

Physical One-Way Functions

by

Pappu Srinivasa Ravikanth

BSEE, Osmania University, 1991

MSEE, Villanova University, 1993

MS in Media Arts and Sciences, MIT, 1995

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

Massachusetts Institute of Technology

March 2001

© MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2001. ALL RIGHTS
RESERVED

Author _____

Pappu Srinivasa Ravikanth

Program in Media Arts and Sciences

Certified by _____

Neil A. Gershenfeld

Associate Professor of Media Arts and Sciences

Program in Media Arts and Sciences

Accepted by _____

Stephen A. Benton

Chair, Departmental Committee on Graduate Students

Program in Media Arts and Sciences

Physical One-Way Functions - Abstract

by

Pappu Srinivasa Ravikanth

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on March 2, 2001 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Abstract

Modern cryptography relies on algorithmic one-way functions - numerical functions which are easy to compute but very difficult to invert. This dissertation introduces *physical one-way functions* and *physical one-way hash functions* as primitives for physical analogs of cryptosystems.

Physical one-way functions are defined with respect to a physical probe and physical system in some unknown state. A function is called a physical one-way function if (a) there exists a deterministic physical interaction between the probe and the system which produces an output in constant time (b) inverting the function using either computational or physical means is difficult (c) simulating the physical interaction is computationally demanding and (d) the physical system is easy to make but difficult to clone.

Physical one-way hash functions produce fixed-length output regardless of the size of the input. These hash functions can be obtained by sampling the output of physical one-way functions. For the system described below, it is shown that there is a strong correspondence between the properties of physical one-way hash functions and their algorithmic counterparts. In particular, it is demonstrated that they are *collision-resistant* and that they exhibit the *avalanche effect*, i.e., a small change in the physical system causes a large change in the hash value.

An inexpensive prototype authentication system based on physical one-way hash functions is designed, implemented, and analyzed. The prototype uses a disordered three-dimensional microstructure as the underlying physical system and coherent radiation as the probe. It is shown that the output of the interaction between the physical system and the probe can be used to robustly derive a unique tamper-resistant identifier at a very low cost per bit. The explicit use of three-dimensional structures marks a departure from prior efforts. Two protocols, including a one-time pad protocol, that illustrate the utility of these hash functions are presented and potential attacks on the authentication system are considered.

Finally, the concept of *fabrication complexity* is introduced as a way of quantifying the difficulty of materially cloning physical systems with arbitrary internal states. Fabrication complexity is discussed in the context of an idealized machine - a Universal Turing Machine augmented with a fabrication head - which transforms algorithmically minimal descriptions of physical systems into the systems themselves.

Thesis supervisor: Neil A. Gershenfeld

Title: Associate Professor of Media Arts and Sciences, Program in Media Arts and Sciences

Doctoral Dissertation Committee

The following people served on the dissertation committee.

Neil A. Gershenfeld

Associate Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Joseph M. Jacobson

Associate Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Nabil M. Amer

Senior Manager, Physical Science Department
IBM Watson and Almaden Research Centers

Daniel R. Simon

Cryptographer
Microsoft Research

I have often pondered over the roles of knowledge or experience, on the one hand, and imagination or intuition, on the other, in the process of discovery. I believe that there is a certain fundamental conflict between the two, and knowledge, by advocating caution, tends to inhibit the flight of imagination. Therefore, a certain naiveté, unburdened by conventional wisdom, can sometimes be a positive asset.

Harish Chandra
1923-1983

Information is physical.

Rolf Landauer
1927-1999

Acknowledgements

Many extraordinary people contributed to this work in extraordinary ways.

I want to thank

Neil Gershenfeld, my thesis advisor, for his vision, energy, enthusiasm, and inspiration. It has been an absolute pleasure to work with him.

Nabil Amer, Dan Simon, and Joe Jacobson for being on my dissertation committee, and giving me some very useful suggestions and advice. I would especially like to thank Nabil for nominating me for an IBM research fellowship during the last two years of my tenure at the Media Lab.

The people in the Bits and Atoms community at the Media Lab: Rehmi Post, Bernd Schoner, Yael Maguire, Matt Reynolds, Ben Vigoda, Josh Smith, Ben Recht, Isaac Chuang, Scott Manalis, Saul Griffith, Jason Taylor, John DeFrancesca, H. Shrikumar, Aggelos Bletsas, Rich Fletcher, Olufemi Omojola, John-Paul Strachan, Aram Harrow, Karen Robinson, Esa Masood, Peter Russo, and others. They are an amazing group of people and I learn from them everyday.

