

Automatic Target Recognition

Third Edition

Tutorial Texts Series

- *Applications of Lock-in Amplifiers in Optics*, Gerhard Kloos, Vol. TT117
- *Systems Engineering for Astronomical Telescopes*, Paul A. Lightsey and Jonathan W. Arenberg, Vol. TT116
- *Low-Level Light Therapy: Photobiomodulation*, Michael R. Hamblin, Ying-Ying Huang, Cleber Ferraresi, James D. Carroll, and Lucas Freitas de Freitas, Vol. TT115
- *Fiber Bragg Gratings: Theory, Fabrication, and Applications*, Marcelo M. Werneck, Regina C. Allil, and Fábio V. B. de Nazaré, Vol. TT114
- *Powering Laser Diode Systems*, Grigoriy A. Trestman, Vol. TT112
- *Optics Using MATLAB®*, Scott W. Teare, Vol. TT111
- *Plasmonic Optics: Theory and Applications*, Yongqian Li, Vol. TT110
- *Design and Fabrication of Diffractive Optical Elements with MATLAB®*, A. Vijayakumar and Shanti Bhattacharya, Vol. TT109
- *Energy Harvesting for Low-Power Autonomous Devices and Systems*, Jahangir Rastegar and Harbans S. Dhadwal, Vol. TT108
- *Practical Electronics for Optical Design and Engineering*, Scott W. Teare, Vol. TT107
- *Engineered Materials and Metamaterials: Design and Fabrication*, Richard A. Dudley and Michael A. Fiddy, Vol. TT106
- *Design Technology Co-optimization in the Era of Sub-resolution IC Scaling*, Lars W. Liebmann, Kaushik Vaidyanathan, and Lawrence Pileggi, Vol. TT104
- *Special Functions for Optical Science and Engineering*, Vasudevan Lakshminarayanan and L. Srinivasa Varadharajan, Vol. TT103
- *Discrimination of Subsurface Unexploded Ordnance*, Kevin A. O'Neill, Vol. TT102
- *Introduction to Metrology Applications in IC Manufacturing*, Bo Su, Eric Solecky, and Alok Vaid, Vol. TT101
- *Introduction to Liquid Crystals for Optical Design and Engineering*, Sergio Restaino and Scott Teare, Vol. TT100
- *Design and Implementation of Autostereoscopic Displays*, Byoung-ho Lee, Soon-gi Park, Keehoon Hong, and Jisoo Hong, Vol. TT99
- *Ocean Sensing and Monitoring: Optics and Other Methods*, Weilin Hou, Vol. TT98
- *Digital Converters for Image Sensors*, Kenton T. Veeder, Vol. TT97
- *Laser Beam Quality Metrics*, T. Sean Ross, Vol. TT96
- *Military Displays: Technology and Applications*, Daniel D. Desjardins, Vol. TT95
- *Interferometry for Precision Measurement*, Peter Langenbeck, Vol. TT94
- *Aberration Theory Made Simple, Second Edition*, Virendra N. Mahajan, Vol. TT93
- *Modeling the Imaging Chain of Digital Cameras*, Robert D. Fiete, Vol. TT92
- *Bioluminescence and Fluorescence for In Vivo Imaging*, Lubov Brovko, Vol. TT91
- *Polarization of Light with Applications in Optical Fibers*, Arun Kumar and Ajoy Ghatak, Vol. TT90
- *Digital Fourier Optics: A MATLAB Tutorial*, David G. Voeltz, Vol. TT89
- *Optical Design of Microscopes*, George Seward, Vol. TT88
- *Analysis and Evaluation of Sampled Imaging Systems*, Richard H. Vollmerhausen, Donald A. Reago, and Ronald Driggers, Vol. TT87
- *Nanotechnology: A Crash Course*, Raúl J. Martín-Palma and Akhlesh Lakhtakia, Vol. TT86
- *Direct Detection LADAR Systems*, Richard Richmond and Stephen Cain, Vol. TT85
- *Optical Design: Applying the Fundamentals*, Max J. Riedl, Vol. TT84
- *Infrared Optics and Zoom Lenses, Second Edition*, Allen Mann, Vol. TT83
- *Optical Engineering Fundamentals, Second Edition*, Bruce H. Walker, Vol. TT82
- *Fundamentals of Polarimetric Remote Sensing*, John Schott, Vol. TT81
- *The Design of Plastic Optical Systems*, Michael P. Schaub, Vol. TT80
- *Radiation Thermometry: Fundamentals and Applications in the Petrochemical Industry*, Peter Saunders, Vol. TT78
- *Matrix Methods for Optical Layout*, Gerhard Kloos, Vol. TT77
- *Fundamentals of Infrared Detector Materials*, Michael A. Kinch, Vol. TT76
- *Practical Applications of Infrared Thermal Sensing and Imaging Equipment, Third Edition*, Herbert Kaplan, Vol. TT75
- *Bioluminescence for Food and Environmental Microbiological Safety*, Lubov Brovko, Vol. TT74

(For a complete list of Tutorial Texts, see <http://spie.org/publications/books/tutorial-texts>.)

Automatic Target Recognition

Third Edition

Bruce J. Schachter

Tutorial Texts in Optical Engineering
Volume TT118

SPIE PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Names: Schachter, Bruce J. (Bruce Jay), 1946- author.

