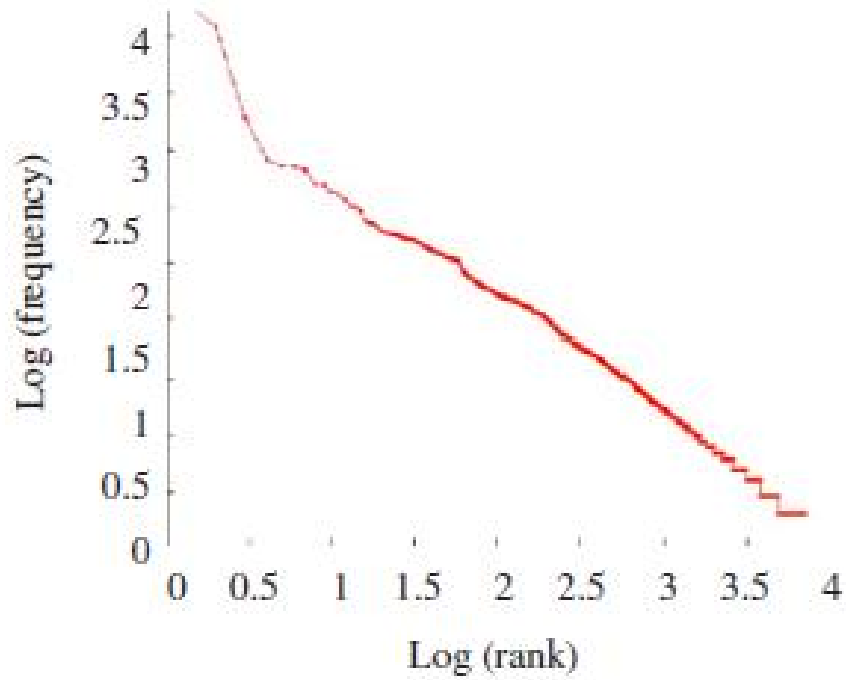
$$\text{int}\left[\frac{8 \times g(x,y)}{255}\right] = c(x,y)$$