Linda Peterson, Susan Bottari, and Liz Hennessey for keeping the MIT academic and administrative dragons at bay.

My friends: Raji, Sujal, Aakash, Anu, Rama, Shanti, Shami, Gaddam, Harish, Bindu, Sudhir, Kishan, Jayadev, Ravindran, Barrett....

Wendy Plesniak, for being my closest friend in the world, for really broadening my horizons, and for having the clearest perspective on the things in life that really matter. Thank you Wen!!!

Finally, I want to thank my family: my father Purushottama Rao, my mother Chayalaxmi, my brothers Vivek and Kartik, and my sister Gayatri for their encouragement, inspiration, good humour, and their undying faith in me.

My parents made many sacrifices so all of us could have the best possible education. I am deeply grateful to them for putting our education ahead of their own comfort. I dedicate this work to them.

Contents

Physical One-Way Functions - Abstract	3
Doctoral Dissertation Committee	5
Acknowledgements	9
1 Introduction	15
1.1 Algorithmic one-way functions	15
1.1.1 The RSA function	15
1.1.2 The Rabin function	16
1.2 The concept of physical one-way functions	16
1.3 Intellectual inspiration	17
1.4 Motivation	17
1.4.1 Authenticating bits with monetary value	17
1.4.2 Silicon-based authentication is not inexpensive enough	18
1.4.3 Asymmetry between 2D and 3D microfabrication	19
1.4.4 Connection between physical systems and cryptography	20
1.5 Research goals	20
1.6 Organization of the dissertation	20
2 Preliminaries	25
2.1 One-way functions	25
2.1.1 The origin of OWFs and OWHFs	25
2.1.2 Formal definitions	26
2.2 Authentication and digital signatures	28
2.2.1 Definitions	28
2.2.2 How one-way hash functions are used in digital signatures	28
2.2.3 Attacks on one-way hash functions	29
2.3 Quantum money	31
2.3.1 Photon polarization	31
2.3.2 Quantum measurement of photons	31
2.3.3 Preparing the quantum banknote	33
2.3.4 Forging a quantum banknote	34
2.3.5 Discussion	34
2.4 Algorithmic and computational complexity	35
2.4.1 Problem size	35
2.4.2 Asymptotic notation	36
2.4.3 Complexity classes	36
2.5 Complexity in physical systems	38
2.5.1 Candidate metrics	38
2.5.2 Kolmogorov complexity	39
2.5.3 Summary	42
3 Related work	43
3.1 Prior art in physical authentication	43
3.1.1 Optically variable devices	43
3.1.2 Authentication using random features	44
3.2 Summary	47

4	Concept, design choices, and problem formulation	49
4.1	System concept and data pipeline:	49
4.2	Requirements of each system component	50
4.2.1	Physical system requirements	50
4.2.2	Requirements for the probe	50
4.2.3	Detector requirements	51
4.2.4	Interaction between physical system and probe	51
4.3	Design choices	51
4.4	Problem formulation	52
4.4.1	System concept	52
4.4.2	System theory and performance	52
4.4.3	Attacks and spoofing	53
4.4.4	Cryptographic framework and future work	53
5	Light transport through disordered media	55
5.1	Assumptions and notation	55
5.2	Length scales and scattering regimes	56
5.3	Coherent multiple scattering	57
5.3.1	Classical speckle theory	57
5.3.2	Born again: the memory effect	58
5.3.3	Experimental observation of the memory effect	59
5.3.4	The C1, C2, and C3 correlations	60
5.3.5	C1, C2, and C3: an engineering view	65
5.3.6	Speckle sensitivity	66
5.3.7	The random matrix formalism	68
5.4	Light transport through nonlinear media	71
5.4.1	The approach	71
5.4.2	Engineering issues	72
5.5	Optical localization	72
5.6	A summary of key ideas	73
6	Theory of physical one-way (hash) functions	75
6.1	Computational one-way functions	75
6.2	General definition of physical one-way functions	76
6.2.1	Definitions	76
6.2.2	Discussion of the definition	77
6.3	Coherent multiple scattering implements a POWF	78
6.3.1	Notation	78
6.3.2	POHFs as sampled speckle patterns	79
6.4	Heuristic arguments	79
6.4.1	Easy to “compute”	79
6.4.2	Hard to invert	80
6.4.3	Simulating the output	84
6.4.4	High-sensitivity	84
6.4.5	Cloning the structure	84
6.5	Summary	84
7	System design and engineering	85
7.1	What needs to be designed?	85
7.2	Token design	85

