Image Analysis Through Local Information Measures

Neil Bruce

Department of Computer Science and Centre for Vision Research

York University, Toronto, Canada, M3J 1P3

neil@cs.yorku.ca

Abstract

The properties of local image statistics are analyzed in a classic information theoretic setting. Local spatiochromatic image elements are projected into a space in which constituent components are independent by way of independent component analysis, allowing a fast and tractable means of considering the joint likelihood of such statistics. Observation of this likelihood allows inferences to be made regarding the informativeness of a particular set of statistics. This operation is shown to illuminate a number of perceptually important image properties, allowing figure-ground segmentation, removal of common or expected image elements, and prediction of regions of interest.

1. Introduction

Information theoretic principles have proven useful in the analysis of statistics arising from a great variety of vastly different disciplines. In the context of image processing, contributions of information theory have been largely focused upon the applications of compression and coding of image statistics[1]. A natural extension of this work is the consideration of what information measures may afford in the analysis of image content, with a focus on the relationship between quantitative measures of informativeness and perceptual relevance of a set of image statistics. It is with these considerations in mind, that we propose a means of quantifying the information afforded by a local arrangement of image statistics, and the utility of such measures in the analysis of image content.

2. Representation of Local Statistics

There are many ways in which the statistics that comprise a local neighborhood of an image might be represented. Possible representations include various color spaces such as RGB, HSI, or CIE colorspaces, wavelet domain representations, or perhaps some other less commonly employed statistical framework. An important consideration in choosing this representation, is

that in drawing inferences based on local statistics, one need consider a joint probability density function with the same dimensionality as number of dimensions in the chosen representation. For example, statistical inference based on the values of a 5 by 5 image patch in RGB space would require the observation of a 75 (5x5x3) dimensional probability space in the absence of any independence assumptions. In practice, inference based on such high dimensional probability functions proves unfeasible requiring prohibitively large degrees of computation and data. It seems logical then, to choose a representation of the desired statistics that yields an independent representation. In doing so, inference is reduced to the evaluation of a number of one dimensional probability functions. Recent endeavors [2,3] have employed a representation based on independent component analysis (ICA) wherein the set of local statistics are transformed into a space in which the constituent components are as independent as possible. It has been demonstrated that independent components provide a reasonable means of representing very general non-Gaussian statistics[4] in a manner that resembles the representation employed by the human primary visual cortex[5]. With these considerations in mind, local statistics in our analysis are represented by Infomax independent components[4]. This independent representation yields a tractable means of evaluating the probability functions required for the analysis presented in section 3.

3. Quantifying Local Informativeness

As mentioned before, the goal of this work is the consideration of the utility of measuring the information afforded by a set of local statistics within a particular context. Such analysis is conducted within a classic information theoretic framework. Information theory has provided a number of quantitative means of describing the information content provided by a set of statistics. Perhaps the most well known of such information measures is Shannon's self-information measure[6] which quantifies the unexpectedness of an event. Given a probability

function p(x), self-information is given by I(x)=-log(p(x)). This transformation has the desirable properties that the informativeness is inversely proportional to the probability of the event, the function I(x) is monotonically decreasing, and the information afforded by two independent events combines additively. Informativeness in our discussion, is then given by the negative log of the likelihood that a particular event, or set of statistics occurs. In this case, the event under consideration is the joint set of basis coefficients corresponding to a local neighbourhood drawn from a natural image. Figure 1. depicts the basic framework for this analysis. The setup depicted in figure 1 is employed in two stages: Learning the basis coefficient distributions, and estimating informativeness based on the learned distributions.

3.1. Learning the basis coefficient distributions

Basis coefficient distributions as shown at the bottom of figure 1 are learned for a number of basis functions (Fig. 1. middle) based on a set of images patches (Fig. 1. top). The learning process may be described as follows:

1. A number of image patches are drawn from a particular context, which might include a single image, images of horses, or, the space of all images.
2. A set of basis functions is produced through Infomax independent component analysis[4] based on a very large set of image patches drawn from a large number of natural images.
3. Each of the initial image patches is projected into basis space by the pseudoinverse of the learned basis set.
4. Coefficients describing the contribution of each of the basis functions are obtained, and the distribution of such coefficients for each basis function is estimated using a Gaussian Parzen window density estimate.

