Tina Memo No. 2008-001
Internal.

Visual Intelligence: The Series.

Neil A. Thacker.

Last updated
25 / 11 / 2007

WWW.TINA-VISION.NET

Imaging Science and Biomedical Engineering Division,
Medical School, University of Manchester,
Stopford Building, Oxford Road,
Manchester, M13 9PT.

Visual Intelligence: The Series.

We have now exceeded the 100 mark with the number of reports maintained on our web site. While this might be considered an impressive figure it also presents a daunting sight (not to say obstacle) to those trying to make sense of the work. It is because of this that these reports have now been organised as volumes (or series), in order to guide the interested reader through related work and concepts.

The material for these series have been accumulated over a number of years with the intention of developing an understanding of how to approach the problems of data analysis encountered in image interpretation. These papers may be seen to differ from much of the literature, in that we do not set out to apply algorithms to arbitrary data sets in order to illustrate use. Indeed, as there are as many possible publications as there are combinations of approach and data, we consider this to be a futile task. Neither have we been very interested in performing algorithmic shoot-outs[1]. We are interested in trying to develop an understanding of the performance of the algorithms we develop. In particular the issues which determine theoretical properties, validity and domains of applicability. Only by doing this do we gain an understanding of when to apply a specific method to a particular data set. We believe this issue to be at the heart of why we are doing algorithmic research, and to do otherwise has no academic closure. We also believe that finding an answer to these issues is important for the constuction of intergrated systems which solve more sophistcated tasks. Therefore, identified constraints should be applicable to our understanding of evolved biological solutions to equivalent tasks. Ultimately, we expect this work to help us formulate a better understanding of human perception and therefore the nature of visual intelligence.

Although these volumes are intended to deal with particular areas of topical research, they are not intended to summarise these fields as would a conventional book or review. In order to do justice to the breadth of these topics and to cover them at any depth such a document would be immense (see for example 2005-009). Rather, approaches which do not fit with the identified principles which emerge from our analysis of validity do not warrant coverage. As you will see some of the principles go back to what we believe to be the scientific definition of each task. These documents are a careful selection of the topics which have demanded consideration during the period of what is now an ongoing research program of 20 years standing. They represent a subset of the total volume of work and some good pieces of work have been excluded in the interest of brevity. The material has been arranged so that the work introduces concepts in the order in which they are needed.

Each selected docment has an index on the cover page, as shown below. This allows the reader to follow the main arguments, evidence and conclusions[2]. There are currently three series, **Statistics and Segmentation**, **Features and Measurement** and **Recognition and Intelligence**. The second two volumes build directly, and are consistent with, the methods and principles established in the first.

You may notice that the dates on these documents are not in order. This is generally because having published a particular piece of work we found it necessary to explain the unwritten principles which had been considered, either to later PhD students or to people outside of the research group. These later works are often unpublished, as sometimes we did not believe a selection of expert reviewers would want to see a document published which aimed to dismiss some ideas and approaches. In these cases the best place for these documents was our own web pages. We can only expect a reader to appreciate the opinions expressed in these documents in the context of our other work.

[1]Shoot-out have their uses, particulary with respect to commercial application, but good performance is always driven by appropriate assumptions regarding data generation. It can be considered as logically flawed to invent an algorithm first and only worry about how it works on data later. Shoot-outs will only identify those algorithms which are based upon assumptions which are valid for the data set selected.

[2]If you are only interested in one aspect of the work remember you can go back and look for other associated documents via the web linked cross-index.

Statistics and Segmentation.

This document forms part of the **Statistics and Segmentation** Series (2008-001) available from www.tina-vision.net.