7.2.1	Creating the microstructure	85
7.2.2	Making the token	85
7.3	Probe design	86
7.3.1	Optical coherence tomography (OCT)	87
7.3.2	Magnetic resonance imaging (MRI)	88
7.3.3	Laser beam	90
7.4	Reader design	91
7.4.1	Mechanical requirements	91
7.4.2	Implementation	92
7.4.3	Performance	95
7.5	The Gabor hash algorithm	99
7.5.1	Desired features	99
7.5.2	Theory of Gabor Transforms	100
7.5.3	Implementation to derive unique identifier	101
7.5.4	An example	103
7.5.5	Tradeoffs	105
7.6	Final system	107
7.7	Potential improvements of the system	107
8	Experiments and results	109
8.1	Proof-of-concept experiment	109
8.1.1	The setup	109
8.1.2	Results	109
8.2	Statistics of Gabor Hash strings	111
8.2.1	The setup	111
8.2.2	Statistical results	111
8.3	Demonstration of tamper resistance	119
8.3.1	The setup	119
8.3.2	Results	119
8.4	Summary	120
9	Protocols	123
9.1	A bit of history	123
9.2	One-time pad protocol	124
9.2.1	Motivation	124
9.2.2	Augmenting the card and terminal	125
9.2.3	Nonlinearity in the microstructure	125
9.2.4	Assertions	125
9.2.5	Notation	126
9.2.6	The protocol for trusted terminals	126
9.2.7	The one-time pad protocol for untrusted terminals	127
9.3	Bit commitment	128
9.3.1	Background	128
9.3.2	The bit-commitment protocol	129
9.4	Summary	130
10	Scaling, attacks, and fabrication complexity	131
10.1	Scaling issues	131
10.1.1	Scaling the size of the physical structure	131
10.1.2	Scaling the number of tokens	132

10.2	Brute-force and birthday attacks	134
10.3	Replay attacks	135
10.4	Fabrication methods	136
10.4.1	Photolithography	136
10.4.2	Electron beam lithography	137
10.4.3	Scanned probe lithography	137
10.4.4	Summary	137
10.5	Fabrication complexity	138
10.5.1	Notation	138
10.5.2	Problem definition	138
10.5.3	A Universal Fabrication Machine (UFM)	138
10.5.4	Kolmogorov complexity of disordered structures	139
10.5.5	Physical resources used in fabrication	140
10.6	Parallel fabrication attack	140
10.7	Summary	141
11	Contributions and future work	143
11.1	Summary and original contributions	143
11.2	Future work	145
12	References	149

1 Introduction

This dissertation introduces *physical one-way functions* and *physical one-way hash functions* as primitives for physical cryptography.

Physical one-way functions are defined with respect to a physical probe and physical system in some unknown state. A function is called a physical one-way function if (a) there exists a deterministic physical interaction between the probe and the system which produces an output in constant time (b) inverting the function using either computational or physical means is difficult (c) simulating the physical interaction is computationally demanding and (d) the physical system is easy to make but difficult to clone. Physical one-way hash functions produce a fixed-length output regardless of the size of the input.

In this chapter, we briefly look at commonly used algorithmic one-way functions, namely the RSA and Rabin functions, before proceeding to present the concept of physical one-way functions in section 1.2. The intellectual inspiration for this work, Quantum Money, is briefly presented in section 1.3. The reasons for studying physical one-way functions are discussed in section 1.4. Finally, in the concluding section, we provide a detailed roadmap of this document.

1.1 Algorithmic one-way functions

Modern asymmetric cryptography rests squarely on the shoulders of one-way functions — numerical functions which are easy to compute but difficult to invert. This asymmetry is naturally embodied in a real-world security concern: it should be easy for a legitimate user to operate a cryptosystem but infeasible for an adversary to foil it. This gap in complexity of effort between legitimate users and adversaries lies at the heart of cryptography. However, *it is not known if one-way functions exist* [11]. Despite this, however, there are several important results from the literature which are predicated on the existence of one-way functions. Examples of such results include: *one-way functions are necessary and sufficient for secure signatures* [15] and *any one-way function can be used to construct a pseudorandom generator* [16].