In our implementation, distributions were learned based on two separate cases. In the first case, 360,000 7x7 image patches were randomly drawn from 3600 natural images to produce distributions representative of the statistics of all natural images. In the second case, local information content is considered in the context of a single image and the set of image patches consists of all local 7x7 neighbourhoods of the image under consideration. To distinguish between these two cases we will call the case based on general statistics Informativeness based on ecological statistics and the latter case Context-based informativeness. The local joint probabilities that describe these sets of statistics shall be referred to as p(x|X) and p(x|k) respectively, where X denotes the space of all images, k the statistics of a single

image under consideration, and x, the coefficients corresponding to the contribution of the set of basis functions in describing the local image statistics.
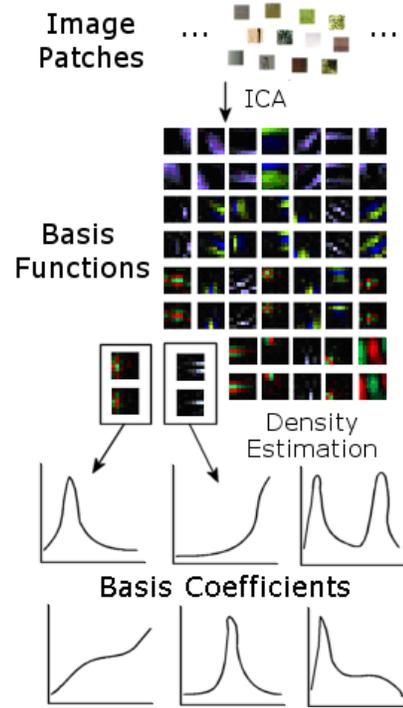
Figure 1: Framework for deriving an independent set of probability distributions that describe local signal content. 28 of the 147 basis functions are shown with positive coefficients of each basis function above their negative counterparts.

3.2. Estimating informativeness

Given a set of basis functions and a set of probability functions that describes the distribution of coefficients corresponding to each of the basis functions, it is straightforward to apply this information to an image to produce an information map as output. This is carried out as follows:

1. Each local neighbourhood of the image is projected into basis space by the pseudoinverse of the basis set.
2. The likelihood of the resulting set of coefficients is given by the product of the probabilities of each of the coefficient values, which may be looked-up from the learned basis coefficient distributions.
3. The joint probability is converted to an information measure via Shannon's self-information measure.
4. Performing this operation on each local neighbourhood gives an estimate of the relative

informativeness at each pixel location, based on local statistics.

3.3. Scale-space considerations

As is the case with any feature detection operation, the filter output is somewhat dependent on window size. In our implementation, this has been dealt with by learning the probability functions based on a 7x7 window at 4 spatial scales based on a Gaussian pyramid representation[7]. When applying the learned distributions as feature detectors, self-information is computed on each level of the Gaussian pyramid and averaged across scale. This provides a reasonable means of doing away with sensitivity to window size.

4. Discussion

As discussed, self-information of local image content is computed within two different contexts: Self-information of the local statistics is measured based on the relative contribution of basis functions to the local statistics as compared with: i. the contribution of such basis functions in the context of all natural images. ii. contribution of such basis functions in the image from which the local neighbourhood is drawn. Results of applying these operations to a number of natural images are presented in this section.

4.1. Informativeness based on ecological statistics

The information operator based on ecological statistics provides a measure of the uniqueness of local statistics in the space of all images. Figure 2. demonstrates the output of this operation when applied to a number of natural images. It is clear that in training, one will encounter patches that correspond to typically background elements such as grass or water much more frequently than those that correspond to foreground content. As such, the coefficient values corresponding to the basis functions in describing such natural texture patches will occur with greater frequency than those corresponding to foreground content. The result of applying this information measure to the images, as one might expect, is the suppression of natural textures with less usual image content producing a stronger response. It is clear that the application of the learned model to a given image frequently results in illustrating meaningful signal content. The information measure based on ecological statistics often provides a fast means of segmenting foreground from background and might be employed as a precursor to guide more involved tasks such as object recognition.

Figure 2: Information maps derived from coefficient distributions based on ecological statistics.

4.2. Context-based informativeness

The context-based information operator provides a measure of the uniqueness of local statistics in the context of a particular image. The result of applying this operator to a number of natural images is shown in figure 3. The output typically results in a strong response in regions of interest in the scene. For example, in the image second from top in figure 3. a particularly strong response results from the one individual with uniquely colored clothing in the context of the image. This operation then illuminates local regions with informative or unique content in the context of the image in question. As mentioned before, this information might be very useful in guiding such tasks as perceptually motivated compression, and quality assessment, as well as guiding higher level image analysis.