Title: Automatic target recognition / Bruce J. Schachter.

Other titles: Tutorial texts in optical engineering ; v. TT 118.

Description: Third edition. | Bellingham, Washington : SPIE Press, [2018] | Series: Tutorial texts in optical engineering ; volume TT 118 | Includes bibliographical references and index.

Identifiers: LCCN 2018005261 | ISBN 9781510618565 (softcover ; alk. paper) | ISBN 1510618562 (softcover) | ISBN 9781510618572 (pdf) | ISBN 1510618570 (pdf) | ISBN 9781510618589 (epub) | ISBN 1510618589 (epub) | ISBN 9781510618596 (Kindle/Mobi) | ISBN 1510618597 (Kindle/Mobi)

Subjects: LCSH: Radar targets. | Optical pattern recognition. | Algorithms. | Image processing.

Classification: LCC TK6580 .S33 2018 | DDC 623.7/348--dc23 LC record available at <https://lcn.loc.gov/2018005261>

Published by

SPIE

P.O. Box 10

Bellingham, Washington 98227-0010 USA

Phone: +1 360.676.3290

Fax: +1 360.647.1445

Email: books@spie.org

Web: <http://spie.org>

Copyright © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE)

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means without written permission of the publisher.

The content of this book reflects the work and thought of the author. Every effort has been made to publish reliable and accurate information herein, but the publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Printed in the United States of America.

First Printing.

For updates to this book, visit <http://spie.org> and type “TT118” in the search field.

SPIE.

Introduction to the Series

Since its inception in 1989, the Tutorial Texts (TT) series has grown to cover many diverse fields of science and engineering. The initial idea for the series was to make material presented in SPIE short courses available to those who could not attend and to provide a reference text for those who could. Thus, many of the texts in this series are generated by augmenting course notes with descriptive text that further illuminates the subject. In this way, the TT becomes an excellent stand-alone reference that finds a much wider audience than only short course attendees.

Tutorial Texts have grown in popularity and in the scope of material covered since 1989. They no longer necessarily stem from short courses; rather, they are often generated independently by experts in the field. They are popular because they provide a ready reference to those wishing to learn about emerging technologies or the latest information within their field. The topics within the series have grown from the initial areas of geometrical optics, optical detectors, and image processing to include the emerging fields of nanotechnology, biomedical optics, fiber optics, and laser technologies. Authors contributing to the TT series are instructed to provide introductory material so that those new to the field may use the book as a starting point to get a basic grasp of the material. It is hoped that some readers may develop sufficient interest to take a short course by the author or pursue further research in more advanced books to delve deeper into the subject.

The books in this series are distinguished from other technical monographs and textbooks in the way in which the material is presented. In keeping with the tutorial nature of the series, there is an emphasis on the use of graphical and illustrative material to better elucidate basic and advanced concepts. There is also heavy use of tabular reference data and numerous examples to further explain the concepts presented. The publishing time for the books is kept to a minimum so that the books will be as timely and up-to-date as possible. Furthermore, these introductory books are competitively priced compared to more traditional books on the same subject.

When a proposal for a text is received, each proposal is evaluated to determine the relevance of the proposed topic. This initial reviewing process has been very helpful to authors in identifying, early in the writing process, the need for additional material or other changes in approach that would serve to strengthen the text. Once a manuscript is completed, it is peer reviewed to ensure that chapters communicate accurately the essential ingredients of the science and technologies under discussion.

It is my goal to maintain the style and quality of books in the series and to further expand the topic areas to include new emerging fields as they become of interest to our reading audience.

*James A. Harrington
Rutgers University*

Contents

<i>Preface</i>	<i>xiii</i>
1 Definitions and Performance Measures	1
1.1 What is Automatic Target Recognition (ATR)?	1
1.1.1 Buyers and sellers	4
1.2 Basic Definitions	4
1.3 Detection Criteria	10
1.4 Performance Measures for Target Detection	13
1.4.1 Truth-normalized measures	13
1.4.1.1 Assigned targets and confusers (AFRL COMPASE Center terminology)	15
1.4.2 Report-normalized measure	15
1.4.3 Receiver operating characteristic curve	15
1.4.4 P_d versus FAR curve	18
1.4.5 P_d versus list length	19
1.4.6 Other factors that can enter the detection equation	19
1.4.7 Missile terminology	19
1.4.8 Clutter level	20
1.5 Classification Criteria	21
1.5.1 Object taxonomy	21
1.5.2 Confusion matrix	25
1.5.2.1 Compound confusion matrix	26
1.5.3 Some commonly used terms from probability and statistics	26
1.6 Experimental Design	29
1.6.1 Test plan	31
1.6.2 ATR and human subject testing	32
1.7 Characterizations of ATR Hardware/Software	33
References	34
2 Target Detection Strategies	35
2.1 Introduction	35
2.1.1 What is target detection?	36
2.1.2 Detection schemes	36
2.1.3 Scale	38