Algorithmic one-way functions, as we currently know them, are mathematical objects which are based on (conjectured) intractable problems. Let us consider two examples which are based on the intractability of integer factorization: it is difficult to factorize a number which is the product of two prime numbers of comparable length.

1.1.1 The RSA function

This is a family of one-way functions named after its inventors Rivest, Shamir, and Adleman. Let P and Q be two prime numbers such that $|\log P - \log Q| \leq 1$, $N = PQ$, and let e be an integer smaller than N and relatively prime to $\phi(N) = (P-1)(Q-1)$. Then the RSA function is defined over the domain $\{1, \dots, N\}$ as

$$RSA_{N,e}(x) = x^e \bmod(N) \quad 1.1.1$$

It is widely believed that inverting $RSA_{N,e}(x)$ is intractable given an (N, e) pair but not the factors P and Q .

1.1.2 The Rabin function

The Rabin function is defined in a very similar way except the function is defined by

$$RABIN_N(x) = x^2 \bmod(N) \quad 1.1.2$$

It can be shown that inverting $RABIN_N(x)$, i.e., finding square roots $\bmod(N)$, is computationally equivalent to factoring N , which is conjectured to be intractable.

1.2 The concept of physical one-way functions

We introduce *physical one-way functions* (POWFs) in this dissertation. Unlike algorithmic one-way functions, which are mathematical objects, physical one-way functions are defined as interactions between physical systems and physical probes. Specifically, a physical one-way function requires

- A physical system with some unknown internal state
- A physical probe
- An interaction between the probe and the system

For physical one-way functions the hard problems are the *difficulty of cloning a physical system with a specific internal state* and *efficiently simulating the interaction between the probe and the physical system*.

There are two asymmetries present in our conceptual picture of a system that employs physical one-way functions. They are:

- The physical system is easy to make but difficult to clone.
- The interaction between the probe and the physical system produces an output quickly, but computationally simulating this interaction is difficult.

In this work, we use *disordered three-dimensional microstructures* as the physical system and *coherent radiation* as the probe. The output of the interaction between the probe and the 3D microstructure is called a speckle pattern and is a very complicated fingerprint of the structural details of the microstructure. The physical mechanism of speckle generation is called *coherent multiple scattering*. We use speckle patterns to generate unique and tamper-resistant identifiers for 3D structures.

Algorithmic hash functions produce a fixed-length output regardless of the

length of the input. We introduce *physical one-way hash functions* in this dissertation. We show that physical one-way hash functions can be obtained by sampling the output of a POWF, i.e., we sample the output speckle patterns on a regular grid to produce a fixed size output. For the system described above, it is shown that there is a strong correspondence between the properties of physical one-way hash functions and their algorithmic counterparts. In particular, it is demonstrated that they are *collision-resistant* and that they exhibit the *avalanche effect*, i.e., a small change in the physical system causes a large change in the hash value.

In this dissertation, we use physical one-way hash functions to obtain *unique*, *tamper-resistant*, and *unforgeable* identifiers from 3D structures. Each of these three desired qualities has a corresponding mathematical/physical manifestation.

- *Unique*: The number of independent degrees of freedom in the output space should be large.
- *Tamper-resistant*: The output of the physical system must be very sensitive to changes in the state of the probe or the system itself.
- *Unforgeable*: It must be very difficult to clone the physical system in such a way that the cloned version produces identical responses to all probe states.

In the rest of this dissertation we will use OW(H)F to denote either algorithmic one-way (hash) functions. We will use POW(H)Fs to denote their physical counterparts.

1.3 Intellectual inspiration

Our work is philosophically inspired by the notion of Quantum Money, first proposed in 1983 by Wiesner [1] in a paper titled *Conjugate Coding*. In this paper, Wiesner presented two ideas. The first one was a verify-only memory, that, with high probability, cannot be read or copied by someone ignorant of its contents. The second idea was a scheme to multiplex two messages in such a way that, with high probability, either message could be recovered at the cost of irreversibly destroying the other. This work resulted in a follow-up paper [2] by Bennett, Brassard, Briedbart, and Wiesner entitled *Quantum Cryptography, or Unforgeable Subway Tokens*.

The basic idea was to use quantum mechanical systems, primarily polarized photons, to produce subway tokens whose validity could be checked by anyone but which no one could counterfeit. The scheme they proposed rested on the impossibility of simultaneously determining rectilinear and diagonal polarization of photons. We will take a closer look at the ideas behind quantum money in the next chapter.