4.3. Comparison of -log(p($x|X$)) and -log(p($x|k$))

Although the output of the two information operators may seem quite similar, there is a key difference between the two. The information measure based on ecological
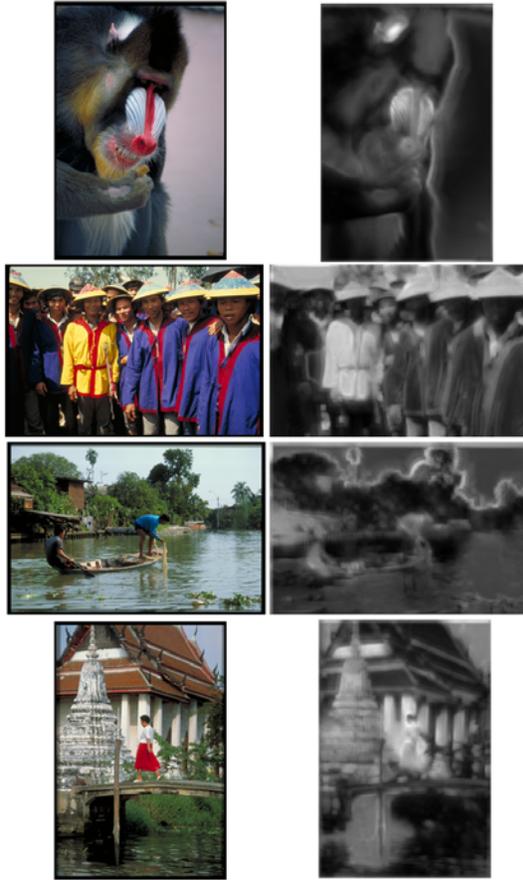
Figure 3: Information maps produced based on basis coefficient distributions learned from the one image under consideration.

Figure 4: A comparison of the output of the two distinct information operators.

statistics provides a measure of the uniqueness of local statistics in the space of natural images. The context-based information operator quantifies uniqueness of local statistics within the image in question. Figure 4. illustrates

the key difference between these two operations. The ecological information measure produces a strong response in the vicinity of the bright background, as such statistics occur rarely in all natural images. The context-based information measure includes knowledge that a large part of this image is yellow, and as such, these regions receive a small response for the context-dependant operator. In many cases, including each of the images in figure 2, the response of the two information operators is very similar. Since statistics unique within a particular image frequently correspond to statistics that are unique in general, one would expect the two operators to frequently yield similar responses in practice. That said, similarities or differences between the two measures may be particularly telling in the context of image analysis.

5. Conclusion

A set of novel image operators has been introduced, which quantify the uniqueness of local image content within a given context. A representation based on independent components is shown to allow for a representation of local statistics in a framework that allows analysis to be performed in a low-dimensional and tractable space. We have demonstrated that the information operators may be very useful in quickly segmenting foreground from background content, filtering natural textures or frequently occurring statistics, and predicting regions of interest. The resulting information maps may have significant utility in guiding more time-intensive imaging algorithms, allowing processing of the image to be focused on regions containing meaningful signal content.

References

[1] Y. Chee (1995), Information theory and its application to image coding, Tech. Rep. 11, Elec. Comp. Eng., Curting Univ.
[2] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski (2002). Face recognition by independent component analysis. IEEE Transactions on Neural Networks 13(6) p. 1450-64.
[3] M.S. Bartlett, S. Makeig, A.J. Bell, T.P. Jung, T.J. Sejnowski (1995). ICA of EEG data. Soc. for Neuroscience Abs., 21:437.
[4] T.W. Lee, M. Girolami, T.J. Sejnowski (1999), ICA using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Computation, 11, 417-441.
[5] D.J. Field, and B. A. Olshausen (1996), Emergence of simple-cell receptive field properties by learing a sparse code for natural images. Nature, 381:607—609.
[6] C. E. Shannon (1948). A mathematical theory of communication. Bell Systems Tech. Jour., 27:379-423, 623-656.
[7] H. Greenspan, S. Belongie, P. Perona, R. Goodman, S. Rakshit, and C. H. Anderson (1994), Overcomplete steerable pyramid filters and rotation invariance, IEEE CVPR, 222-228.