2.1.4	Polarity, shadows, and image form	38
2.1.5	Methodology for algorithm evaluation	40
2.1.5.1	Evaluation criteria for production systems	40
2.1.5.2	Target detection: machine versus human	41
2.2	Simple Detection Algorithms	41
2.2.1	Triple-window filter	41
2.2.2	Hypothesis testing as applied to an image	42
2.2.3	Comparison of two empirically determined means: variations on the T-test	43
2.2.4	Tests involving variance, variation, and dispersion	46
2.2.5	Tests for significance of hot spot	48
2.2.6	Nonparametric tests	49
2.2.6.1	Percent-bright tests	50
2.2.7	Tests involving textures and fractals	51
2.2.8	Tests involving blob edge strength	51
2.2.9	Hybrid tests	52
2.2.10	Triple-window filters using several inner-window geometries	53
2.3	More-Complex Detectors	53
2.3.1	Neural network detectors	53
2.3.2	Discriminant functions	54
2.3.3	Deformable templates	55
2.4	Grand Paradigms	55
2.4.1	Geometrical and cultural intelligence	56
2.4.2	Neuromorphic paradigm	57
2.4.3	Learning-on-the-fly	57
2.4.4	Integrated sensing and processing	58
2.4.5	Bayesian surprise	59
2.4.6	Modeling and simulation	60
2.4.7	SIFT and SURF	61
2.4.8	Detector designed to operational scenario	61
2.5	Traditional SAR and Hyperspectral Target Detectors	62
2.5.1	Target detection in SAR imagery	63
2.5.2	Target detection in hyperspectral imagery	64
2.6	Conclusions and Future Direction	66
	References	67
	Appendices	69
3	Target Classifier Strategies	77
3.1	Introduction	77
3.1.1	Parables and paradoxes	77
3.2	Main Issues to Consider in Target Classification	80
3.2.1	Issue 1: Concept of operations	81
3.2.2	Issue 2: Inputs and outputs	81
3.2.3	Issue 3: Target classes	82

3.2.4	Issue 4: Target variations	83
3.2.5	Issue 5: Platform issues	84
3.2.6	Issue 6: Under what conditions does a sensor supply useful data?	85
3.2.7	Issue 7: Sensor issues	86
3.2.8	Issue 8: Processor	87
3.2.9	Issue 9: Conveying classification results to the human-in-the-loop	87
3.2.10	Issue 10: Feasibility	88
3.3	Feature Extraction	92
3.4	Feature Selection	96
3.5	Examples of Feature Types	100
3.5.1	Histogram of oriented gradients	101
3.5.2	Histogram of optical flow feature vector	103
3.6	Examples of Classifiers	103
3.6.1	Simple classifiers	104
3.6.1.1	One-class classifiers	104
3.6.1.2	Two-class linear classifiers	104
3.6.1.3	Support vector machine	105
3.6.2	Basic classifiers	108
3.6.2.1	Single-nearest-neighbor classifier	108
3.6.2.2	Naïve Bayes classifier	110
3.6.2.3	Perceptron	111
3.6.2.4	Learning vector quantization family of algorithms	113
3.6.2.5	Feedforward multilayer perceptron trained with backpropagation of error	114
3.6.2.6	Mean-field theory networks	114
3.6.2.7	Model-based classifiers	116
3.6.2.8	Map-seeking circuits	116
3.6.2.9	Ensemble classifiers	118
3.6.3	Contest-winning and newly popular classifiers	118
3.6.3.1	Hierarchical temporal memory	120
3.6.3.2	Long short-term memory recurrent neural network	120
3.6.3.3	Convolutional neural network	121
3.6.3.4	Sentient ATR	124
3.7	Discussion	125
	References	131
4	Unification of Automatic Target Tracking and Automatic Target Recognition	135
4.1	Introduction	135
4.2	Categories of Tracking Problems	138
4.2.1	Number of targets	138
4.2.2	Size of targets	139

4.2.3	Sensor type	140
4.2.4	Target type	141
4.3	Tracking Problems	141
4.3.1	Point target tracking	141
4.3.2	Video tracking	146
4.3.2.1	Correlation tracking (video data)	147
4.3.2.2	Feature-vector-aided tracking (video data)	148
4.3.2.3	Mean-shift-based moving object tracker (video tracking)	149
4.4	Extensions of Target Tracking	149
4.4.1	Activity recognition (AR)	150
4.4.2	Patterns-of-life and forensics	152
4.5	Collaborative ATT and ATR (ATT $\leftrightarrow$ ATR)	153
4.5.1	ATT data useful to ATR	153
4.5.2	ATR data useful to ATT	154
4.6	Unification of ATT and ATR (ATT $\cup$ ATR)	155
4.6.1	Visual pursuit	155
4.6.2	A bat's echolocation of flying insects	157
4.6.3	Fused ATT $\cup$ ATR	158
4.6.3.1	Spatiotemporal target detection	158
4.6.3.2	Forecast of features and classes	161
4.6.3.3	Detection-to-track association	163
4.6.3.4	Track maintenance	163
4.6.3.5	Incorporation of higher-level knowledge	164
4.6.3.6	Implementation	165
4.7	Discussion	165
	References	166
5	Multisensor Fusion	169
5.1	Introduction	169
5.2	Critical Fusion Issues Related to ATR	172
5.3	Levels of Fusion	176
5.3.1	Data-level fusion	176
5.3.2	Feature-level fusion	184
5.3.2.1	Sensor selection for feature-level sensor fusion	184
5.4	Multiclassifier Fusion	187
5.4.1	Fusion of classifiers making hard decisions	189
5.4.1.1	Majority voting	189
5.4.1.2	Combined class rankings	190
5.4.1.3	Borda count method	190
5.4.1.4	Condorcet criterion	190
5.4.2	Fusion of classifiers making soft decisions	191
5.4.2.1	Simple Bayes average	191
5.4.2.2	Bayes belief integration	192

