How to know what you don't see

Quantifying visual information

Alex Striff
Advisor: Vincent Meunier

Reed College

July 8, 2020

Outline

What are we after?

Information and entropy

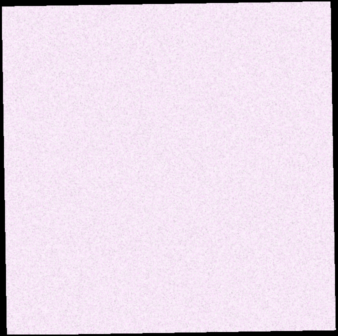
Image fluctuations

Wang-Landau algorithm results

Other approaches taken

Next steps

What do you see?



Information from a random variable X being $x \in \mathcal{X}$?

Shannon, C. E. A mathematical theory of communication. The Bell system technical journal **27,** 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x (1948)

Information from a random variable X being $x \in \mathcal{X}$?

- If $P(x) = 1$, $I(x) = 0$
- If $P(x) < P(x')$, $I(x) > I(x')$
- For independent events x and y, $I(x, y) = I(x) + I(y)$

Shannon, C. E. A mathematical theory of communication. The Bell system technical journal **27**, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x (1948)

Information from a random variable X being $x \in \mathcal{X}$?

- If $P(x) = 1$, $I(x) = 0$
- If $P(x) < P(x')$, $I(x) > I(x')$
- For independent events x and y,
  $I(x, y) = I(x) + I(y)$

$$I(x) = -\log P(x)$$

Shannon, C. E. A mathematical theory of communication. The Bell system technical journal **27**, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x (1948)

Information from a random variable X being $x \in \mathcal{X}$?

- If $P(x) = 1$, $I(x) = 0$
- If $P(x) < P(x')$, $I(x) > I(x')$
- For independent events x and y, $I(x, y) = I(x) + I(y)$

$$I(x) = -\log P(x)$$

$$\begin{aligned} S(X) &= \langle I \rangle_X \\ &= -\sum_{x \in \mathcal{X}} P(x) \log P(x) \end{aligned}$$

Shannon, C. E. A mathematical theory of communication. The Bell system technical journal **27**, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x (1948)

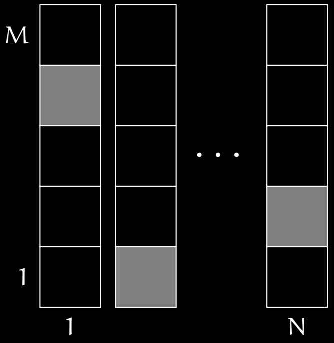
The maximum entropy principle

How to assign probabilities $p(x)$ to a random variable X with given $\langle \mathbf{f}(x) \rangle$ in an unbiased way?

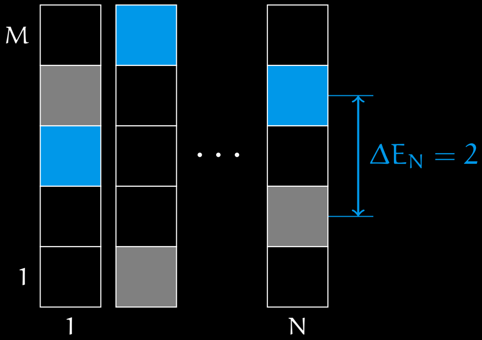
- Maximize the entropy subject to moment constraints
- For mean energy, this gives the canonical distribution
- Subjective and objective probabilities

Jaynes, E. T. Information Theory and Statistical Mechanics. Phys. Rev. **106**, 620–630. doi:10.1103/PhysRev.106.620 (4 May 1957)