Compressed image examples



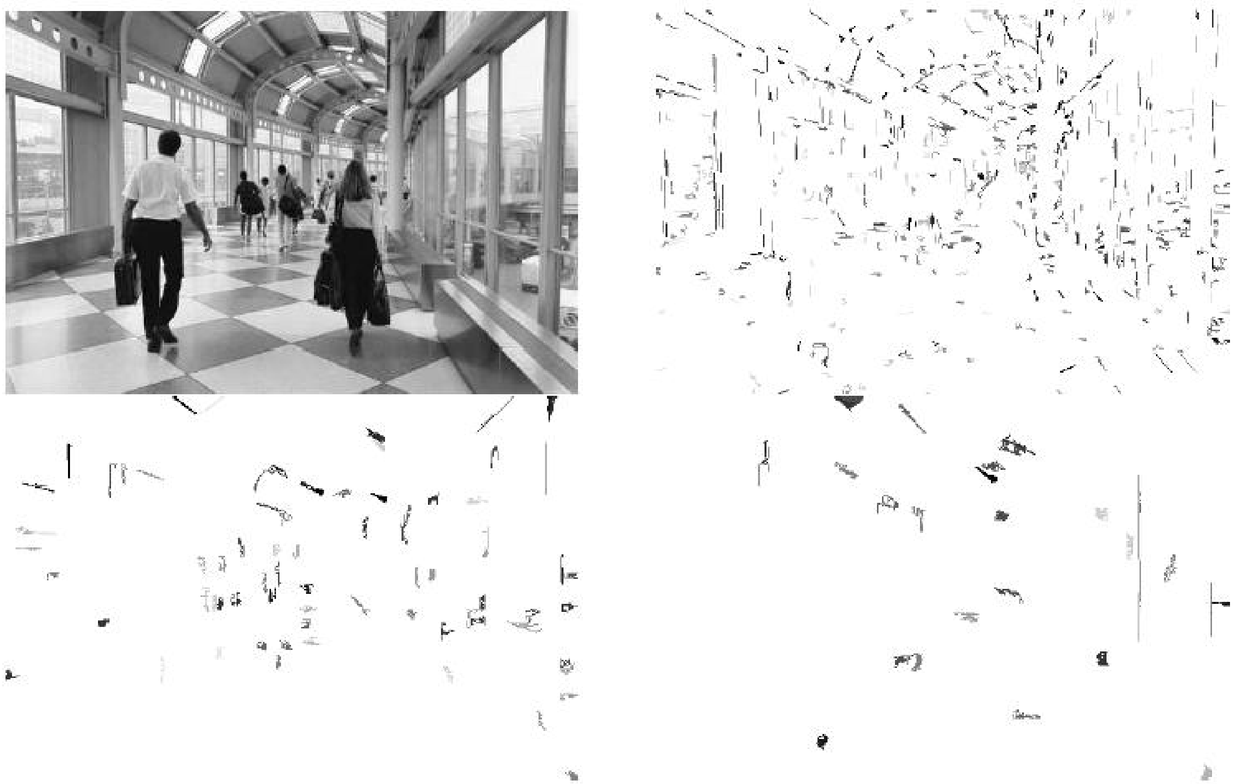

10^5 10^4 10^3 10^2 10^1 10^0
10^0 10^1 10^2 10^3
10^5 10^4 10^3 10^2 10^1 10^0
10^4 10^3 10^2 10^1 10^0
10^0 10^1 10^2 10^3 10^4 10^5
$10^{0.3}$ $10^{0.2}$ $10^{0.1}$ 10^0
10^0 10^1 10^2 10^3 10^4 10^5

Observed power law

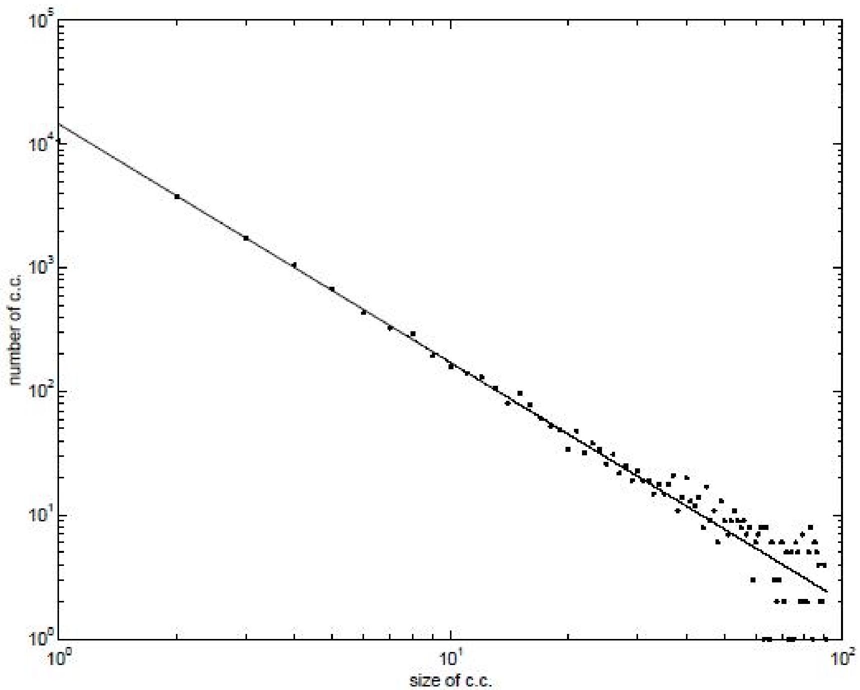


The distribution of bilevels in natural images

$$I_l(i,j) = \begin{cases} 1 & \text{if } I(i,j) \in \left[(l-1)\frac{N}{k}, l\frac{N}{k}\right), \\ 0 & \text{otherwise,} \end{cases}$$



The distribution of bilevels in natural images



$$f(a) \approx \frac{C}{a^{\alpha}}$$

Generating pow-law graph for natural images

1. Generate super-pixel segmentation for the image. (omit if the original image is not very large, this process is not needed for small-scale images)
2. Cluster the super-pixels (original pixels) based on K-nearest-neighbor algorithm.
3. Connect any pair of pixels by letting similar pairs having larger weights.
4. Merge pixels together according to the clustering result, form a power-law graph.

1. Super-pixel generation

- If the original image is large, we first do a super-pixel segmentation. Super-pixel algorithms group original image pixels into perceptually meaningful atomic regions.
- SLIC (Simple linear iterative clustering) super-pixel (Achanta et al. 2011) is a very fast super-pixel segmentation algorithm.
- We implement an algorithm similar to the SLIC super-pixel generating algorithm.
- Image is transformed to CIELAB form (In CIELAB format, L represents the intensity of the pixel, a and b are coefficients representing the color of the pixel). A distance measurement between two pixels is defined as (m is a coefficient, S is the size of the super-pixel).