2007-008	Tutorial: Defining Probability for Science.
2001-007	Performance Characterisation in Computer Vision: The Role of Statistics in Testing and Design.
2002-007	The Effects of an Arcsin Square Root Transform on a Binomial Distributed Quantity.
2001-010	The Effects of a Square Root Transform on a Poisson Distributed Quantity.
2004-004	Shannon Entropy, Renyi Entropy, and Information.
2002-002	Validating MRI Field Homogeneity Correction Using Image Information Measures.
2004-001	Empirical Validation of Covariance Estimates for Mutual Information Coregistration.
2004-005	The Equal Variance Domain: Issues Surrounding the Use of Probability Densities in Algorithm Design.
2009-008	Avoiding Zero and Infinity in Sample Based Algorithms.
2001-008	Derivation of the Renormalisation Formula for the Product of Uniform Probability Distributions and Extension to Non-Integer Dimensionality.
2001-005	Model Selection and Convergence of the EM Algorithm.
2003-007	Noise Filtering and Testing for MR Using a Multi-Dimensional Partial Volume Model.
2002-004	A Novel Method for Non-Parametric Image Subtraction: Identification of Enhancing Lesions in Multiple Sclerosis from MR Images.
2001-014	Bayesian and Non-Bayesian Probabilistic Models for Image Analysis.
1997-001	The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data.
1999-001	The Bhattacharyya Measure requires no Bias Correction.
1999-004	B-Fitting: An Estimation Technique With Automatic Parameter Selection.
2005-008	Tutorial: Beyond Likelihood.

The main documents in this series are 2007-008, 2004-005, 2001-014, 1999-004 and 2005-008. The other documents (often published papers) are provided as introductory material.

This series begins by making a case for the quantitative use of probability in science and engineering (2007-008). It was written following observations regarding the practical analysis of data in many areas, and the consequences of adopting various approaches. Eventually, the common difficulties observed in individual projects were tracked back to the very definition of probability itself and this document summarises what I believe could already have been concluded from a consideration of the philosophy of probability and its application to science. As this does not require any reference to practical examples, but is consistent with what is done in the remainder of the body of work, it makes sense to put this at the start. The main argument, that useful analysis of data is necessarily quantitative, recurs throughout the other documents. You may not think this point is particularly contentious until you see the conclusions which follow. Then you might begin to realise why we do not subscribe to some methodologies which are seen frequently in journals and at conferences. As we considered this work to be unpublishable in our area, the document has been sent out for comments to a large number of fellow academics. Some of the resulting comments were used to refine the text, others have been included as appendices[3].

The next document, 2001-007, was published, and outlines the link between quantitative probability, statistics, algorithm design and testing. After reading this document we expect many will have got the message that a theory of data analysis should be capable of predicting key performance measures. One can never truly understand something which is treated as a black box[4].

This is followed by two background documents, which analyse the effects of variance normalising transforms on Binomial and Poisson distributions (2002-007 and 2001-101). The reasons for these analyses come from the frequentist interpretation of probability and do not become apparent until the "equal variance" paper 2004-005, the next few documents (2004-004, 2002-002) also help to set the scene. The aim of 2004-001 was specifically to show that entropy based algorithms are not what they seem. We consider the equal variance work to be fundamentally important to the scientific use of probability in algorithm design. In particular it explains that the common definition of many methodologies, including Likelihood[5], is ambiguous, and how this can be resolved in a way

[3] Paul Bromiley's comments add a new perspective to the original arguments, and are often more conscise.

[4] Patrick Courtney was a key motivator for getting this work into a public domain. Without him these ideas would not have got beyond lectures and tutorials.

[5] Throughout these papers I break with convention and use upper case to indicate that Likelihood is the name of a specific design methodology and therefore a real noun and not to be confused with conventional definitions of the word.

which is consistent with quantitative use of probability. It also explains how unambiguous measures of probability similarity can be constructed for frequentist probability. These are both potential sources of multiple (contradictory) approaches to the analysis of data in pattern recognition and computer vision. Here these ambiguities are eliminated by applying the conclusions and general philosophy outlined in 2007-008. The approach constitutes a novel theory for the construction of algorithms, which is motivated by a more restricted view of probability than is currently encountered in the literature. To date our attempts to get this understood by reviewers in our preferred journals have met with a lack of success[6]. It sits on the web pages as a major research output while we reconsider our strategy.

