Learning Visual Operators from Examples: A New Paradigm in Image Processing

Hans Knutsson *Magnus Borga*

Computer Vision Laboratory
Linköping University
email: `knutte@isy.liu.se`

Abstract

This paper presents a general strategy for designing efficient visual operators. The approach is highly task oriented and what constitutes the relevant information is defined by a set of examples. The examples are pairs of images displaying a strong dependence in the chosen feature but are otherwise independent. Particularly important concepts in the work are mutual information and canonical correlation. Visual operators learned from examples are presented, e.g. local shift invariant orientation operators and image content invariant disparity operators. Interesting similarities to biological vision functions are observed.

1 Introduction

The need for a generally applicable method for learning is evident in problems involving vision. The dimensionality of typical inputs often exceed 10^6 effectively ruling out any type of complete analysis. In common practice vision problems are handled by reducing the dimensionality to typically < 10 by throwing away almost all available information in a basically *ad hoc* manner. This approach is however likely to fail if, as is frequently the case for vision problems, the mechanisms by which the necessary information can be extracted is not well understood. For this reason designing system capable of learning the relevant information extraction mechanisms appears to be the only possible way to proceed.

The development in the field of neural computation provides consistent evidence that incorporation of knowledge gained in more mature fields of research can help speed up progress in understanding important learning mechanisms. Much of the knowledge developed within the areas of information theory, signal theory, control theory and computer science is in fact at the core of learning and many researchers in the field, including us, are seeking to integrate pertinent theory and principles from these areas. Particularly important concepts in our recent work is **mutual information** [1] and **canonical correlation** [2]. Further examples of our work along these lines are given by [3, 4, 5]

Model complexity A primary question in signal processing in general and image analysis in particular is: What is the appropriate model complexity for the analysis? A naive approach would indicate a need for an extreme complexity. Consider, for example, a visual input as small as a 5x5 window of 4 bit pixels giving $2^{100} \approx 10^{30}$ possible different situations.

'Empty' spaces The following observation is, however, enlightening and somewhat reassuring. Suppose new neighborhoods can be attained at a rate of 10 MHz. It will take 10^{20} seconds, i.e. more than 3000 billion years, to see them all! This shows that the distribution of real world event samples will necessarily, even for relatively low dimensional systems, be very sparse indeed. There is simply not time enough for everything to happen. The number of actual events will leave the space almost empty.

Constraint manifolds What then, if anything, can in general be said about the distribution of real world samples? Fortunately it turns out that there is, as a rule, a manifold defined by nonlinear constraints, on which the values are jointly restricted to lie [6]. (Would a human cope otherwise?) Our hope is that it is possible to find a model generating structure adaptive enough to model only this manifold.

Mutual information The most general measure of the interdependency between input and output is the amount of mutual information. A system consisting of several local

models can approximately find the global mutual information between the input and output spaces if each local model is valid and finds the mutual information in it's region. For continuous signals locally linear models will suffice and for a linear system the mutual information is maximized when correlation is maximized. It is crucial, however that the local models are base independent. Standard correlation analysis will serve to exemplify this fact.

Base independent models Correlation analysis only measures correlations in fixed directions in the parameter space, i.e. the projections of the signal onto the basis vectors in a given coordinate system. Consider the following simple example, where the process is described by eight parameters:

Suppose that a linear combination of the first four parameters correlates fully with a linear combination of the last four. In the coordinate system given by the process parameters, the correlation matrix looks like the first matrix below. It is hard to see that the first four parameters have much to do with the last four. However, if we make the analysis in a coordinate system where the two linear combinations define two of the basis vectors, the relation is obvious (and the correlation equals one) according to the second matrix below.

$$\begin{pmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.00 & 0.00 & 1.00 & 0.00 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.00 & 1.00 & 0.00 & 0.00 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.00 & 0.00 & 1.00 & 0.00 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$\begin{pmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & \mathbf{1.00} & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 0.00 & 0.00 \\ \mathbf{1.00} & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

Canonical correlation Relations of this kind, where linear combinations of parameters are correlated and a standard correlation analysis gives weak indications of correlation, can be found if we instead look for *canonical correlations* [2, 7]. Canonical correlation analysis (CCA) is base independent and finds correlations between inner products, i.e. $\mathrm{Corr}(\mathbf{x}^{\mathbf{T}}\mathbf{w}_x\ ,\ \mathbf{y}^{\mathbf{T}}\mathbf{w}_y)$. We have developed an efficient algorithm for estimation of canonical correlation ,[5].