1.4 Motivation

1.4.1 Authenticating bits with monetary value

This work was motivated by a simple practical question. Is there an inexpensive way to authenticate bits with monetary value?

The very properties that make bits extremely useful for *content representation* make them unsuitable for *value representation*. Consider the electronic cash scenario. Using a set of bits to represent a sum of money engenders several problems—e.g., multiple spending and counterfeiting—and the solutions to these problems lead to lack of privacy, lack of anonymity, and make payment traceability possible. These problems occur largely due to the fact that the monetary value is not tied to a physical representation. One way of doing this is to make use of the physical structure of the card in the transaction authentication process.

Consider another application where bits possess value. In the United States Postal Service’s Information Based Indicia Program (IBIP), postage stamps are allowed to be purchased and printed from a personal computer. The indicium (stamp) includes a two-dimensional (2D) barcode that is machine readable, along with human readable information. The indicium conveys mail processing and security related data. However, there is nothing that prevents a user from photocopying an indicium and reusing it several times. This problem may be mitigated by using an indicium derived not only from the user’s identifying data, but also from the *physical structure of the envelope* that carries the piece of mail.

The preceding example points at a broad class of emerging applications where there is an increasing need to be able to provide everyday objects with tamper-resistant and unforgeable serial numbers without significantly adding to their cost.

Smart cards, credit-card-sized devices with a single embedded chip, are currently being proffered for, among other applications, electronic wallets, authentication, and storing medical records. A single smart card transaction usually takes place within several data systems owned by different parties: the cardholder, the terminal (merchant or service provider); and the bank. In one class of attacks that can be perpetrated—by the cardholder against the terminal—counterfeit or modified cards running modified software can be used to subvert the security of the protocol between the card and the terminal. According to Bruce Schneier [72],

*“Good protocol design mitigates the risk of these kinds of attacks, which can be made more difficult by **hard-to-forge physical aspects of the card** (e.g., the hologram on the Visa and MasterCard cards), which can be checked by the terminal owner manually. Note that digital signatures on the software are not effective since a rogue card can always lie about its signature, and **there is no way for the terminal to peer inside the card.**”* [emphasis added]

The system proposed in this work relies literally on hard-to-forge physical aspects and gives the terminal the ability to peer inside the card.

1.4.2 Silicon-based authentication is not inexpensive enough

What does it cost to provide an uncopiable silicon-based serial number to an object? Despite the relentless onslaught of Moore’s Law, it costs on the order of a dollar for Silicon Serial Number DS2401 chip from Dallas

Semiconductor [83]. This chip provides a 48-bit unique identifier which can be read in approximately 3 milliseconds with a minimal electronic interface, typically a single pin of a microcontroller. Uncopiability is a little more expensive to come by. Minimal crypto-processors, which implement one-way hash functions, cost on the order of a few dollars. Clearly, there are many situations where adding a few dollars to the cost of an object is both economically and practically unacceptable. Our primary purpose for studying POWFs is to build systems which enable the identification and authentication of everyday objects in these situations.

1.4.3 Asymmetry between 2D and 3D microfabrication

Another interesting observation which motivates our work is the asymmetry between 2D and 3D microfabrication. The means to fabricate 2D structures have been steadily evolving over the past century culminating in a rather extreme example shown in figure 1.1.