The next few documents (2001-008, 2001-005, 2003-007 and 2002-004) are all practical examples of the use of quantitative probability in segmentation. They set the scene for the summary paper 2001-014, which arrives at conclusions regarding the general utility of Bayes Theorem in algorithm design. We were lucky enough to get this published following a one day meeting of the BMVA which produced a special edition in Image and Vision Computing[7]. One of the reviewers insisted that we could only claim that the theoretical points illustrated were valid for medical image analysis, rather than computer vision in general. We did not feel it would help get the paper published by starting an argument regarding the philosophy of science, so we modified a few sentences. The version on the web pages is in the form we wanted to publish. We made sure that 2007-008 contained enough discussion of the philosophy of science so that an interested reader will uderstand why we think this is valid.

The next three documents (1997-001, 1999-001, and 1999-004) follow the application of quantitative probability to the problem of model selection. 1999-001 was unpublished, but is a missing piece in the theory of probability overlap comparison. Together these documents explain why model selection is **the** key problem in the construction of automatic analysis systems. The final document assesses the perceived limitations of Likelihood based methodologies and possible extensions. 2005-008 is necessarily dismissive, and some would say contentious, in its attempt to reconcile philosophical requirements of the scientific method with the goals of automated data analysis. In particular it justifies using frequentist probability as the basis for quantitative optimisation of generalisation, which is the main idea presented in 1999-004. At the time this paper was novel, though as it represented an obvious extension to the practices universally accepted in the field of neural networks, I would expect that similar ideas have surfaced in this literature since. 2005-008 finishes with some recommendations regarding the valid use of variations around Likelihood for data analysis. Currently these are the main conclusions of this part of the research programme. The ideas constrain the work presented in the following two series.

[6] Manuel Trucco was pulled in as a co-author, specifically to help present the ideas in a way that a computer vision researcher might appreciate.

[7] Tim Cootes usefully provided a one page appendix for us which replaced a 15 page derivation, this got him onto the author list.

Features and Measurement

This document forms part of the **Features and Measurement** Series available from www.tina-vision.net.

2002-005	Tutorial: An Empirical Design Methodology for the Construction of Machine Vision Systems.
1996-001	Algorithmic Modelling for Performance Evaluation.
2006-002	A Statistical Framework for Detection of Connected Features.
2006-007	Quantitative Verification of Projected Views Using a Power Law Model of Feature Detection.
1997-003	Tutorial: Supervised Neural Networks in Machine Vision.
1995-003	Invariance Network Architecture.
1992-001	Combining the Opinions of Several Early Vision Modules using a Multi-Layer Perceptron.
2005-006	Curve Fitting and Image Potentials: A Unification within the Likelihood Framework.
1996-002	Tutorial: The Likelihood Interpretation of the Kalman Filter.
1994-003	Using a Switchable Model Kalman Filter.
2004-012	Tutorial: Computing 2D and 3D Optical Flow.
2005-011	Comparing the Performance of Least-Squares Estimators: when is GTLS Better than LS?
1994-001	Tutorial: Overview of Stereo Matching Research.
1992-002	Online Stereo Camera Calibration.
1995-002	Calibrating a 4 DOF Stereo Head.
2000-009	An Evaluation of the Performance of RANSAC Algorithms for Stereo Camera Calibration.
2001-011	The Evolution of the TINA Stereo Vision Sub-System.
2007-011	A Methodology for Constructing View-Dependent Wireframe Models.

This series builds directly upon the principles outlined in Series 1 and their application in the area of structural analysis of images. As well as trying to cover the basic problems in this field, feature detection and "shape from", most of the material is intended to show the link that exists between popular algorithms and Likelihood. Once this association has been established, and the assumptions which would generate the chosen algorithm from Likelihood are known, we are faced with some stark questions:

- Is Likelihood the true origin of the algorithm, or have we just invented a new principle for data analysis?
- If (as is nearly always the case) it is the former, are the required assumptions realistic?