2 Learning from Examples

As indicated above descriptors for higher order features are in practice impossible to design by hand due to the overpowering amount off possible signal combinations. In [4] it is shown how canonical correlation analysis can be used to find operators that represent relevant local features in images.

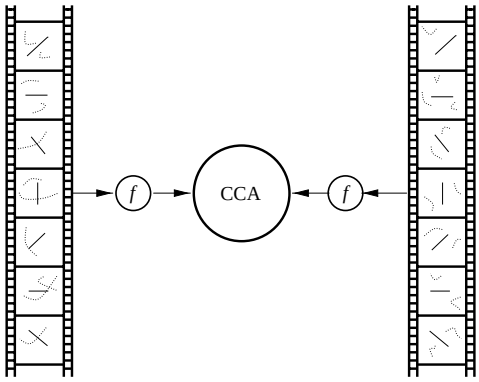

Figure 1: A symbolic illustration of the method of using CCA for finding feature detectors in images. The desired feature (orientation: here illustrated by a solid line) is varying equally in both image sequences while other features (here illustrated with dotted curves) vary in an uncorrelated way. The input to the CCA is a function f of the image.

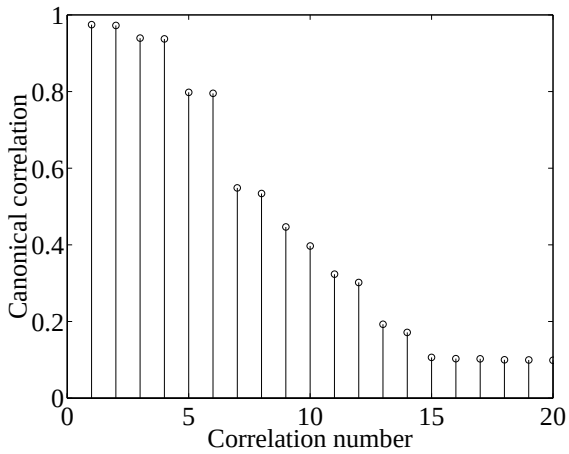

Figure 2: Plot of canonical correlation values.

The basic idea behind the CCA approach, illustrated in figure 1, is to analyze two signals where the feature that is to be represented generates dependent signal components. The signal vectors fed into the CCA are image data mapped through a function f. If f is the identity operator (or any other full-rank linear function), the CCA finds the linear combinations of pixel data that have the highest correlation. In this case, the canonical correlation vectors can be seen as linear filters. In general, f can be any vector-valued function of the image data, or even different functions f_x and f_y, one for each signal space. The choice of f is of major importance as it determines the representation of input data

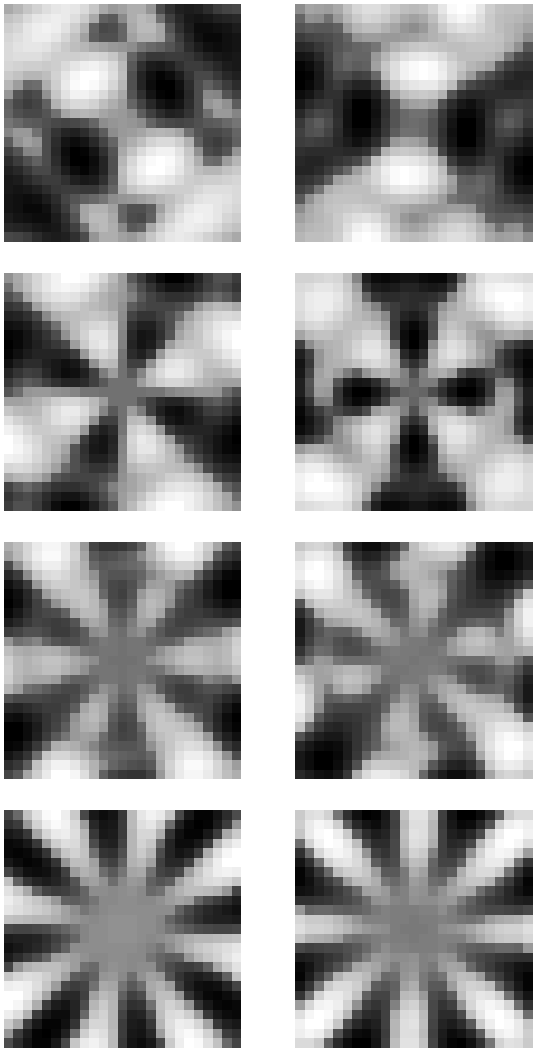

Figure 3: Projections of Fourier components on canonical correlation vectors 1 to 8. The result shows that angular operators of orders 2, 4, 6 and 8 are important information carriers.

for the canonical correlation analysis.