Images as discrete-level systems

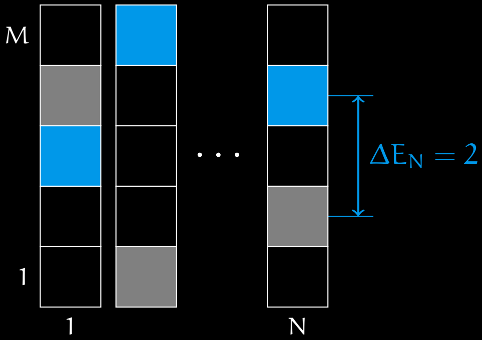


Images as discrete-level systems

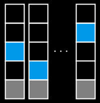


$$E = \sum_i \Delta E_i$$

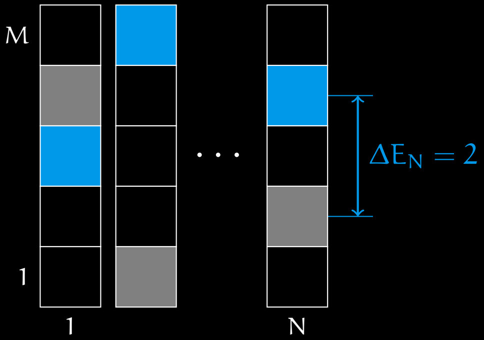
Images as discrete-level systems



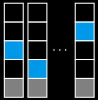
$$E = \sum_i \Delta E_i$$



Images as discrete-level systems

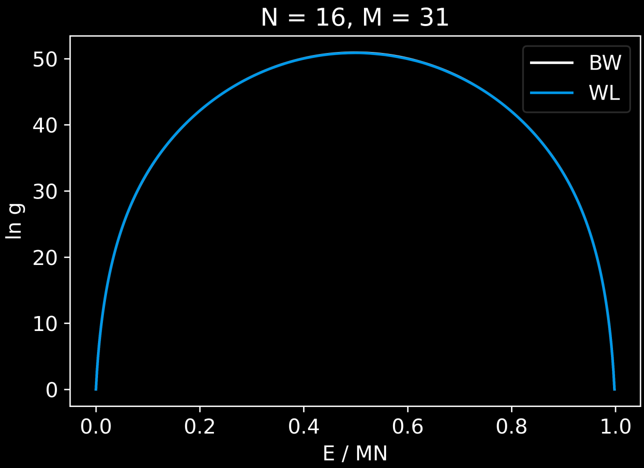


$$E = \sum_i \Delta E_i$$



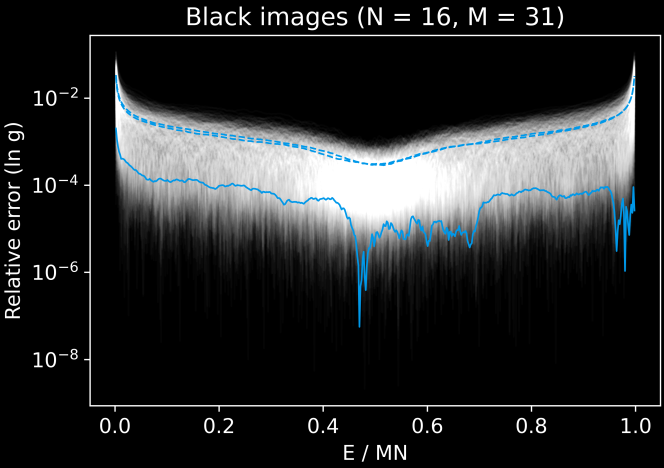
$$g(E) = \sum_k (-1)^k \binom{N}{k} \binom{N + E - kM - 1}{E - kM}$$

Wang-Landau algorithm results



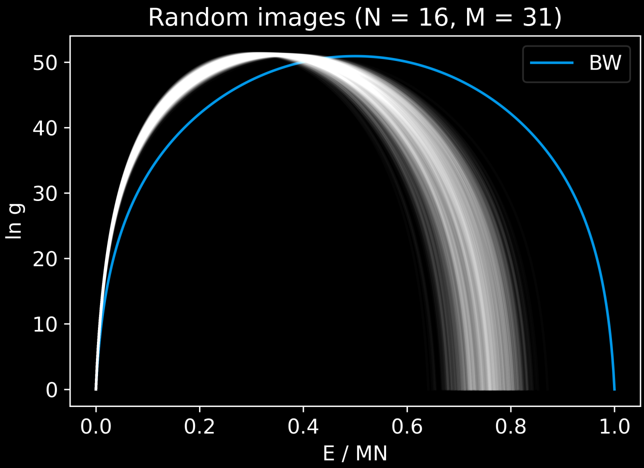
BW: Black/white (exact)   WL: Wang-Landau result

Wang-Landau algorithm results



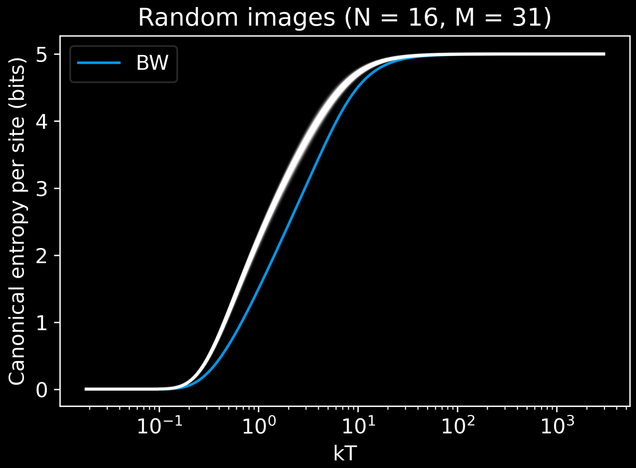
Black image simulations   Mean $\pm$ standard deviation

