# CS231A Course Project Proposal

Niru Maheswaranathan
Stanford University
nirum@stanford.edu

## Abstract

*The space of natural images is both high-dimensional and has striking non-Gaussian, higher-order statistical structure [ref?]. It has been argued on information theoretic grounds that much of early visual processing consists of removing statistical regularities in visual input, as reducing the entropy maximizes information transmission. Therefore, a key part of understanding the redundancy of natural scenes is understanding the entropy of image patches. Entropy allows us to place a number on the information content of image patches, and bounds optimal compression schemes. This project aims to compare different methods for estimating the entropy of natural image patches, and to compare how these estimates scale with increasing patch sizes*



## 1. Problem statement

It has long been argued [cite:Barlow] that the human visual system has evolved to more efficiently process the statistical structure of natural images. This suggests that efficient computer vision systems can also be designed given this statistical structure. One example of where knowledge of the statistical structure of natural images would be useful is in redundancy reduction, as information is maximized when redundancies in the data are removed.

The difficulty in estimating the entropy of image patches arises due to the curse of dimensionality: for patches of size $n \times n$, we must estimate a probability distribution with dimension $n^2$, which quickly becomes infeasible as $n$ grows. Despite these limitations, it is possible to obtain estimates of the entropy of high dimensional probability distributions. This project will compare different methods for making these estimates for increasing image patch sizes.

## 2. Background

A few previous studies have developed methods for estimating the entropy of image patches.

Chandler and Field use an approximation that relies on nearest-neighbor (NN) distances. It has been shown [Kozachenko and Leonenko 1987] that the average log NN distance can be used to estimate the entropy without estimating the probability distribution $p(x)$.

One approach to place an upper bound on the entropy involves fitting maximum entropy models to natural images. These models are guaranteed to have the largest entropy given certain statistics observed from data. This was the approach taken by [Bethge and Berens], who fit an approximate maximum entropy model and used it to estimate the entropy of image patches.

Proximity distribution estimates, using nearest neighbor techniques (Chandler and Field) DG near-maxent model fitting (Bethge and Berens) Gaussian Scale Mixtures (Hosseini, Sinz, and Bethge)

## 3. Methods

First, I will generate samples of image patches that will be used to fit model parameters. In addition to natural images, patches drawn from distributions with known probability distributions will be used to benchmark the entropy estimation algorithms listed below. The fixed distributions used will be Gaussians with either diagonal covariance structure (white noise), or fixed covariance structure, where the covariance estimated from natural images (colored noise). The entropy of these samples is known given the formula for the entropy of a multivariate Gaussian. For natural images, samples of patches of different sizes will be drawn from images from the van Hateren natural image database [cite].

The entropy of image patches drawn from these distributions will be computed using two methods discussed in the previous literature. The first is involves using nearest-neighbor methods [Chandler and Field] and the second involves approximate maximum entropy models [Bethge and Berens]. Entropy estimates will be computed both for (1) increasing sample sizes and (2) increasing dimensionality (patch sizes). One novel contribution of this work will be to compare and contrast the two methods. Currently, it is not clear how the reliability of the entropy estimates will scale as the dimensionality of the patches increases, but I will be able to say something about this given my method.

A secondary aim that I am thinking about pursuing involves performing dimensionality reduction (for example, with PCA) on the set of natural iamge patches, and then computing the entropy of the distributions in the low-dimensional space. It is not clear how entropy will be affected by different dimensionality reduction techniques, and by the size of the low-dimensional space. The above methods of estimating entropy will yield more reliable estimates of the entropy given low-dimensional representations, and will allow me to extrapolate the entropy estimate to larger image patch sizes.

## 4. Evaluation

The ultimate goal of this project is to obtain plots of the entropy (in bits per pixel) of image patches for different patch sizes.

To validate the methods, a comparison of how the various metrics perform when estimating image patches drawn from known distribuitons with fixed statistical structure (white/colored noise) will be used as a baseline for comparison, since the entropy of these distributions is known analytically.

Once the methods have been validated, entropy estimates of the different measures on natural image patches will be compared. Different curves will be shown for measures of the entropy made using different methods. Another criterion for evaluation is comparing the consistency of different entropy estimates for increasing sample sizes.

Finally, these results will be compared with recent entropy estimates from the literature [refs].

## 5. Criteria

What is the computer vision problem that you will be investigating? Why is it interesting?

What image or video data will you use? If you are collecting new datasets, how do you plan to collect them?

What method or algorithm are you proposing? If there are existing implementations, will you use them and how? How do you plan to improve or modify such implementations?

Which reading will you examine to provide context and background?

How will you evaluate your results? Qualitatively, what kind of results do you expect (e.g. plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results (e.g. what performance metrics or statistical tests)?

## 6. Introduction

Please follow the steps outlined below.

### 6.1. Language

All manuscripts must be in English.

### 6.2. The ruler

The LaTeX style defines a printed ruler which should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. If you are preparing a document using a non-LaTeX document preparation system, please arrange for an equivalent ruler to appear on the final output pages. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler.

### 6.3. Mathematics

Please number all of your sections and displayed equations. It is important for readers to be able to refer to any particular equation. Just because you didn't refer to it in the text doesn't mean some future reader might not need to refer to it. It is cumbersome to have to use circumlocutions like "the equation second from the top of page 3 column 1". (Note that the ruler will not be present in the final copy, so is not an alternative to equation numbers). All authors will benefit from reading Mermin's description of how to write mathematics.