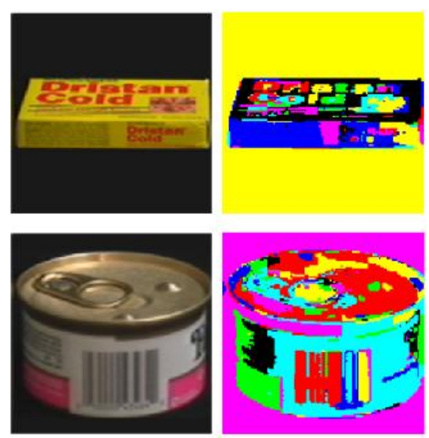
$$d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}$$
$$d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$
$$D = d_{lab} + \frac{m}{S} d_{xy}$$

Algorithm 1 Efficient superpixel segmentation

1: Initialize cluster centers $C_k = [l_k, a_k, b_k, x_k, y_k]^T$ by sampling pixels at regular grid steps S.
2: Perturb cluster centers in an $n \times n$ neighborhood, to the lowest gradient position.
3: **repeat**
4:    **for** each cluster center C_k **do**
5:       Assign the best matching pixels from a $2S \times 2S$ square neighborhood around the cluster center according to the distance measure.
6:    **end for**
7:    Compute new cluster centers and residual error E {$L1$ distance between previous centers and recomputed centers}
8: **until** $E <$ threshold

2. Clustering the super-pixels (original pixels)

- The second step is to cluster the super-pixels into larger components.
- For each super-pixel, we compute its average [L, a, b, x, y] values.
- If the image is not very large (less than 150*150), we can directly cluster the original pixels together.
- The clustering algorithm is a K-nearest-neighbor algorithm. The pixels are clustered according to their color information. We then segment the image based on the connectivity of the pixel clustering labels. (Pixels with the same label which are connected directly (or by other same-label pixels) are segmented into the same area.)



3. Generating the power-law graph

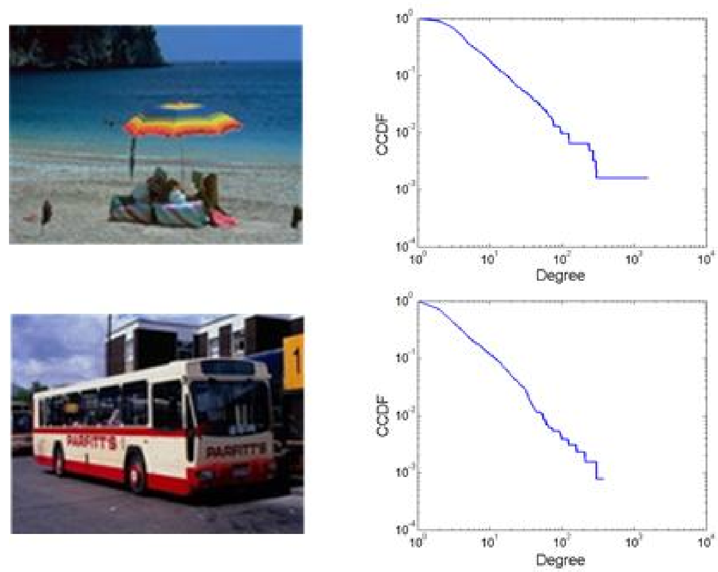
- The weight between two pixels is defined as