Wang-Landau algorithm results



Random gray image simulations BW: Black/white (exact)

Wang-Landau algorithm results



Random gray image simulations BW: Black/white (exact)

Other approaches taken

- All light entering the eye (receptive fields)
- Entropy for continuous coordinates (KL divergences)

Next steps

- Include spatial information (infer from neighbors)
- Include color (CIE Lab)
- Information as how to draw the image (Kolmogorov complexity)

Acknowledgements

Questions?

Appendix

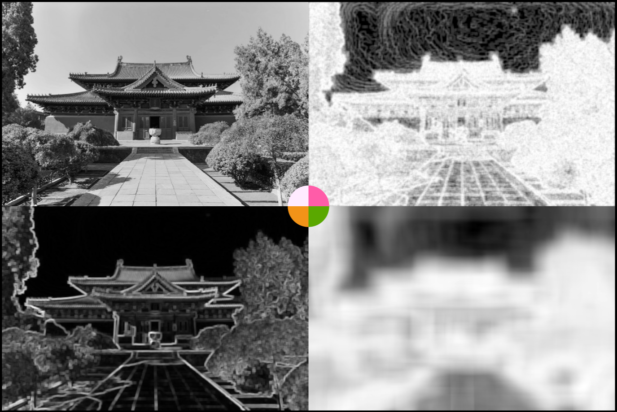
Intensity entropy

Inference and the maximum entropy method

Information dimension

References

Intensity entropy



Original 5×5 IE 5×5 SD 41×41 IE

IE: Intensity entropy

SD: Standard deviation

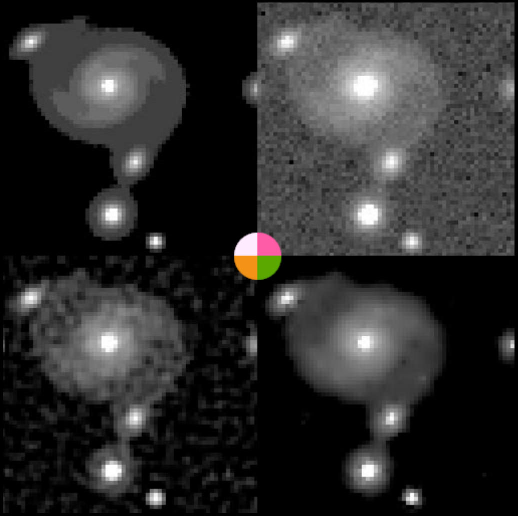
Inference

- A noisy measurement of an image (I) produces data (D).

Bayes' rule

$$P(I \mid D) = \frac{P(D \mid I)P(I)}{P(D)}$$

- We infer the posterior from the likelihood and prior, and normalize by the evidence.
- MAP estimate: maximize $\ln P(D \mid I) + \ln P(I)$
- Maximum entropy method: $P(I) = e^{\lambda S(I)}$



Original “Measured” MEM Multiscale MEM

Pantin, E. & Starck, J.-L. Deconvolution of astronomical images using the multiscale maximum entropy method. Astron. Astrophys. Suppl. Ser. **118,** 575–585. doi:10.1051/aas:1996221 (1996)

How to assign an information dimension to the random variable X?

$$X_m = \frac{\lfloor mX \rfloor}{m}$$

$$d(X) = \lim_{m \to \infty} \frac{S(X_m)}{\log m}$$

- Lebesgue decomposition theorem: $d(X)$ is the fraction of P_X that is discrete.
- For a n-vector X with finite $H(\lfloor X \rfloor)$, $0 \leqslant d(X) \leqslant n$.

References

- Shannon, C. E. A mathematical theory of communication. The Bell system technical journal **27,** 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x (1948).
- Jaynes, E. T. Information Theory and Statistical Mechanics. Phys. Rev. **106,** 620–630. doi:10.1103/PhysRev.106.620 (4 May 1957).
- Pantin, E. & Starck, J.-L. Deconvolution of astronomical images using the multiscale maximum entropy method. Astron. Astrophys. Suppl. Ser. **118,** 575–585. doi:10.1051/aas:1996221 (1996).
- Wang, F. & Landau, D. P. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. Phys. Rev. Lett. **86,** 2050–2053. doi:10.1103/PhysRevLett.86.2050 (10 Mar. 2001).