Tina Memo No. 2006-001
Internal Memo

A Guide to Sound Scientific Methodologies for Machine Vision Oriented PhD Students

Simon Coupe and Neil Thacker

Last updated
15 / 12 / 2005

WWW.TINA-VISION.NET

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

Abstract

This essay presents various factors which are suggested to warrant consideration for machine vision oriented PhD students. The factors suggest principles of best practice which are intended to ensure the validity of the work undertaken for the PhD. These ideas also have value in the assessment of research material when conducting literature surveys. The views reflect the methodologies upon which the TINA [1] open source computer vision system is founded.

Introduction

Despite the large amount of interest (and effort) directed towards developing intelligent computer vision systems, a generalized system exhibiting human-like levels of perception still eludes us, as does a computational model of the process. This can primarily be attributed to the sheer complexity of the task at hand and the relative novelty of able computing technologies. However, it is proposed that by ensuring all associated research adheres to a set of related, self evident scientific methodologies, we should be better able to more quickly realize the full potential of our developments.

Essentially, the proposed methodologies are intended to ensure the production of accurate, consistent, dependable, reusable, well engineered and worthwhile computer vision systems.

Outline of Methodologies

- The scientific method
- Problem reduction
- Misapplication of scientific principles
- Deterministic vs. stochastic theories
- Ambiguity
- Errors
- Statistics
- Evaluation and testability
- Open source development

The Scientific Method

The scientific method is the approach around which all scientific understanding is based. The method is one of evidential reasoning, which aims to test the validity of any scientific theories and identify areas of weakness in our understanding.

The first step of the scientific method is to formulate a quantitative theory which accounts for some observed phenomenon. The theory is then used to make predictions relating to the existence of the phenomenon, against which any actual observations can be quantitatively compared. If theory and observation consistently coincide (within a statistical framework (see Deterministic vs. Stochastic Theories below)), then the theory is validated up to the limits of the statistical test. This does not mean the theory is correct, only that it cannot be refuted using the available data. Any theories are accepted as potential descriptions of the process until reapplication of the scientific method proves otherwise.

Although an apparently simple and intuitive approach, the scientific method is the basis of modern science. The theories used in and developed for the science of machine vision should be no different and should all be formulated in adherence to the scientific method.

Problem Reduction

In order to solve any problem we must first be able to define it. This requires us to specify the required task and the data available to solve it. Unable to solve certain computer vision problems, some scientists may be willing to abandon their fundamental objectives and reduce their tasks to more manageable ones. This may inlcude restricting the class of problem (or image data set) for which the algorithm is to work so that the problems can be solved by hacks [1]. We believe that as professional scientists we should be focused on providing sound solutions to our underlying scientific problems. In doing so, we should be better able to advance the science of computer vision and ensure that our endeavors stand the test of time. We should therefore evaluate any work on the basis of what we consider the real challenges to be, and not readily accept redefinitions which trivialise such tasks and restrict the scope of application.

When presented with a piece of work, one question we should ask ourselves is; will this research still be relevant in 20 years' time?

Misapplication of Scientific Principles

Decades of computer vision research have resulted in the formulation of a multitude of diverse algorithms, intended to model associated processes or, at least, provide us with clues relating to the solutions of our problems. It is often the case that principles which were formulated as solutions to specific scientific problems (typically in the realm of physics) are naively adopted to provide solutions to other, unrelated problems. In such cases, the assumptions behind the original use of the theory are neither identified nor considered. As professional and credible computer scientists, it is proposed that we should be strictly focused on understanding and modelling authentic solutions to our problems. In particular, we need to be aware of the disticntion between deterministic and stochastic theories and the circumstances when each are appropriate.

Deterministic vs. Stochastic Theories

Deterministic theories pertain to mathematical formulas whose results are precisely determined from their input. Most theories of classical physics are deterministic (e.g. mechanics).

Stochastic theories are those whose results incorporate a random element, whereby theoretical predictions are statistical in nature, such as distributions. Most theories which describe measurement processes are necessarily stochastic (e.g. parameter estimation tasks in compter vision). Fundamentally, the presence of measurement noise gives rise to possible distributions of parameter estimates.

Problems arise when theories are assumed to be deterministic (such as geometric or optical/graphics based models), but in fact are influenced by, for example, measurement errors, uncontrolled environmental conditions or noise. In such circumstances, the theories required are stochastic ones, where quantitative statistical analysis is required to account for these potential sources of error. Failure to adhere to such a requirement can result in unreliable estimates of parameters and an inability to estimate associated uncertainty (see Errors below).

Ambiguity

To further reinforce the authenticity of our algorithms, ambiguity should be avoided in designing computer vision systems. For example, an algorithm which requires specific sets of control parameters may be seen as a powerful and flexible image analysis tool. Equally, such approaches require specific a priori knowledge of scene contents. If we assume that the field of computer vision is defined in order to provide solutions to scene interpretation tasks, then this corresponds to requiring the answer as an input (see Problem Reduction above). We should therefore seek solutions to problems which avoid the use of arbitrary control parameters. Algorithmic control parameters should be strictly definable by theory or experiment. Unconstrained and arbitrary control parameters result in problems when attempting to characterise the performance of an algorithm for practical use (see Testability, below).

[1] hack: An inelegant and usually temporary solution to a problem.