These questions provide a platform for the assessment of novel algorithms, and provide an avenue for academic closure within each topic.

The first document, 2002-005, is based upon tutorial material which has now been given at machine vision summer schools (and international conferences) for over a decade. It outlines methods for the quantitative testing and development of computer vision systems. This includes a flow chart which explains the key issues which must be addressed during various stages of algorithm development.

The next set of papers and tutorials cover a quantitative approach to the detection and location of features. The neural network material (1997-003, 1992-001) is included as an illustration in learning non-linear classification systems. These ideas resurface in Series 3. Here they are used to solve the problem of data fusion, which occurs frequently in the construction of image interpretation systems. 2005-006 is written to illustrate the link between Likelihood and popular methods for shape extraction from features, which make use of non-statistical algorithmic methods such as image potentials and "snakes".

The two documents 1996-002 and 1994-003 illustrate the use of Kalman Filter based techniques in object tracking. While the first shows the links of this filter to statistical estimation, the second uses the results for model selection outlined in Series 1 in order to choose the most appropriate predictive model. The temporal theme is continued in 2004-012, with an assessment of mature approaches to the extraction of shape from optical flow[8].

Many approaches to structural estimation in the computer vision literature are based more upon geometry than statistics. You might say that they are, at best, approximately statistical. In 2005-011 [9] we had the chance to

[8] This document is based directly upon the experience of John Barron, and exists only because of his enthusiasm and effort.
[9] Again in collaboration with Manuel Trucco.

collaborate on a paper which illustrates the theoretical limitations of these approaches and potential consequences. We show how the "linearise and invert" approach to parameter estimation rarely embodies appropriate statistical measures for the data. This should be heeded as a warning to researchers in this area. Mathematical sophistication, and closed form mathematical algebraic solutions, may make for quick algorithms but are not a guarantees of validity. Conversely, any algorithm which can be seen to make appropriate statistical assumptions and apply a valid methodology must be taken as the "correct" solution, even if it is mathematically trivial (such as optimising the correct robust cost function with your favourite numerical method)[10].

The final set of papers and tutorials (1994-001, 1992-002, 1995-002, 2000-009) assess the problems associated with the construction of quantitative stereo vision systems. A tutorial is provided simply to set the scene, the following work is done once again in the context of applying quantitative statistical approaches in a way which makes appropriate assumptions regarding the data. In particular we wish to take appropriate account of the relative errors on input data and compute error covariances on the result.

The remaining papers (2001-011 and 2007-011) explain the architecture for the TINA stereo vision system. They revolve about the calibration of working systems and the quantitative use of data in object modelling[11]. It provides one final justification for taking this approach to vision. Understanding the characteristics of input and output data not only aid system construction, they are fundamentally necessary for it. The information available in an image is fixed at the time it is acquired, and some information (such as dense depth estimates on smooth surfaces) is not present. We should therefore not ask "How do we estimate X?", on the assumption that X can be usefully computed, but rather "What is an appropriate X which provides a quantitative (therefore practically useful) summary of image content?".

[10] This observation appears to run counter to the reviewing practices found for conferences and journals. Mathematically sophisticated papers often being preferred as though complexity is, in itself, a measure of theoretical rigor.

[11] This approach, pioneered by Steve Pollard and John Porrill during their years together in Sheffield, illustrates the TINA (there is no alternative) philosophy.

Recognition and Intelligence

This document forms part of the **Recognition and Intelligence** Series available from www.tina-vision.net.