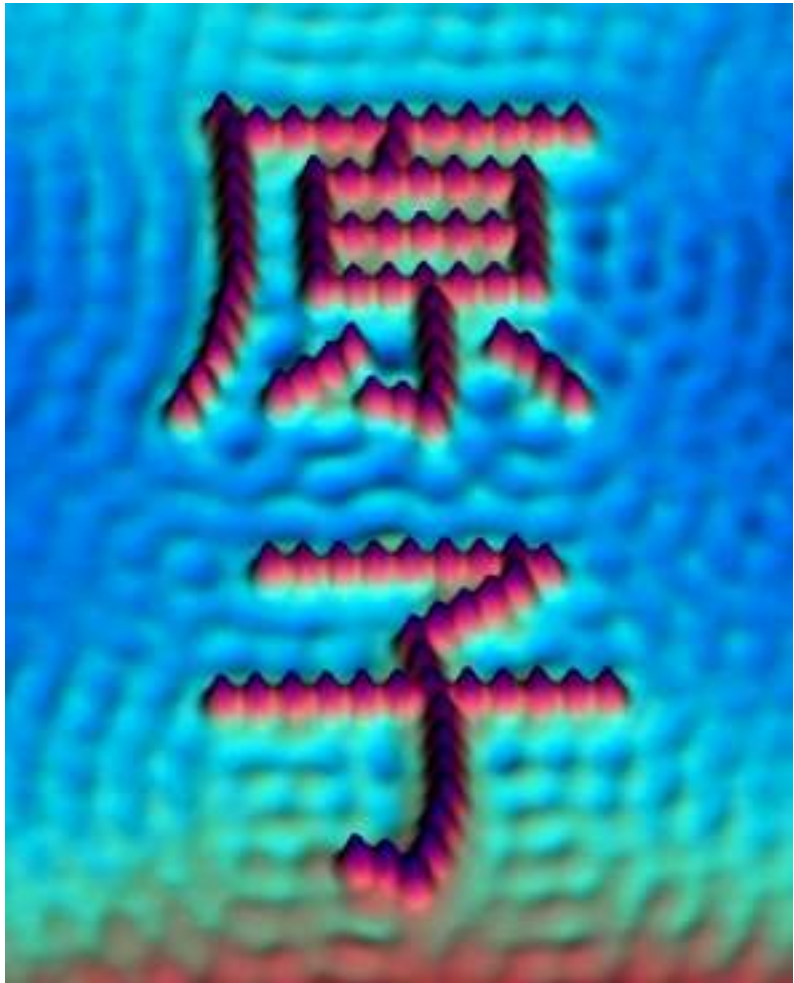


FIGURE 1.1 IRON ATOMS ON COPPER. THE KANJI TEXT READS *ATOM*. IMAGE BY DON EIGLER, IBM.

3D microfabrication, on the other hand, has been almost exclusively studied

in the context of Very Large Scale Integration (VLSI) and, more recently, Micro Electro-Mechanical Systems (MEMS). The standard fabrication method used to make essentially all microelectronic devices is photolithography where feature sizes are approaching 0.1 micron. However, this process is extremely expensive and current 90% yield main-line fabrication plants cost on the order of two billion dollars [81][82]. Further, standard top-down fabrication techniques are all geared toward producing *regular* structures at the submicron scale — producing arbitrarily random structures at these scales is still a very challenging problem.

1.4.4 Connection between physical systems and cryptography

During the course of this work, it became evident that there is a strong correspondence between physical one-way hash functions and algorithmic hash functions. As we look at physical one-way functions from a theoretical perspective, we are left wondering if there are any deeper connections to be found between (non-quantum) physical systems and cryptosystems. Specifically, since the same computational complexity theory is used to study both physical systems and cryptosystems, it would be very useful if the design and analysis of physical cryptosystems could be performed in the same (principled and rigorous) framework as algorithmic cryptosystems. In this spirit, we provide definitions of physical one-way functions that mirror the definitions for their algorithmic counterparts. Of course, we do not preclude the flow of concepts and ideas in the opposite direction: from physics to cryptography. This is certainly in keeping with Rolf Landauer's exhortation: *information is physical*. By means of this work, we offer an avenue by which the connections between physics, information theory, computational complexity, and cryptography can be further explored.

1.5 Research goals

In the service of coherence and cogency we defer discussion of our research goals to section 4.4.

1.6 Organization of the dissertation

This section provides a fairly detailed map of this dissertation.

Chapters 2 and 3 present some background and discuss related work. Specifically, we discuss algorithmic one-way functions and provide formal definitions for them. We use these definitions as templates in defining physical one-way functions in a later chapter. We then discuss the role of one-way (hash) functions in authentication and digital signatures and discuss common attacks on one-way hash functions. This is followed by a detailed exposition of Quantum Money where we look at how a quantum banknote is prepared and how a counterfeiter might approach the problem of cloning it. In the penultimate section of chapter 2, we provide a brief introduction to computational complexity theory and outline the various complexity classes. These classes make an appearance when we discuss the theory of physical one-way functions. Finally, we take a look at the various measures of physical complexity and focus on one such measure: the Kolmogorov complexity (also referred to as the algorithmic information content, algorithmic entropy, and algorithmic randomness) which we will use in a later chapter.