### 6.4. Miscellaneous

Compare the following:
$conf_a$
$\mathit{conf}_a$
See The TeXbook, p165.

The space after *e.g.*, meaning "for example", should not be a sentence-ending space. So *e.g.* is correct, *e.g.* is not. The provided macro takes care of this.

When citing a multi-author paper, you may save space by using "et alia", shortened to "*et al.*" (not "*et. al.*" as "*et*" is a complete word.) However, use it only when there are three or more authors. Thus, the following is correct: " Frobnication has been trendy lately. It was introduced by Alpher [**?**], and subsequently developed by Alpher and Fotheringham-Smythe [**?**], and Alpher *et al.* [**?**]."

This is incorrect: "... subsequently developed by Alpher *et al.* [**?**] ..." because reference [**?**] has just two authors. If you use the macro provided, then you need not worry about double periods when used at the end of a sentence as in Alpher *et al.*

For this citation style, keep multiple citations in numerical (not chronological) order, so prefer [**?**, **?**, **?**] to [**?**, **?**, **?**].

## 7. Formatting your paper

All text must be in a two-column format. The total allowable width of the text area is $6\frac{7}{8}$ inches (17.5 cm) wide by $8\frac{7}{8}$ inches (22.54 cm) high. Columns are to be $3\frac{1}{4}$ inches (8.25 cm) wide, with a $\frac{5}{16}$ inch (0.8 cm) space between them. The main title (on the first page) should begin 1.0 inch (2.54 cm) from the top edge of the page. The second and following pages should begin 1.0 inch (2.54 cm) from the top edge. On all pages, the bottom margin should be 1-1/8 inches (2.86 cm) from the bottom edge of the page for $8.5 \times 11$-inch paper; for A4 paper, approximately 1-5/8 inches (4.13 cm) from the bottom edge of the page.

### 7.1. Margins and page numbering

All printed material, including text, illustrations, and charts, must be kept within a print area 6-7/8 inches (17.5 cm) wide by 8-7/8 inches (22.54 cm) high.

### 7.2. Type-style and fonts

Wherever Times is specified, Times Roman may also be used. If neither is available on your word processor, please use the font closest in appearance to Times to which you have access.

MAIN TITLE. Center the title 1-3/8 inches (3.49 cm) from the top edge of the first page. The title should be in Times 14-point, boldface type. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave two blank lines after the title.

Figure 1. Example of a short caption, which should be centered.

AUTHOR NAME(s) and AFFILIATION(s) are to be centered beneath the title and printed in Times 12-point, non-boldface type. This information is to be followed by two blank lines.

The ABSTRACT and MAIN TEXT are to be in a two-column format.

MAIN TEXT. Type main text in 10-point Times, single-spaced. Do NOT use double-spacing. All paragraphs should be indented 1 pica (approx. 1/6 inch or 0.422 cm). Make sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

Figure and table captions should be 9-point Roman type as in Figures **??** and **??**. Short captions should be centred.

Callouts should be 9-point Helvetica, non-boldface type. Initially capitalize only the first word of section titles and first-, second-, and third-order headings.

FIRST-ORDER HEADINGS. (For example, **1. Introduction**) should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one blank line after.

SECOND-ORDER HEADINGS. (For example, **1.1. Database elements**) should be Times 11-point boldface, initially capitalized, flush left, with one blank line before, and one after. If you require a third-order heading (we discourage it), use 10-point Times, boldface, initially capitalized, flush left, preceded by one blank line, followed by a period and your text on the same line.

| Method | Frobnability |
|---|---|
| Theirs | Frumpy |
| Yours | Frobbly |
| Ours | Makes one's heart Frob |

Table 1. Results. Ours is better.

### 7.3. Footnotes

Please use footnotes[1] sparingly. Indeed, try to avoid footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence). If you wish to use a footnote, place it at the bottom of the column on the page on which it is referenced. Use Times 8-point type, single-spaced.

### 7.4. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [**?**]. Where appropriate, include the name(s) of editors of referenced books.

### 7.5. Illustrations, graphs, and photographs

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths which render effectively in

[1] This is what a footnote looks like. It often distracts the reader from the main flow of the argument.

print. Many readers (and reviewers), even of an electronic copy, will choose to print your paper in order to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

When placing figures in LaTeX, it's almost always best to use , and to specify the figure width as a multiple of the line width as in the example below

### 7.6. Color

Color is valuable, and will be visible to readers of the electronic copy. However ensure that, when printed on a monochrome printer, no important information is lost by the conversion to grayscale.

## 8. Appendix

If your course project is part of a larger project from another class or research lab, please fill in this section and clearly spell out the following items:

1. Explicitly explain what the computer vision components are in this course project;

2. Explicitly list out all of your own contributions in this project in terms of:

   (a) ideas

   (b) formulations of algorithms

   (c) software and coding

   (d) designs of experiments

   (e) analysis of experiments

3. Verify and confirm that you (and your partner currently taking CS231A) are the sole author(s) of the writeup. Please provide papers, theses, or other documents related to this project so that we can compare with your own writeup.

4. Please explicitly list out:

   (a) Is this project shared with some other course? If so, what class the project is being shared with (for all members of the team)

   (b) The exact contribution of each person

   (c) The exact portion of the project that is being counted for CS231A

Failure to do the above is an honor code violation.