2.1 Local orientation

It is shown in [4] and [8] that if f is an **outer product** and the image pairs contain sine wave patterns with equal orientations but different phase, the CCA finds linear combinations of the outer products that convey information about local orientation and are invariant to local phase. Figures 2, 3 and 4 show results from a similar experiment this time using image pairs of edges having equal orientation and different, independent positions. Independent white Gaussian noise to a level of 12 dB SNR was added to all images. Figure 2 shows the values of the 20 first canonical correlations. The values appear to come in pairs, the first two values being ≈ 0.98 demonstrating that the mutual information mediated through local orientation is high.

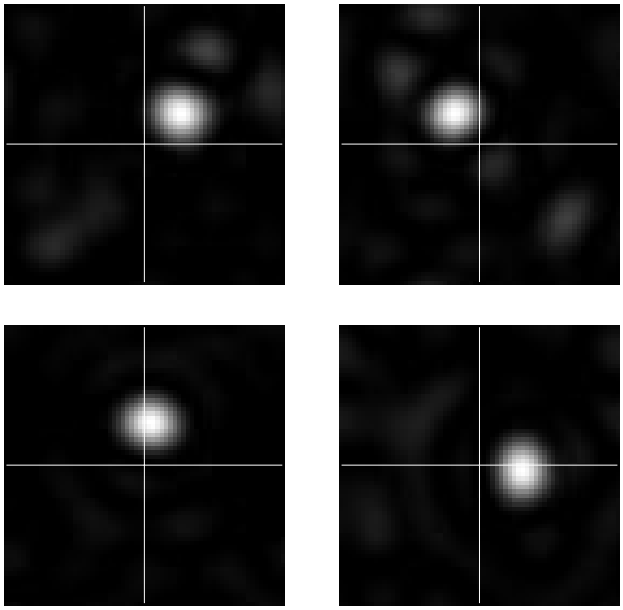

Figure 4: Spectra of eigenimages interpreted as complex quadrature filter pairs. Top two from canonical correlation vector 1. Bottom two from canonical correlation vector 2.

Figure 3 show the projections of Fourier components on canonical correlation vectors 1 to 8. The result shows that angular operators of orders 2, 4, 6 and 8 has been formed and are important information carriers. The magnitude of the projections are close to shift-invariant having a position dependent variation in the order of 5 %. Comparing to figure 2 it can be seen that the decrease in the canonical correlation values corresponds to an increase in angular order of the operators.

'Complex cells' Performing an eigenvalue decomposition of the canonical correlation vectors the corresponding linear combinations, in the outer product space, can be seen as quadratic combinations of linear filters [4]. The linear filters (eigenimages) obtained display a clear tendency to form pairs of odd and even filters having similar spectra. Such quadrature filter pairs allow for a local shift-invariant feature and are functionally similar to the orientation selective 'complex cells' found in biological vision. Figure 4 shows the spectra of four such filter pairs. The top two are from canonical correlation vector one and display selectivity to orientations 45 and 135 deg. The bottom two are from canonical correlation vector two and display selectivity to orientations 0 and 90 deg.

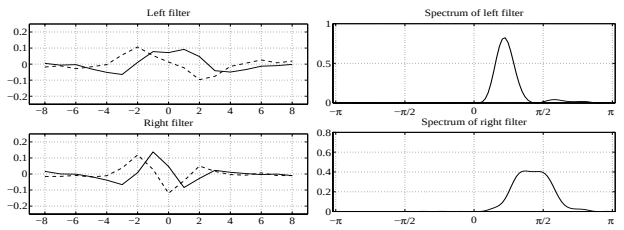

Figure 5: The filters created by CCA. Solid lines show the real parts and dashed lines show the imaginary parts.

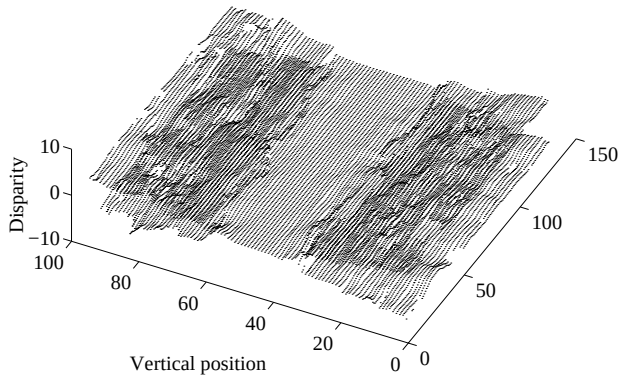

Figure 6: The result of the stereo algorithm for two random dot images corresponding to two semi-transparent crossing planes.