$$w_{ij} = e^{-d_c^2/\sigma_I} \cdot e^{-ds^2/\sigma_X}$$

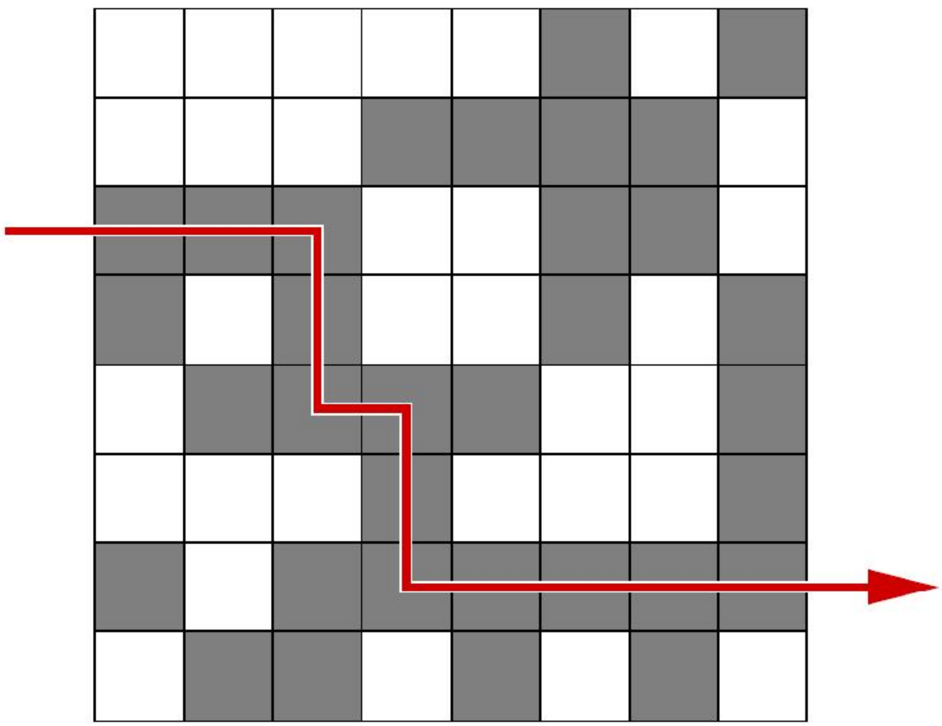
- Based on the clustering result, the weight of the edge is computed by

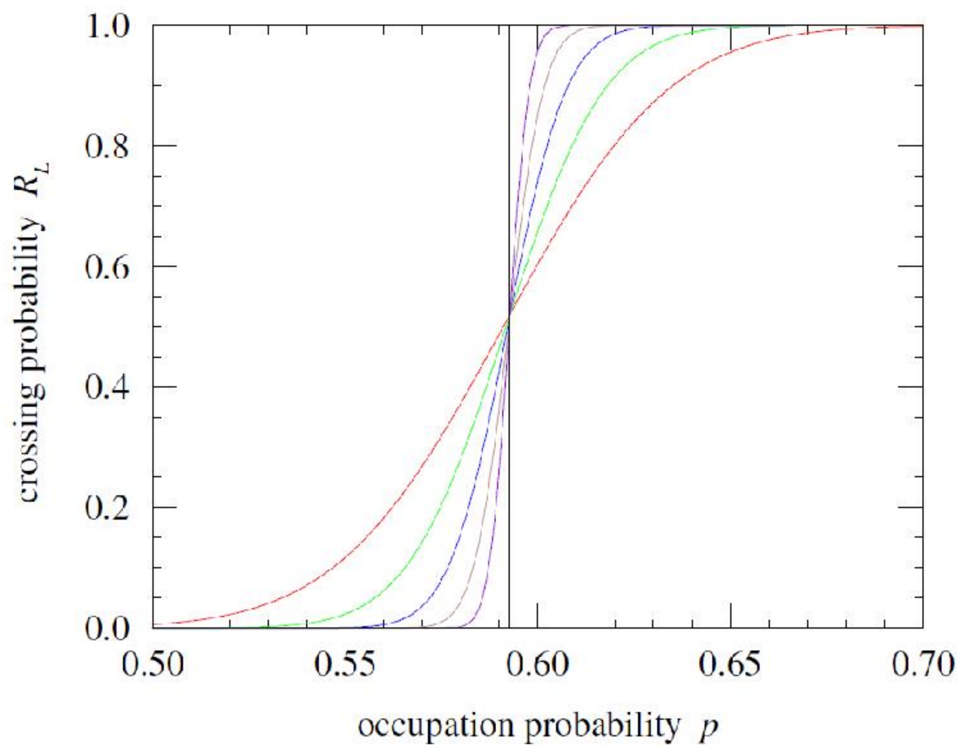
$$A(m,n) = \sum_{p_i \in cc_m} \sum_{p_j \in cc_n} w_{ij}.$$