Errors

In modelling and interpreting information from the physical world, errors are perhaps the most important consideration. Our assumptions, calculations and interpretations are prone to errors and uncertainty. As the scope for application of our advances evolves, the need for precision, reliability and resultant dependability and safety becomes paramount. Sources of error are widespread and of course application dependant. To develop robust, autonomous computer vision systems, all potential sources of error need to be accounted for within a unified framework. In addition, and as a consequence, a module intended for a larger system must provide estimates of errors on output data if it is to be used. We can think of this as an example of the application of the scientific method in the context of scene interpretation. Any algorithm which does not provide estimates of error makes any quantitative comparison with a theoretical interpretation impossible. It therefore becomes impossible to ignore the errors on the outputs from an algorithm once we decide to make practical use of the data.

Statistics

Naturally, the framework with which we can account for variability and predict the accuracy and bounds of our computations is probabilistic (see Deterministic vs Stochastic Theories above). Using quantitative statistical analysis we can therefore aim to fully understand, and if required, compensate for, any limitations of our data and calculations whilst optimizing the performance of our vision systems. In addition, well founded theories, based upon probability theory and statistics, are not expected to generate arbitrary control terms (see Ambiguity above).

Computer vision can be thought of as a branch of applied statistics. The core task of scene interpretation, for example, essentially involves forming a decision pertaining to the most likely distribution of supposed subject matter within a scene.

Statistical analysis should be borne in mind at all stages of system development. In the first place, the quantitative statistical properties of the required input data should be analyzed, tested and modelled. Algorithms can then be designed to best and more comprehensively cater for the problems at hand.

The sheer complexity of computer vision as a problem (see Problem Reduction above) necessitates a modular system framework. Component consistency is therefore vital for the accuracy and dependability of any such system. This can be secured by ensuring that all computed modules operate in a quantitative statistical framework whereby any assumptions are consistently applied throughout.

An Illustrative Example

A typical vision guided, mobile robotics application serves to highlight the statistical considerations of the previous sections:

For the robot to be able to pass cleanly through a doorway, avoiding damage to itself and the doorframe, it needs to be certain that it can do so. Rather than naively expecting all its calculations to be watertight, the robotic system should be designed to appreciate any prospective errors. For instance, if the robot was to perceive the door to be exactly 1 meter wide, considering the range of potential errors to which it was susceptible, it is more likely that the door would be physically sized somewhere within a range of associated values, for example, 0.95 and 1.05 meters. Quantitative statistics can be used to estimate such ranges and their associated probability distributions. When an entire system is constructed in such a fashion, any assumptions and degrees of uncertainty can be propagated throughout, thereby ensuring any resultant decisions will account for appropriate levels of uncertainty. If, for example, the robot therefore concluded that it was only 55% certain it would fit through the doorway; logic would suggest that another route, or at least a reevaluation, would be required.

Evaluation and Testability

Although the fundamental implication of any new research publication describing a novel algorithm is that the algorithm is somehow better than another approach, on some criteria or under some circumstances, the area has found it remarkably difficult to identify and converge upon the most appropite approaches for the analysis of data. Often this is due to an over-reliance upon empirical testing as the only way to understand performance, an approach which is always dependent upon data sample selection. However, computer vision systems all too often suffer from a lack of robustness to changes in data sample. In designing computer vision systems, it is therefore suggested that greater emphasis be placed on gearing our algorithms to deal appropriately with sets of test data which more faithfully represent the diverse variability typical of the real world.

Designing algorithms using statistical methods offers the potential for being able to make quantitative predictions of expected performance, in terms of error rates and accuracy. An important consequence of such a methodology is that all resultant computer vision systems stand on a level test bed, whereby comparisons and evaluations can be meaningfully made. Furthermore, the most appropriate ways to analyse data can be determined by the statistical definition of the problem and any quantitative predictions.

The area of computer vision is inherently mathematical and contains many branches, including, for example, geometry and graphical optics. However, if we are looking for the area of computer vision which could realistically be regarded as the theoretical component of a scientific subject (against which real world measurements or observations can be compared), then this area requires the quantitative use of probability and statistics. Though geometry and optical modelling may be required as part of the algebra, quantitative predictions of the performance of algorithms should be regarded as the basis for a 'theory of computer vision'. Any approach which precludes meaningful testing of underlying theory, therefore being outside the scope of the scientific method, should of course be considered as un-scientific.

Open Source Development

As well as ensuring that a student has a complete understanding of their chosen subject matter, a PhD should serve to make novel scientific advances. Well founded open source development environments promote reuse and therefore avoid duplication of work. Not having to 'reinvent the wheel' every time a PhD is undertaken, allows for more time to be invested in the primary objective of scientific discovery. Similarly, by making our discoveries and advances freely available to others, an environment more conducive to progress and innovation is formed. The pressures of peer scrutiny could, for instance, motivate researchers to maximize the quality and integrity of their work.

Closing Remarks

This paper has hopefully raised the awareness of machine vision oriented PhD students for the need for a self-consistent scientific methodology. A quantitative statistical framework is suggested to underpin such a methodology. Not only should students' ongoing work benefit from such insights, but they should resultantly be better able to evaluate the work of others and therefore make best use of their time. Whilst machine vision students have been the intended audience, the theories should also bear significance for students of other related scientific disciplines.

Reference

[1] TINA (Open Source Image Analysis Environment) www.tina-vision.net

This essay was written to document a tutorial given by Drs. Neil Thacker and Paul Bromiley in association with TINA in October 2005.