EPILOGUE

The hardest part about writing this book was stopping. At the end of each chapter, we could see many other directions to explore, either segments of the recent literature that were not adequately represented or new problems for which we had assembled the tools but not yet begun to use them. Our inclination was to follow as many of these paths as possible, stretching our methodologies to their limits, but of course there would have been no end. We had to draw a line and send this tome to press. To give ourselves the courage to do so, we submit this modest epilogue, suggesting to the reader the paths we might have followed had we had several other lifetimes to do so, and we invite the reader to accompany us on the portions of the journey that we still hope to take.

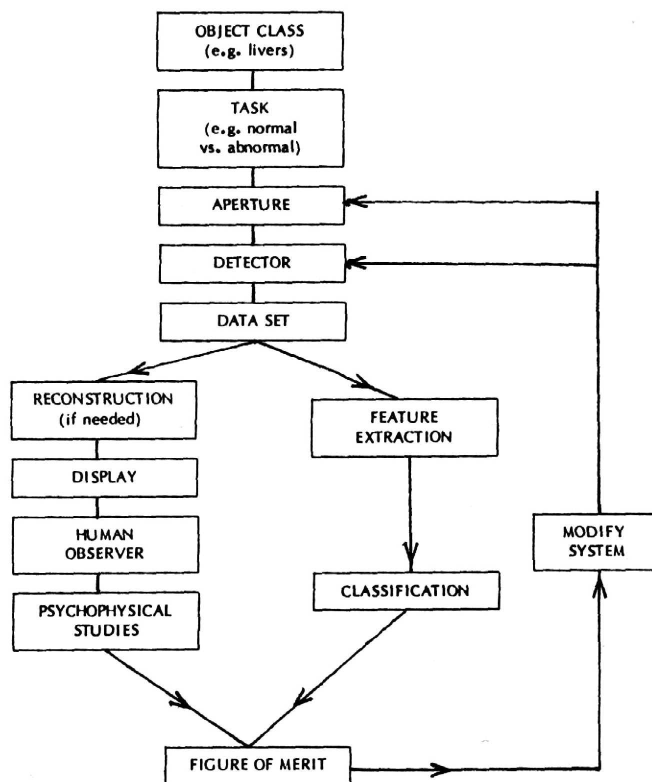

Fig. 1 Roadmap for the systematic optimization of gamma-ray imaging systems (from Myers *et al.*, 1986).

Systematic system optimization We begin the journey with a backward glance and a look at the roadmap to see where we are. In 1985, as the concept of task-based assessment of image quality was taking hold in the medical-imaging community, the diagram in Fig. 1 was presented at a conference at Georgetown University. This diagram embodies two basic principles that we have tried to stress in this book. First, it recognizes that meaningful measures of system quality must be based on the performance of specific observers on specific tasks of practical interest. Second,

it envisions an imaging system as an integrated whole, with all components contributing to task performance and hence needing to be considered together.

A crucial feature of this diagram is the feedback path. After we have learned how to evaluate imaging systems, the next logical step is to incorporate the evaluation into a program of systematic, iterative optimization of task performance.

In the two decades since we first sketched this diagram, many researchers have contributed to the methodology implied in it. We now know much more about image formation, noise in imaging systems, image reconstruction algorithms and the properties of human and model observers.

The challenge for the future of image science is thus twofold: We must continue to refine our understanding of every block in diagrams like the above and we must think seriously about closing the loop and carrying out a full system optimization.

The search for excellence A basic difficulty in actually carrying out the plan suggested in the figure is that the specification of an imaging system is so complex. Consider the example of lens design, where the routine procedure is to assume some merit function and design the lens to optimize it. We would, of course, advocate that the merit function be chosen in relation to specific tasks, but no matter what merit function is used, the search for an optimum is difficult.

Even after the designer has selected the number and arrangement of the individual elements in the lens, there are still many parameters to be chosen, including interelement spacings, curvatures and indices of refraction. Iterative algorithms such as simulated annealing and genetic algorithms can be used to find a more-or-less global optimum in this parameter space, but there is no guarantee that some completely different arrangement of elements with a different set of free parameters might not perform better. Task-based assessment allows objective comparison of competing designs but does not obviate the role of human creativity in selecting the competitors in the first place.