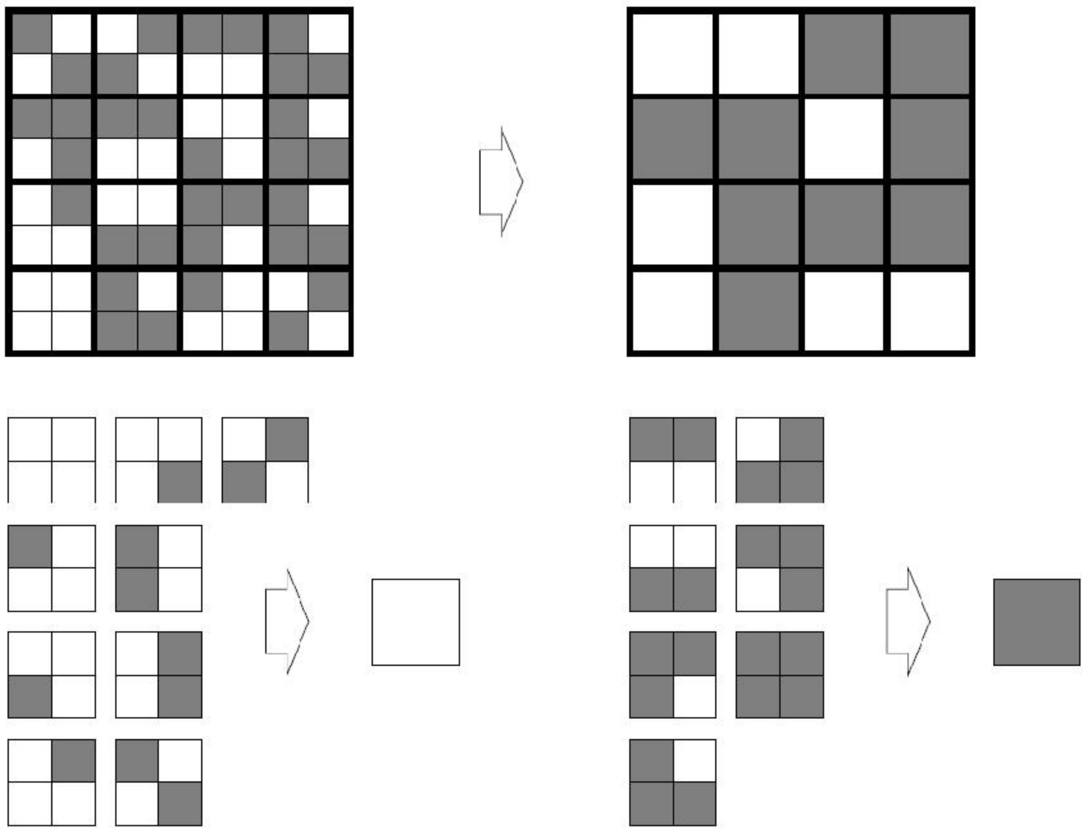
- For natural image, the generated graph has power-law property.



Percolation calculation for power law/self-similarity phenomena at the critical point







Derivation of size distribution power law at critical point

Important stuff: Consider the following scaling argument. If we change the *scale* on which we measure areas on our lattice by a factor of b, then all clusters change size according to $s \to bs$. Of course, the physics of the system hasn't changed, only how we measure it, so this change of variables cannot change the distribution n_s, except by a numerical factor to keep the normalization correct.

The argument of $f(x)$ doesn't change anyway, because s and $\langle s \rangle$ both change by the same factor b. But the argument of $g(x)$ does change. Thus $g(x)$ must satisfy

$$g(bx) = k(b)g(x), \tag{2}$$

where $k(b)$ is the numerical factor, which can depend on b but not x. Let us choose the normalization of g so that $g(1) = 1$. Then, setting $x = 1$ above we have

$$g(b) = k(b) \tag{3}$$

for all b and hence $k(x)$ and $g(x)$ are the same function. Thus

$$g(xy) = g(x)g(y). \tag{4}$$

To solve this equation, we take the derivative with respect to y:

$$\frac{\partial}{\partial y} g(xy) = xg'(xy) = g(x)g'(y), \tag{5}$$

then set $y = 1$ to get

$$xg'(x) = g(x)g'(1), \tag{6}$$

whose solution is

$$\log g(x) = g'(1) \log x + c, \tag{7}$$

where c is an integration constant. Given $g(1) = 1$, we must have $c = 0$, and hence

$$g(x) = x^{-\tau}, \tag{8}$$

where $\tau = -g'(1)$. This functional form is called a **power law**. The quantity τ is a **critical exponent**.

> The distribution of cluster sizes becomes a power law exactly at the critical point. Indeed, the same arguments imply that all distributions will become power laws at the critical point. This is one of the characteristic features of phase transitions.

Conclusions

- Natural images exhibit power law cluster size distributions and self-similarity;
- A possible cause could be that the vision systems evolved to a critical point in order to take advantage of the self-similarity for pattern searching;
- Percolation theory could provide mathematical analysis tools for the phenomena;
- Image data could provide a huge resource for rare event data to facilitate the development of tools for the challenges in the NRC report on Frontiers in Massive Data Analysis (2013).