Chapter 3, *Related work*, is devoted to exploring the landscape of biometric authentication and looking at prior work in physical authentication. We spend a fair amount of time on optically variable devices where the physical authentication token modulate incident light depending on the angle of incidence. Familiar examples include the holograms on credit cards. The common feature of all these optically variable devices is that they are regular and are exactly the same on each object to be authenticated. We also look at authentication using random features in this chapter.

In chapter 4, *Concept, design choices, and problem formulation*, we present a first look at a conceptual physical authentication system. We look at the components of a physical authentication system and the data pipeline in such a system. This leads into prescribing the ideal requirements for each component of the system. Specifically, we lay out the prerequisites for the physical system, the probe, and the detector. Crucially, we also specify the desired characteristics of the *interaction* between the probe and the physical system. We then present design choices which satisfy the requirement for an exemplary embodiment of a physical authentication system. The last section of this chapter outlines the problems tackled in this dissertation.

Chapter 5, *Light transport through disordered media*, presents a detailed look at the physics of coherent multiple scattering. In the pre-quantum-mechanics era, the scattering of light by small particles occupied the minds of almost all the great masters of mathematical physics — Fresnel, Maxwell, Cauchy, Green, Poisson, Kirchoff, Stokes, and Lord Rayleigh. Of course, *coherent* multiple scattering in the visible region of the spectrum was not observed till the invention of the laser in 1962. Since then, however, there have been several important advances in the study of multiple scattering.

We present an overview of classical speckle theory, primarily formulated by Goodman. We then look at a very interesting memory effect, which has theoretical implications for the study of multiple scattering as well engineering implications for physical authentication systems. It is worth noting that many of the key ideas about multiple scattering in the post-laser age actually came from the study of disordered *electronic* structures. The similarities between diffusion in electronic and optical disordered media have been the focus of a lot of recent theoretical and experimental activity, with many pioneering ideas introduced by Rolf Landauer.

We then take a look at some very recent work involving coherent light transport through *nonlinear* disordered media. Although we do not explicitly use nonlinear disordered media in this dissertation, we make use of the theoretical development to strengthen the argument for physical one-way functions. We then look at *optical localization*, which is the cessation of light diffusion through the structure. It occurs when the mean free path between scatterers approaches the wavelength of light. Localization may be viewed as a phase transition in the medium. It is well known in complexity theory that dramatic changes in computational cost, analogous to physical phase transitions, occur at the boundary between under- and over-constrained

problems [73][74]. This connection between physical phase transitions and computational complexity suggests that physical one-way functions could become harder to invert as the regime of operation moves closer to localization. We conclude this chapter with a summary of key ideas.

Chapter 6, *Theory of physical one-way (hash) functions*, presents a more formal approach to defining POWFs where we broaden our vision and define physical one-way functions in terms of a general physical system. Our goal for this chapter is to define POWFs by augmenting definitions of OWFs with physical definitions. The advantage of looking at POWFs through this lens is that it enables (and indeed, motivates) structured and succinct mathematical descriptions. We can then look at POWFs as a physical layer encapsulating (in the ideal case, at least) an underlying OWF. We also partition POWFs into two classes, weak and strong, depending on the computational complexity of simulating the interaction between the probe and the physical system. In this chapter, we show that the combination of coherent multiple scattering and inhomogeneous 3D microstructures implements a collision-resistant physical one-way hash function.

The next three chapters are concerned with implementation of an exemplary physical authentication system.

In chapter 7, *System design and engineering*, we document the design and implementation of a prototype physical authentication system. We progress through the design of various system components described in chapter 4, taking care to document (briefly) some of the instantiations of each component that we experimented with along the way. Especially important to note is the fact that coherent multiple scattering was just one of a number of different probes we considered. The three others that we seriously scrutinized for our application were optical coherence tomography, confocal microscopy, and magnetic resonance imaging. The similarities between coherent multiple scattering and OWFs were too striking to ignore which is what led to our ultimate choice. We discuss the mechanical design of the token reader as well as the thresholding algorithm, the Gabor hash algorithm, in the remainder of this chapter. We conclude the chapter by listing a set of tradeoffs that different physical authentication systems must consider and potential improvements in future versions of the system.

We discuss *Experiments and results* in chapter 8. The first experiment is a proof-of-principle experiment. We are primarily interested in showing that a unique identifier can be obtained from an inhomogeneous 3D microstructure repeatably by probing it with a laser beam. The second experiment asks questions related to the statistics of the identifiers. Here we deal with a large number of speckle patterns and look at how distinguishable they are from one another. In the final experiment, we focus on determining the effect of small change in the microstructure on the identifier.