5.4.2.3	Trainable classifier as a combiner	193
5.4.2.4	Dempster–Shafer theory	193
5.5	Multisensor Fusion Based on Multiclassifier Fusion	194
5.6	Test and Evaluation	195
5.7	Beyond Basic ATR Fusion	196
5.7.1	Track fusion	196
5.7.2	Multifunction RF systems	199
5.7.2.1	Multifunction RF	199
5.7.2.2	Cognitive radar	200
5.7.3	Autonomous land vehicles	201
5.7.4	Intelligent preparation of the battlespace	202
5.7.4.1	Dissemination and integration	203
5.7.5	Zero-shot learning	204
5.7.5.1	Manned–unmanned teaming	205
5.7.5.2	Expert systems	207
5.8	Discussion	209
	References	210
6	Next-Generation ATR	215
6.1	Introduction	215
6.2	Hardware Design	216
6.2.1	Hardware recommendations for next-generation neuromorphic ATR	223
6.2.1.1	What <i>shouldn't</i> be copied from biology?	224
6.2.1.2	What <i>should</i> be copied from biology?	225
6.3	Algorithm/Software Design	228
6.3.1	Classifier architecture	229
6.3.1.1	Decision tree	229
6.3.1.2	Embodied and situated (ES)	232
6.3.1.3	Adaptivity and plasticity (PI)	233
6.3.2	Embodied, situated, plastic RNN [M = ES-PI-RNN(Q)] coupled with a controller C	233
6.3.2.1	Training the controller C	236
6.3.3	Software infrastructure	236
6.3.4	Test results	237
6.4	Potential Impact	237
	References	238
7	How Smart is Your Automatic Target Recognizer?	241
7.1	Introduction	241
7.2	Test for Determining the Intelligence of an ATR	243
7.2.1	Does the ATR understand human culture?	244
7.2.2	Can the ATR deduce the gist of a scene?	244
7.2.3	Does the ATR understand physics?	245

7.2.4	Can the ATR participate in a pre-mission briefing?	247
7.2.5	Does the ATR possess deep conceptual understanding?	247
7.2.6	Can the ATR adapt to the situation, learn on-the-fly, and make analogies?	248
7.2.7	Does the ATR understand the rules of engagement?	249
7.2.8	Does the ATR understand order of battle and force structure?	251
7.2.9	Can the ATR control platform motion?	252
7.2.10	Can the ATR fuse information from a wide variety of sources?	253
7.2.11	Does the ATR possess metacognition?	254
7.3	Sentient Versus Sapient ATR	255
7.4	Discussion: Where is ATR Headed?	256
	References	260
	Appendix 1: Resources	261
	Appendix 2: Questions to Pose to the ATR Customer	295
	Appendix 3: Acronyms and Abbreviations	299
	<i>Index</i>	305

Preface

An automatic target recognizer (ATR) is a real-time or near-real-time image/signal-understanding system. An ATR is presented with a stream of data. It outputs a list of the targets that it has detected and recognized in the data provided to it. A complete ATR system can also perform other functions such as image stabilization, preprocessing, mosaicking, target tracking, activity recognition, multi-sensor fusion, sensor/platform control, and data packaging for transmission or display.

In the early days of ATR, there were fierce debates between proponents of signal processing and those in the emerging field of computer vision. Signal processing fans were focused on more advanced correlation filters, stochastic analysis, estimation and optimization, transform theory, and time-frequency analysis of nonstationary signals. Advocates of computer vision said that signal processing provides some nice tools for our toolbox, but what we really want is an ATR that works as well as biological vision. ATR designers were less interested in processing signals than understanding scenes. They proposed attacking the ATR problem through artificial intelligence (AI), computational neuroscience, evolutionary algorithms, case-based reasoning, expert systems, and the like. Signal processing experts are interested in tracking point-like targets. ATR engineers want to track a target with some substance to it, identify what it is, and determine what activity it is engaged in. Signal processing experts keep coming up with better ways to compress video. ATR engineers want more intelligent compression. They want the ATR to tell the compression algorithm which parts of the scene are more important and hence deserving of more bits in the allocation. ATR, in and of itself, can be thought of as a data reduction technique. The ATR takes in a lot of data and outputs relatively little data. Data reduction is necessary due to bandwidth limitations of the data link and workload limits of the time-strapped human operator. People are very good at analyzing video until fatigue sets in or they get distracted. They don't want to be like the triage doctor at the emergency ward, assessing everything that comes in the door, continually assigning priorities to items deserving further attention. Pilots and ground station operators want a machine to relieve their burden as long as it rarely makes a mistake. Trying to do this keeps ATR engineers employed. As often told to the author, pilots and

image analysts are not looking for machines to replace them entirely. However, such decisions will be made higher up in the chain of command as ATR technology progresses.