If it doesn't fit, get a bigger hammer! Imaging systems acquire huge amounts of data and often require formidable data-processing resources even to produce one image. Since image quality is inherently a statistical concept, accurate evaluation of a single system requires a huge number of images, and systematic optimization requires, in effect, evaluation of a huge number of imaging systems.

Daunting though this prospect may seem, we can be encouraged by the rapid advances in computational capabilities. In its most expansive form, Moore's law says that every aspect of computer power (chip density, CPU speed, memory, data-transfer rates) will double every 18 months or so. For three decades shortsighted prognosticators have been predicting the end of this trend, but they have consistently been proven wrong. If the trend holds for another three decades, the gain in all facets of computer power will be about $2^{20} \simeq 10^6$. If we crudely benchmark a typical personal computer at the turn of the millennium (January 1, 2001, of course) as a single 1 GHZ processor with 100 MB of memory, we can envision the image scientist of 2031 having access to 1 PHz of processing speed and 100 TB of memory (T = tera = 10^{12}, P = peta = 10^{15}). These numbers might imply a huge number of processors and memory spread around the world, but they do not require an inordinate amount of technological optimism; the pessimistic view is that they will not be correct until 2041 or 2051.

What will image scientists of the not-so-distant future do with such computer power? Possibilities abound. Three-dimensional reconstructions on $1000 \times 1000 \times 1000$ grids will be routine, and rapid temporal sequences with hundreds or thousands of 3D frames will make high-resolution 4D imaging possible. Huge databases will be available for image archiving, and rapid access to the databases will facilitate automated image interpretation. Simulation tools will make the phrase "photographic realism" obsolete, and the authors of this book will quit chiding their colleagues for megalopinakophobia.

Most importantly, the computational arguments against meaningful system evaluation will disappear, and we can begin to close the loop and really optimize imaging systems. Along the way we have to increase our understanding of every component of the system, from radiation source to final observer, and we have to be creative in devising ways of improving them.

It's the data, stupid! A famous American politician told his campaign staff, "It's the economy, stupid!" thereby imploring them not to lose sight of the most significant factor in their endeavor. In imaging, the key factor is the data; robust data are as important to an image scientist as a robust economy is to a politician. In this analogy, image processing plays the role of spin — it can put a good face on a bad data set but cannot really overcome its limitations.

A data-driven approach to image science would obey three dicta: 1. Use all of the image data; 2. Get more image data; 3. Get more nonimaging data for use in conjunction with image data.

One might think that modern imaging does use all of the image data, but in fact there are information losses at several stages. As a simple example, a time exposure of a time-varying object discards potentially useful temporal data. A more complex example concerns position-sensitive photon-counting detectors where multiple sensor outputs are reduced to estimates of x and y coordinates for each photon. Information loss can occur in the estimation step and again when the estimates are binned into an image matrix. In fact, any binning or quantization of data is a potential source of information loss; moreover, we often use lossy algorithms for data compression. The challenge is to understand how serious these losses are, in terms of specific imaging tasks, and to devise ways of minimizing them.

Getting more imaging data may mean improving the spatial or temporal resolution of a system or it may mean acquiring information from different angular views, different wavelength bands or completely different imaging modalities. Often these advances come only after intensive research and considerable economic investment, so it is imperative to understand how they facilitate the intended application of the images.

Auxiliary nonimaging data are routinely used in conjunction with image data. A radiologist uses clinical indications and patient history along with radiographs in making a diagnosis. Analysts of landsat images use weather and climate data to aid in assessing the health of agricultural fields, and astronomers supplement their direct observations with data from celestial mechanics calculations and astrophysical simulations.