2007-001	Retinal Sampling, Feature Detection and Saccades: A Statistical Perspective.
2006-008	Statistical Principles for Selection of Computer Vision Algorithms as Modules for Visual Perception - Show Me the Errors.
1991-001	Designing a Layered Network for Context Sensitive Pattern Classification.
1997-002	Supervised Learning Extensions to the CLAM Network.
1996-003	Tutorial: Algorithms For 2-Dimensional Object Recognition.
1997-005	Speechreading Using Probabilistic Models.
2000-002	Solving Shape Based Object Recognition from a Computational Standpoint - Practical and Physiological Constraints.
1995-004	Assessing the Completeness Properties of Pairwise Geometric Histograms.
1996-004	Robust Recognition of Scaled Shapes Using Pairwise Geometric Histograms.
1996-005	Multiple Shape Recognition Using Pairwise Geometric Histogram Based Algorithms.
2007-007	Automatic Identification of Morphometric Landmarks in Digital Images.
1999-002	A Feature Representation for Map Building and Path Planning.
2001-015	Colour Image Segmentation by Non-Parametric Density Estimation in Colour Space.
2001-006	What is Intelligence?: Generalised Serial Problem Solving.
1994-002	A Correlation Chip for Stereo Vision.
1995-001	Specification and Design of a General Purpose Image Processing Chip.

This series describes an ambitious effort to understand the nature of visual intelligence. It uses the framework specified in the first two series to help define goals approaches and solutions. The first document (2007-001) is an interpretation of biological structure as a process of measurement with required invariance, specifically illumination, scale and orientation invariances. The main conclusions of this analysis are then used in an attempt to understand attentional focus in human psychophysical experiments. The work here is much later than the published papers which follow in this series. Some of the goals and assumptions, which we initially assumed would be obvious and shared by the physiological modelling communities, were not and needed justifying and explaining.

The following document (2006-008) sets the scene for visual recognition by considering the properties that visual data is expected to have (non-linear disjoint distributions on manifolds of variable dimensionality) and the capabilities of current approaches to pattern recognition. The general conclusions of this document are then applied in the design of a self generating learning system which is consistent with a simple interpretation of neuronal function (1991-001 and 1997-002). The work shows how (frequentist) conditional probabilities of classification can be generated using neural tissue. This interpretation of function places constraints on the way that data can be represented and processed in the brain. The theoretical interpretation is again consistent with the statistical principles outlined in Series 1, and in particular the use of the Bhattacharyya measure for frequency distribution (histogram) comparison. We suggest this as the theoretical statistical basis for the comparison of neuronal patterns in the brain.

The next group of papers (1995-003, 2000-002, 1995-004, 1996-004, 1996-005, and 2007-007) then outline an approach to object recognition, which is consistent with both the biology and statistical behaviour. They show how the same invariances as considered in 2007-001 lead to a unique representation of visual shape. The papers contain key theoretical results, particularly the observation of completeness of the shape representation which we call Pairwise Geometric Histograms. The method for detailed recognition (location) of grey level structure (1997-005) which may well be very similar to physiological mechanisms for face recognition) is largely provided for completeness and perhaps illustrates, at the same time, the capabilities and fundamental restrictions of such approaches in the context of a general vision system.

The next series of papers build up to a theory of intelligence. Starting with 1999-002 and 2001-015, the aim is to show how maps of visual similarity can be used to solve a variety of tasks, including colour segmentation and map building and path planning. 2001-006 then assesses the definition of artificial intelligence in order to show how visual representations together with solutions to the model selection problem (from Series 1), may be combined to solve non-trivial (intelligent) planning tasks in the context of learning systems. We suggest that a Turing machine, capable of identifying and performing sequences of actions, would be very useful for performing visual control tasks but only if this system could learn from observations of continuous variables in noisy environments. We call this concept "Generalised Serial Problem Solving". To date we have not succeeded in constructing the complete system, but this document shows that we have starting points for key aspects of the theory. In particular the role of probability theory and statistics seems to offer a natural additional constraint to those generally considered when building models of perception. The document finishes with a consideration of computational requirements, it therefore seemed natural to finish the series with two published VLSI architectures for computer vision.