The human vision system is not “designed” to analyze certain kinds of data such as rapid step-stare imagery, complex-valued signals that arise in radars, hyperspectral imagery, 3D LADAR data, or fusion of signal data with various forms of precise metadata. ATR shines when the sustained data rate is too high or too prolonged for the human brain, or the data is not well suited for presentation to humans. Nevertheless, most current ATRs operate with humans-in-the-loop. Humans, at present, are much better than ATRs at tasks requiring consultation, comprehension, and judgement. Humans still make the final decision and determine the action to be taken. This means that ATR output, which is statistical and multi-faceted by nature, has to be presented to the human decision makers in an easily understood form. This is a difficult man-machine interface problem. Marching toward the future, more autonomous robotic systems will necessarily rely more on ATRs to substitute for human operators, possibly serving as the “brains” of entire robotic platforms. We leave this provocative topic to the end of the book.

Systems engineers took notice once ATRs became deployable. Systems engineers are grounded in harsh reality. They care little about the debate between signal processing and computer vision. They don’t want to hear about an ATR being brain-like. They are not interested in which classification paradigm performs 1% better than the next. They care about the concept of operations (ConOps) and how it directs performance and functionality. They care about mission objectives and mission requirements. They want to identify all possible stakeholders, form an integrated product team, determine key performance parameters (KPPs), and develop test and evaluation (T&E) procedures to determine if performance requirements are met. Self-test is the norm for published papers and conference talks. Independent test and evaluation, laboratory blind tests, field tests, and software regression tests are the norm for determining if a system is deployable. The systems engineer’s focus is broader than ATR performance. Systems engineers want the entire system, or system of systems, to work well, including platform, sensors, ATR, and data links. They want to know what data can be provided to the ATR and what data the ATR can provide to the rest of the system. They want to know how one part of the system affects all other parts of the system. Systems designers care a lot about size, weight, power, latency, current and future costs, logistics, timelines, mean time between failure, and product repair and upgrade. They want to know the implications of system capture by the enemy.

At one time, ATR was the sole charge of the large defense electronics companies, working closely with the government labs. Only the defense companies and government have fleets of data collection aircraft, high-end sensors, and access to foreign military targets. Although air-to-ground has

been the focus of much ATR work, ATR actually covers a wide range of sensors, operating within or between the layers of space, air, ocean/land surface, and undersea/underground. Although the name ATR implies recognition of targets, ATR engineers have broader interests. ATR groups tackle any type of military problem involving the smart processing of imagery or signals. The government (or government-funded prime contractor) is virtually the only customer. So, some of the ATR engineer's time is spent reporting to the government, participating in joint data collections, taking part in government-sponsored tests, and proposing new programs to the government.

Since the 1960s, the field of ATR has advanced in parallel with similar work in the commercial sector and academia, involving industrial automation, medical imaging, surveillance and security, video analytics, and space-based imaging. Technologies of interest to both the commercial and defense sector include low-power processors, novel sensors, increased system autonomy, people detection, robotics, rapid search of vast amounts of data (big data), undersea inspection, and remote medical diagnosis. The bulk of funding in some of these areas has recently shifted from the defense to the commercial sector. More money is spent on computer animation for Hollywood movies than for the synthesis of forward-looking infrared (FLIR) and synthetic aperture radar (SAR) imagery. The search engine companies are investing much more in neural networks compared to the defense companies. Well-funded brain research programs are investigating the very basis of human vision and cognitive processing. The days of specialized military processors (e.g., VHSIC) are largely over. Reliance is now on chips in high-volume production: multi-core processors (e.g., Intel and ARM), FPGAs (e.g., Xilinx and Intel/Altera), and GPUs (e.g., Nvidia and AMD). Highly packaged sensors (visible, FLIR, LADAR, and radar) combined with massively parallel processors are advancing rapidly for the automotive industry to meet new safety standards (e.g., Intel/MobilEye). Millions of systems will soon be produced per year. Current advanced driver assistance systems (ADAS) can detect pedestrians, animals, bicyclists, road signs, traffic lights, cars, trucks, and road markers. These are a lot like ATR tasks. The rapid advancement of ADAS will lead to driverless cars.

Some important differences between ATRs and commercial systems are worth noting. ATRs generally have to detect and recognize objects at much longer ranges than commercial systems. Enemy detection and recognition are non-cooperative processes. Although a future car might have a LADAR, radar, or FLIR sensor, it won't have one that can produce high-quality data from a 20,000-ft range. An ADAS will detect a pedestrian but won't report if he is carrying a rifle. Search engine companies need to search large volumes of data with an image-based search, but they don't have the metadata to help the search, such as is available on military platforms. That being said, the cost

and innovation rate of commercial electronics can't be matched by military systems. The distinction between commercial and military systems is starting to blur in some instances. Cell phones now include cameras, inertial measurement units, GPS, computers, algorithms, and transmitters/receivers. Slightly rugged versions of commercial cell phones and tablet computers are starting to be used by the military, even with ATR apps. "Toy" drones are approaching the sophistication of the smallest military unmanned air vehicles. They are now produced in volumes of a million per year. ATR engineers are in tune with advances in the commercial sector and their applicability to ATR. Even their hobbies tend to focus on technology, e.g., hobbies such as quadcopters, novel cameras, 3D printers, computers, phone apps, robots, etc.