2.2 Local disparity

An important problem in computer vision that is suitable to handle with CCA is stereo vision, since data in this case naturally appear in pairs. In [4] a novel stereo vision algorithm that combines CCA and phase analysis is presented. It is demonstrated that the algorithm can handle traditionally difficult problems such as: **1.** Producing multiple disparity estimates for semi-transparent images, see figure 6, **2.** Maintain accuracy at disparity edges, and **3.** Allowing differently scaled images.

Canonical correlation analysis is used to create adaptive linear combinations of quadrature filters. These linear combinations are new quadrature filters that are adapted in frequency response and spatial position maximizing the correlation between the filter outputs from the two images. Figure 5 shows the filters obtained for two white noise images where disparity changes linearly with horizontal position. Note that the obtained filters have adapted to the effect of the disparity gradient through a relative offset in center frequency. The disparity estimate is obtained by analyzing the phase of the scalar product of the adapted filters. A result for depth estimates using semi-transparent images is shown in figure 6.

3 The Future

The concept of mutual information provides a solid and general basis for the study of a broad spectrum of problems including signal operator design and learning strategies.

A broadly applicable general approach, illustrated in figure 7, is to maximize mutual information subject to constraints given by a chosen model space. This could be done by varying not only the linear projections ,i.e. the CCA part, but also the functions f_x and f_y.

Finding suitable function classes and efficient parameterization/implementations for these functions is still the central issue and will be an important theme in our continued investigations.

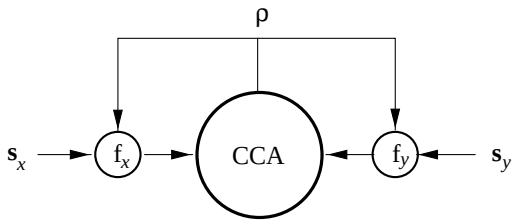

Figure 7: A general approach for finding maximum mutual information.

3.1 Simplicity vs Complexity

A possible criticism of the learned operator approach presented in this paper is the sometimes apparent complexity. The assumption that simple models of the world are more likely to be useful or robust can be traced back to the year 1320 A.D. and the monk William of Occam [9] and is often referred to as *Occam's razor*. It should be noted, however, that simplicity of a model implies that the model is simple to describe in a certain language. But what is simple to express in one language may be complex to express in another and vice versa.

Occam's razor or simply evolution? What then, are the implications of simplicity? Viewing Occam's razor in an evolutionary setting provides an interesting perspective. It is plausible that languages competing for efficiency, i.e. being beneficial for the group of individuals using it to communicate, have evolved in such a way that frequently useful statements have become easy to express. The evolution of this simplicity can be seen as making efficient use of a common context which is implicitly understood by all individuals involved. Hence, the importance of model simplicity in science should be taken with more than a grain of salt since what is complex today may be simple tomorrow.

Acknowledgment

We like to thank The Swedish Research Council for Engineering Sciences (TFR) and The Swedish National Board for Industrial and Technical Development.

References

[1] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 1948. Also in N. J. A. Sloane and A. D. Wyner (ed.) *Claude Elwood Shannon Collected Papers*, IEEE Press 1993.

[2] H. Hotelling. Relations between two sets of variates. *Biometrika*, 28:321–377, 1936.

[3] T. Landelius. *Reinforcement Learning and Distributed Local Model Synthesis*. PhD thesis, Linköping University, Sweden, SE-581 83 Linköping, Sweden, 1997. Dissertation No 469, ISBN 91-7871-892-9.

[4] M. Borga. *Learning Multidimensional Signal Processing*. PhD thesis, Linköping University, Sweden, SE-581 83 Linköping, Sweden, 1998. Dissertation No 531, ISBN 91-7219-202-X.

[5] H. Knutsson, M. Borga, and T. Landelius. Learning Canonical Correlations. Report LiTH-ISY-R-1761, Computer Vision Laboratory, SE-581 83 Linköping, Sweden, June 1995.

[6] C. Bregler and S. M. Omohundro. Surface learning with applications to lipreading. In *Advances in Neural Information Processing Systems 6*, pages 43–50, San Francisco, 1994. Morgan Kaufmann.

[7] H. D. Vinod and A. Ullah. *Recent advances in regression methods*. Marcel Dekker, 1981.

[8] M. Borga, H. Knutsson, and T. Landelius. Learning Canonical Correlations. In *Proceedings of the 10th Scandinavian Conference on Image Analysis*, Lappeenranta, Finland, June 1997. SCIA.

[9] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Occam's razor. *Information Processing Letters*, 24:377–380, 1987.