There are many other opportunities to integrate imaging and nonimaging data. The nonimaging data may concern the particular object being imaged or it may concern one or more classes of objects. In a medical context, for example, one might acquire further measurements on the particular patient being studied or one might

compile a database of disease characteristics and use it to guide the acquisition and interpretation of medical images. The challenge is to learn what specific supplementary information is most useful to the goal of the imaging and to devise ways of acquiring it systematically and optimally.

How good does it have to be? It is not uncommon for the designer of an imaging system to ask a user what resolution is needed. The reply may well be that there is no need for resolution better than some value "because there is nothing to see" at a finer scale. Both the question and the answer are misguided.

The question is misguided since performance on any task will always improve with better data, either lower noise or finer resolution. The user of the imaging system or the marketing manager of the company that manufactures it might then say that the performance is "good enough" at some noise level and resolution, so there is no benefit in making it better. For example, a physician might be interested in detecting a tumor 1 cm in diameter and might feel that some particular system is adequate for this task. Pressed to say how he knows, the diligent physician might do an ROC study and get, say, 0.9 for the area under the curve. Is that good enough? Not for the patient whose tumor is missed! And even if economic cost-benefit arguments are adduced for not trying to improve the performance on this task, there will always be other tasks, such as detecting smaller lesions or distinguishing benign from malignant ones.

The issue is more subtle when we look at characteristics of individual components rather than the overall imaging systems. For example, if the point response function of a lens is much smaller than the size of a detector pixel, one might conclude that further improvements in the lens are not needed. Similarly, for decades it has been conventional wisdom in nuclear medicine that detector improvements are unnecessary since the limitation is the collimator.

In neither of these examples, however, does the conclusion hold up under close scrutiny. If a lens has better resolution than the detector with which it is used, the lens designer might choose to use a larger numerical aperture even though that measure increases the aberrations and degrades the lens contribution to the spatial resolution. A gain in light-collection efficiency could then be achieved with negligible loss in overall resolution, and new applications of the system in low light levels could become possible. Similarly, in the nuclear medicine example, improved detector resolution might indeed be useless with conventional parallel-hole collimators, but new approaches to image formation with multiple pinholes or coded apertures could take advantage of the improved detector capability.

It is not an overstatement to say that increased technological capability in any component of an imaging system will always lead to improved task performance when the design of the overall system is approached in an integrated and creative fashion.

Integrated computational imaging systems In 2001 the Optical Society of America conducted a topical meeting called ICIS — Integrated Computational Imaging Systems. The premise is that image processing and image acquisition are becoming increasingly indistinguishable. It is no longer the case that a hardware designer develops a camera and some kind of computer interface while a software specialist develops ways of processing and displaying the data. The data-acquisition system is almost always under computer control, and the control signals are frequently de-

rived from the images themselves. Design of the hardware must take into account the needs and capabilities of the image analysis, and the nature of the processing can dictate what kind of data are acquired.

To a degree, this viewpoint is already used in the design of tomographic or other indirect imaging systems, where no image at all is obtained without processing, but there can still be a great symbiosis when image acquisition and processing are viewed as a whole and optimized in a coordinated way. As understanding of image quality advances, we can envision optimizing a system not for a class of objects but for a particular object; since we do not know what that object is before taking the image, we will have to modify and optimize the system during the image acquisition. The term *adaptive optics* can then take on a new meaning, adapting to the object being imaged and not just to corrupting influences such as the atmosphere.

To solve an inverse problem, concentrate on the forward problem Improved performance in an indirect imaging system can really come from only two sources: better data or better modeling of the data-acquisition system. The function of the reconstruction algorithm is only to narrow down the range of possible objects that are consistent with the data and the model of the system, and no algorithm can go beyond the limitations imposed by the data and the model.

To put this observation into practice, we need greatly improved models in many kinds of imaging. We need to move away from simplified system matrices with coarse voxels chosen purely for computational speed. In many applications we need accurate modeling of scattered and background radiation, and in imaging through turbid media we must move beyond the Born and Rytov approximations. We must avoid mathematical simplifications such as shift invariance and linearity when they are not justified. In optical systems, for example, we should account for shift-variant radiometric variations and off-axis aberrations as well as nonlinearities of the detector, and we should include the nonlinearities arising from partial coherence.