ATR is not limited to a device; it is also a field of research and development. ATR technology can be incorporated into systems in the form of self-contained hardware, FPGA code, or higher-level language code. ATR groups can help add autonomy to many types of systems. ATR can be viewed very narrowly or very broadly, borrowing concepts from a wide variety of fields. Papers on ATR are often of the form: "Automatic Target Recognition using XXX," where the XXX can be any technology such as super-resolution, principal component analysis, sparse coding, singular value decomposition, Eigen templates, correlation filters, kinematic priors, adaptive boosting, hyperdimensional manifolds, Hough transforms, foveation, etc. In the more ambitious papers, the XXX is a *mélange* of technologies, such as fuzzy-rule-based expert systems, wavelet neural genetic networks, fuzzy morphological associative memory, optical holography, deformable wavelet templates, hierarchical support vector machines, Bayesian recognition by parts, etc. Get the picture? Nearly any type of technology, everything but the kitchen sink, can be thrown at the ATR problem, with scant large-scale independent competitive test results to indicate which approach really works best, supposing that "best" can be defined and measured. This book is not a comprehensive survey of every technology that has ever been applied to ATR. This book covers some of the basics of ATR. While some of the topics in this book can be found in textbooks on pattern recognition and computer vision, this book focuses on their application to military problems as well as the unique requirements of military systems.

The topics covered in the book are organized in the way one would design an ATR. The first step is to understand the military problem and make a list of potential solutions to the problem. A key issue is the availability of sufficiently comprehensive sets of data to train and test the potential solutions. This involves developing a sound test plan, specifying procedures and equations, and determining who is going to do the testing. Testing isn't open ended. Exit criteria are needed to determine when a given test activity has been successfully completed. The next steps in ATR design are choosing the detector and classifier. The detector focuses attention on the regions-of-interest

in the imagery requiring additional scrutiny. The classifier further processes these regions-of-interest and is the decision engine for class assignment. It can operate at any or all levels of a decision tree, from clutter rejection to identifying a specific vehicle or activity. Detected targets are often tracked. Target tracking has historically been treated as a separate subject from ATR, mainly because point-like targets contain too little information to apply an ATR. However, as sensor resolution improves, the engineering disciplines of target tracking and ATR are starting to merge. The ATR and tracker can be united for efficiency and performance. The fifth chapter covers the basics of multisensor fusion. Then it broadens the topic to a variety of other forms of fusion. A strawman design is provided for a more advanced ATR, but with no claim that this is the only way to construct a next-generation ATR. The strawman design should be thought of as a brainstormed simple draft proposal intended to generate discussion of its advantages and disadvantages, and to trigger the generation of new and better proposals. Future ATRs will have to combine data from multiple sources. The last chapter points out how primitive current ATRs really are, as compared to biological systems. It suggests ways for measuring the intelligence of an ATR. This goes far beyond the basic performance measurement techniques covered in Chapter 1. The first appendix lists the many resources available to the ATR engineer. Many of the listed agencies supply training and testing data, perform blind tests, and sponsor research into compelling new sensor and ATR designs. The second appendix advances the notion that a problem that is well described is half solved. The third appendix explains the acronyms and abbreviations used in the book.

CHAPTER 1: ATR technology has benefited from a significant investment over the last 50 years. However, the once-accepted definitions and evaluation criteria have been displaced by the march of technology. The first chapter updates the language for describing ATR systems and provides well-defined criteria for evaluating such systems. This will move forward collaboration between ATR developers, evaluators, and end-users.

ATR is used as an umbrella term for a broad range of military technology beyond just the recognition of targets. In a more general sense, ATR means *sensor data exploitation*. Two types of definitions are included in the first chapter. One type defines fundamental concepts. The other type defines basic performance measures. In some cases, definitions consist of a list of alternatives. This approach enables choices to be made to meet the needs of particular programs. The important point to keep in mind is that within the context of a particular experimental design, a set of protocols should be adopted to best fit the situation, applied, and then kept constant throughout the evaluation. This is especially important for competitive testing.

The definitions given in Chapter 1 are intended for evaluation of end-to-end ATR systems as well as the prescreening and classifier stages of the systems.

Sensor performance and platform characteristics are excluded from the evaluation. It is recognized that sensor characteristics and other operational factors affect the imagery and associated metadata. A thorough understanding of data quality, integrity, synchrony, availability, and timeline are important for ATR development, test, and evaluation. Data quality should be quantified and assessed. However, methods for doing so are not covered in this book. The results and validity of ATR evaluation depend on the representativeness and comprehensiveness of the development and test data. The adequacy of development and test data is primarily a budgetary issue. The ATR engineer should understand and be able to convey the implications of limited, surrogate, or synthetic data. The ATR engineer should be able to damp down naïve proposals centered around the use of an off-the-shelf deep-learning neural network as a miraculous cure to the alleged ATR affliction.

Chapter 1 formalizes definitions and performance measures associated with ATR evaluation. All performance measures must be accepted as ballpark predictions of actual performance in combat. More carefully formulated experiments will provide more meaningful conclusions. The final measure of effectiveness takes place in the battlefield.

CHAPTER 2: Hundreds of simple target detection algorithms were tested on mid- and longwave FLIR images, as well as X-band and Ku-band SAR images. Each algorithm is briefly described. Indications are given as to which performed well. Some of these simple algorithms are loosely derived from standard tests of the difference of two populations. For target detection, these are typically populations of pixel grayscale values or features derived from them. The statistical tests are often implemented in the form of sliding triple-window filters. Several more-elaborate algorithms are also described with their relative performances noted. These algorithms utilize neural networks, deformable templates, and adaptive filtering. Algorithm design issues are broadened to cover system design issues and concepts of operation.

Since target detection is such a fundamental problem, it is often used as a test case for developing technology. New technology leads to innovative approaches for attacking the problem. Eight inventive paradigms, each with deep philosophical underpinnings, are described in relation to their effect on target detector design.

CHAPTER 3: Target classification algorithms have generally kept pace with developments in the academic and commercial sectors since the 1970s. However, most recently, investment into object classification by Internet companies and various large-scale projects for understanding the human brain has far outpaced that of the defense sector. The implications are noteworthy.

There are some unique characteristics of the military classification problem. Target classification is not solely an algorithm design problem, but is part of a larger system design task. The design flows down from a

ConOps and KPPs. Required classification level is specified by contract. Inputs are image and/or signal data and time-synchronized metadata. The operation is often real-time. The implementation minimizes size, weight, and power (SWaP). The output must be conveyed to a time-strapped operator who understands the rules of engagement. It is assumed that the adversary is actively trying to defeat recognition. The target list is often mission dependent, not necessarily a closed set, and can change on a daily basis. It is highly desirable to obtain sufficiently comprehensive training and testing data sets, but costs of doing so are very high, and data on certain target types are scarce or nonexistent. The training data might not be representative of battlefield conditions, suggesting the avoidance of designs tuned to a narrow set of circumstances. A number of traditional and emerging feature extraction and target classification strategies are reviewed in the context of the military target classification problem.

CHAPTER 4: The subject being addressed is how an automatic target tracker (ATT) and an ATR can be fused so tightly and so well that their distinctiveness becomes lost in the merger. This has historically not been the case outside of biology and a few academic papers. The biological model of ATTUATR arises from dynamic patterns of activity distributed across many neural circuits and structures (including those in the retinae). The information that the brain receives from the eyes is “old news” at the time that it receives it. The eyes and brain forecast a tracked object’s future position, rather than relying on the perceived retinal position. Anticipation of the next moment—building up a consistent perception—is accomplished under difficult conditions: motion (eyes, head, body, scene background, target) and processing limitations (neural noise, delays, eye jitter, distractions). Not only does the human vision system surmount these problems, but it has innate mechanisms to exploit motion in support of target detection and classification. Biological vision doesn’t normally operate on snapshots. Feature extraction, detection, and recognition are spatiotemporal. When scene understanding is viewed as a spatiotemporal process, target detection, target recognition, target tracking, event detection, and activity recognition (AR) do not seem as distinct as they are in current ATT and ATR designs. They appear as similar mechanisms taking place at varying time scales. A framework is provided for unifying ATT, ATR, and AR.

CHAPTER 5: Predatory animals detect, stalk, recognize, track, chase, home in on, and if lucky, catch their prey. Stereo vision is generally their most important sensor asset. Most predators also have a good sense of hearing. Some predators can smell their prey from a mile away. Most creatures combine data from multiple sensors to eat or avoid being eaten. Different creatures use different combinations of sensors, including sensors that detect vibration, infrared radiation, various spectral bands, polarization, Doppler, and magnetism.

Biomimicry suggests that a combination of diverse sensors works better than use of a single sensor type. Sensor fusion intelligently combines sensor data from disparate sources such that the resulting information is in some ways superior to the data from a single source. Chapter 5 provides techniques for low-level, mid-level, and high-level information fusion. Other forms of fusion are also of interest to the ATR engineer. Multifunction fusion combines functions normally implemented by separate systems into a single system. Zero-shot learning (ZSL) is a way of recognizing a target without having trained on examples of the target. ZSL provides a vivid description of a detected target as a fusion of its semantic attributes. The commercial world is embracing multisensor fusion for driverless cars. New sensor and processor designs are emerging with applicability to autonomous military vehicles.

CHAPTER 6: Traditional feedforward neural networks, including multilayer perceptrons (MLPs) and the newly popular convolutional neural networks (CNNs), are trained to compute a function that maps an input vector to an output vector. The N -element output vector can convey estimates of the probabilities of N target classes. Nearly all current ATRs perform target classification using feedforward neural networks. These can be shallow or deep. The ATR detects a candidate target, transforms it to a feature vector, and then processes the vector unidirectionally, step by step; the number of steps is proportional to the number of layers in the neural network. Signals travel one way from input to output. A recurrent neural network (RNN) is an appealing alternative. Its neurons send feedback signals to each other. These feedback loops allow RNNs to exhibit dynamic temporal behavior. The feedback loops also establish a type of internal memory. While feedforward neural networks are generally trained in a supervised fashion by backpropagation of output error, RNNs are trained by backpropagation through time.