Achieving this level of accuracy in modeling will require careful system analysis and calibration and, of course, increased computer power, but it is the way to make progress in indirect imaging.

The last refuge Oscar Wilde said, "Consistency is the last refuge of the unimaginative." In inverse problems, however, consistency conditions are a largely untapped resource that can supplement incomplete or inaccurate data. As we saw in Chap. 15, consistency conditions can be derived by characterizing the range of an imaging operator, and in principle they can be used to reduce noise or correct for motion or other unknown characteristics of an imaging system.

Most known consistency conditions are based on continuous-to-continuous models of the imaging system, and most apply to forms of tomography where there are null functions of the adjoint operator. They thus apply to what one might call mathematical tomography rather than the real world of discrete, noisy data. A major theoretical challenge is to elucidate the relationship between mathematical tomography and real tomography and to understand what the continuous consistency conditions tell us about consistency of discrete data.

Tasks and observers A task-based assessment of image quality is useful in practice only if the task is meaningful. In Chaps. 16 – 19 we saw several examples where

oversimplified tasks led to misleading conclusions about system design. Progress in objective assessment will require more complicated and more realistic tasks. A key question concerns extrapolation of conclusions from simple tasks to more complicated ones. For example, will a system optimized for detection of specified nonrandom signals also be optimal for tasks where the possible outcomes are not defined in advance?

As simulation approaches reality Image simulations are becoming ever more realistic, largely because of the entertainment industry. These words are being written shortly after the telecast of Super Bowl XXXVII, the American football championship. Graphics shown during that game cause one to wonder whether Superbowl LXVII will be played with real athletes or with simulations.

How can image scientists take advantage of techniques developed at Lucas Films and Dreamworks? Accurate depiction of the motion of animate objects will advance our understanding of dynamic imaging systems, and accurate modeling of variations in surface reflectance will aid in design and analysis of optical systems for viewing opaque objects.

Accurate simulation of fine-scale or textural variations requires a detailed understanding of the underlying object statistics and how they influence the image statistics. We need these statistical descriptions for several purposes, including evaluation of image quality, development of Bayesian reconstruction algorithms, image synthesis and pattern recognition.

New models of object textures must be devised, along with experimental methods for estimating the parameters of the models from real image data. Since objects are infinite-dimensional but only a finite number of image samples will be available, parsimonious low-dimensional descriptions must be sought, and ways of assessing whether they capture the essential features of the object variability must be developed.

Our perception of perception We also need to understand better the human perceptual process and to integrate it into the design and evaluation of imaging systems, including processing algorithms and display. The role of image reconstruction and enhancement is to match the raw image data to the human perceptual system, and effective design and meaningful assessment of these elements of the imaging chain require exploring the links between image science and cognitive science. Inclusion of frequency-selective channels and internal noise in our observer models is a first, halting, step to bring knowledge from visual perception into image science, but much more sophisticated models are possible. How are we to account for the strong nonlinearities of the visual system? What can we learn from the study of visual illusions that will affect the design of imaging systems?

Man vs. machine Computerized image analysis is becoming more powerful and commonplace. How do we optimize an image-acquisition system when the end user is a computer? Which tasks are best performed by the computer and which by humans? When humans perform well, can we decipher their strategy and build it into an algorithm? What is the tradeoff between computerized and computer-aided image analysis, and can we be quantitative about answering this question? For example, what is a meaningful performance measure for a computer algorithm that segments an image and presents it to humans in cartoon form? When does the

cartoon enhance the ability of the human to grasp essential details (perform tasks) and when does it result in real information loss?

New technologies Although this book has dealt with the mathematics and physics of imaging, it has said very little about the technologies so essential to modern imaging. Some technological advances, such as optical and x-ray detector arrays, have been developed in direct response to the needs of imaging systems; others, such as optical data storage and computer displays, have been developed for more generic uses but have obvious applicability to imaging.

The greatest opportunity for creativity, however, comes from technologies developed in other fields that have no initially obvious connections to imaging. For example, lasers were an extension of masers, microwave amplifiers pursued initially for communications purposes. Superconducting magnets were developed long before their need in magnetic resonance imaging was apparent. And artificial radioisotopes were the result of wartime work on nuclear weapons, with no vision for their use in nuclear medicine. Yet in all of these cases the new technologies were quickly put to use in imaging systems.

Some emerging technologies with potential in imaging include femtosecond optical pulses, entangled photon states, diffractive and reflective x-ray optics, Josephson junctions and other superconducting devices, ultrasensitive seismometers and novel interferometers. Not to be overlooked are stunning advances in software in the areas of database management, artificial intelligence and data mining. A modern image scientist/technologist needs to be conversant with far more than the imaging literature and to remain alert to seemingly unrelated developments that can impact imaging.

New signals Science always searches for new ways of probing the universe, and increasingly the result of the probe is a multidimensional data set that can be manipulated and displayed as an image. One way to make progress in image science, or science in general, is to think up new things to map and new ways of probing them. Tables I and II in the Prologue should provide lots of hints in the search for new signals to image, but we can also look for new ways of applying ideas from image science to things we might not initially think of as images.

Sometimes ideas can come full circle, originating in another area, then being appropriated by image scientists, and finally being returned with embellishments to the original field. For example, in this book we have made considerable use of the Wigner distribution function (WDF) as a tool for signal and image analysis, although it originated in quantum mechanics. Now, however, there is considerable interest in imaging the quantum-mechanical Wigner function itself, and it turns out that tomographic reconstruction algorithms are the way to do it.

New dimensions We have long since passed the point where the word image implied a static, 2D construct. Modern imaging systems are almost always 3D ($x-y-z$ or $x-y-t$) or 4D ($x-y-z-t$), but there are many options for creative addition of yet other dimensions. The challenges lie in acquiring the data sets, doing high-dimensional image reconstruction and displaying the results.

To return to an example just given, the quantum-mechanical Wigner distribution function has been applied so far to 1D quantum states, such as the state of a single mode of the radiation field. If we want to measure the WDF of an

N-dimensional quantum state, it requires a $2N$-dimensional measurement/imaging system.

Similarly, if the quantity of interest is a statistical correlation function for a 2D field, it requires four variables for full specification, and correlations of 3D fields require six variables. There is no reason not to regard the correlation function, rather than the field itself, as the object to be imaged.

New pedagogies This book arose as an educational endeavor. It had its roots many years ago in a course on radiological imaging, and it eventually spawned several other courses on image science and noise. As the courses and the authors' educational and research interests evolved, it became apparent that the pedagogy of imaging was as challenging as the science and technology.

Students wanting to contribute at the cutting edge of image science must master a breadth of material comparable to — and perhaps even exceeding — the scope of mature disciplines such as chemistry and electrical engineering. They not only must study mathematics, statistics, physics and electronics, they must also understand how they interrelate and contribute to the whole gestalt of image science. They must be able to follow many diverse literatures and to pick and choose ideas and methodologies from them to solve their own problems in imaging. Like statisticians, they must appreciate the needs of their "clients," the end users of the images, but they must also be involved in optimizing the systems and developing the uses themselves.

Accommodating this breadth will require new departmental organizations, new interdisciplinary seminars and new professional societies. Academic departments and professional organizations devoted to optics and photonics will expand their base to include imaging, and umbrella interdisciplinary programs will arise to coordinate imaging activities across departments. Even social interactions among students and researchers perusing different aspects of the imaging elephant will help to broaden viewpoints and ease misconceptions.

We — the authors of this volume — have our own experiences in this educational process and our own thoughts about the critical educational needs in image science, and it is not a particularly difficult inverse problem to reconstruct those views from the contents of the book. We also realize, however, that optimization of the process is an ongoing challenge, where diverse views and imaginative approaches can be of immense value. We encourage a dialogue among educators and students in image science, and we hope that this book can stimulate that exchange in some small way. We look forward to observing the further development of image science as an academic discipline and making whatever contributions to it that we can.