Although feedforward neural networks are said to be inspired by the architecture of the brain, they do not model many abilities of the brain, such as natural language processing and visual processing of spatiotemporal data. Feedback is omnipresent in the brain, endowing both short-term and long-term memory. The human brain is thus an RNN—a network of neurons with feedback connections. It is a dynamical system. The brain is plastic, adapting to the current situation. The human vision system not only learns patterns in sequential data, but even processes still frame (snapshot) data quite well with its RNN, jerking the eyes in saccades to shift focus over key points on a snapshot, turning the snapshot into a movie.

An improved type of RNN, called long short-term memory (LSTM), was developed in the 1990s by Jürgen Schmidhuber and his former Ph.D. student Sepp Hochreiter. LSTM and its many variants are now the predominant RNN. LSTM is said to be in use in billions of commercial devices.

Brains don't come in a box like a desktop computer or supercomputer. All natural intelligence is embodied and situated. Many military systems, such as unmanned air vehicles and robot ground vehicles, are embodied and situated. The body (platform) maneuvers the sensor systems to view the battlespace from different situations. An ATR based on an RNN, that is embodied and situated [ES], adaptive and plastic [PI], and of limited precision (e.g., 16-bit floating point), will be denoted by the model $\mathbf{M}=\text{ES-PI-RNN}(\mathbb{Q}_{16})$. A recurrent ATR is more powerful in many ways than a standard ATR. Both computationally more powerful and biologically more plausible than other types of ATRs, an RNN-based ATR understands the notion of events that unfold over time. Its design can benefit from ongoing advances in neuroscience.

Professor Schmidhuber has made an additional improvement to his model. He tightly couples a controller $\mathbf{C}$ to a model $\mathbf{M}$. Both can be RNNs or composite designs incorporating RNNs. Following Schmidhuber's lead, we propose a strawman ATR that couples a controller $\mathbf{C}$ to our model $\mathbf{M}=\text{ES-PI-RNN}(\mathbb{Q}_{16})$ to form a complete system ($\mathbf{C} \cup \mathbf{M}$) that is more powerful in many ways than a standard ATR. $\mathbf{C} \cup \mathbf{M}$ can learn a never-ending sequence of tasks, operate in unknown environments, realize abstract planning and reasoning, perform experiments, and retrain itself on-the-fly. This next-generation ATR is suitable for implementation on two chips: a single custom low-power chip (<1 W) for effecting $\mathbf{M}$, hosted by a standard processor serving as the controller $\mathbf{C}$. A heterogeneous chip design incorporating high-speed I/O, multicore ARM processors, logic gates, GPU, codec, and neural section is also appropriate. This next-generation ATR is applicable to various military systems, including those with extreme size, weight, and power constraints.

CHAPTER 7: ATRs have been under development since the 1960s. Advances in computer processing, computer memory, and sensor resolution are easy to evaluate. However, the time horizon of the truly smart ATR seems to be receding at a rate of one year per year. One issue is that there has never been a way to measure the intelligence of an ATR. This is fundamentally different from measuring detection and classification performance. The description of what constitutes an ATR, and in particular a smart ATR, keeps changing. Early ATRs did little more than detect fuzzy bright spots in first-generation FLIR video or ten-foot-resolution SAR data. Sensors are getting better, computers are getting faster, and the ATR is expected to take over more of the workload. With unmanned systems there is no human onboard to digest information. The ATR is compelled to transmit only the most important information over a limited-bandwidth data link. The ATR or robotic system can be viewed as a substitute for a human. What constitutes intelligence in artificial humans has long been debated, starting with stories of golems,

continuing to the Turing test, and including current dire predictions of super-intelligent robots superseding humans. Chapter 7 provides a Turing-like test for judging the intelligence of an ATR.

APPENDIX 1: The first appendix lists the many resources available to the ATR engineer and includes a brief historical overview of the technologies involved in ATR development.

APPENDIX 2: A successful project starts with a clear description of the problem to be solved. However, a well-defined ATR problem is surprisingly hard to come by. The second appendix provides some questions to pose to a customer to help get a project going.

APPENDIX 3: The third appendix defines all of the acronyms and abbreviations used in this book.

Acknowledgments

Special thanks to the United States Army Night Vision and Electronic Sensors Directorate (NVESD), Air Force, Navy, DARPA, and Northrop Grumman for supporting this work over the years. This book benefited from critique and suggestions made by the reviewers and SPIE staff.

Author's Disclaimer

The views and opinions expressed in this book are solely those of the author in his private capacity and do not represent those of any company, the United States Federal Government, any entity of the U.S. Federal Government, or any private organization. Links to organizations are provided solely as a service to our readers. Links do not constitute an endorsement by any organization or the Federal Government, and none should be inferred. While extensive efforts have been made to verify statements and facts presented in this book, any factual errors or errors of opinion are solely those of the author. No position or endorsement by the U.S. Federal Government, any entity of the Federal government, or any other organization regarding the validity of any statement of fact presented in this book should be inferred.

Author's Contact Information

Comments on this book are welcome. The author can be contacted at Bruce.Jay.Schachter@gmail.com.

Bruce J. Schachter
February 2018