In chapter 9, *Protocols*, we consider how a physical authentication system might be used in practice. In existing cryptosystems, protocols are built by using cryptographic primitives such as OWFs. In this chapter, we devise two

protocols: a one-time pad protocol and a bit commitment protocol to demonstrate how POWFs might be used. We stress here that these protocols are very simple and are only intended to convey the flavor of how POWFs might be employed.

Chapter 10, *Scaling, attacks, and fabrication complexity*, addresses three separate issues of importance. We discuss scaling of physical authentication systems including scaling the number of tokens, and scaling the size of the physical system encapsulated in a token. We then discuss several attacks an adversary might use to compromise a physical authentication system. Finally, we address the question of how hard it is to clone a 3D microstructure. We briefly look at available methods of microfabrication and attempt to get a feel for the resources required to construct a physical structure of the kind we use in this dissertation. In the interest of strengthening our view that a POWF is an underlying OWF with a physical encapsulation, we propose an idealized physical system cloning machine which is simply a Universal Turing Machine augmented with a fabrication head. We then introduce the notion of *fabrication complexity* which is a simple way to calculate the total computational and physical resources required to clone an arbitrary physical system.

Finally, in *Contributions and future work*, we provide a summary original contributions of this dissertation, and consider how it might be extended.

2 Preliminaries

A project that seeks to extend the application domain of algorithmic cryptography and biometric authentication must necessarily draw on ideas from several areas of research. In this and a later chapter, we present a precis of the various concepts and results which are relevant to our work. First, we will present a detailed discussion of One-Way functions. The formalism developed by Goldreich (among others) [11] will be introduced in section 2.1 with a view to defining physical one-way functions in a similar way. In section 2.2, we will look at authentication, digital signatures, and the concept of challenge-response protocols.

In section 2.3, we will look at the notion of quantum money. Much of the discussion is drawn from Wiesner's seminal 1983 paper [1] (written originally in 1970 but not published for over a decade) which introduced the concept of quantum money, and led to Bennett and Brassard [3] to quantum cryptography.

Section 2.4 presents a catalog of computational complexity classes in preparation for a discussion one-way functions. The complexity of physical systems is the subject of section 2.5. Here we will look at different ways of defining complexity in physical systems. The field of physical complexity focuses on making connections between physical systems and computational complexity. The goal of researchers in the field of physical complexity is to prove statements like “*predicting lattice gases is P-complete*”. Such statements provide clear connections between physical phenomena and the computational complexity of simulating them. The reason for our interest in physical complexity will become evident when we define physical one-way functions in chapter 6.

2.1 One-way functions

Our work on physical one-way functions constructs an analogy with algorithmic one-way functions. In this section, we will take a close look at the formal definitions and properties of cryptographic one-way functions, with a view to using a similar approach in the physical case. A few acronyms that we will use repeatedly in the rest of this dissertation are given below:

- OWFs - algorithmic one-way functions
- OWHFs - algorithmic one-way hash functions
- POWFs - physical one-way functions
- POWHFs - physical one-way hash functions

2.1.1 The origin of OWFs and OWHFs

One-way functions are central to modern public-key cryptography [9]. The notion of a one-way function first made its appearance in a very practical context. Consider the “login” procedure in a multiuser computer system (e.g. a network of Unix workstations). When an account is set up, the user chooses a password which is entered into the system's password file. Upon each successive login, the user is asked for the password, which is compared to the stored password. The stored password must be kept secret, in order to prevent

impersonation of a user by an (perhaps malicious) adversary. The security of the user authentication system hinges on the security of the password file. Needham [10] realized that it would be possible to allow the system to judge the authenticity of a password without actually knowing it. His system worked as follows. When a user first enters a password PW , the computer system automatically calculates a function $f(PW)$ and stores this, not PW , in the password file. When a user offers a password X on a successive login, the computer compares $f(X)$ with $f(PW)$ and allows the login if they are equal. The crucial insight was that if f is a one-way function i.e., for any argument it is easy to compute but extremely hard to invert, then, even if f and $f(PW)$ were made public, it would be nearly impossible for a reasonable adversary to compute the password from $f(PW)$. Here, a reasonable adversary is one that does not have access to exponential computing resources.

2.1.2 Formal definitions

In this section, we present definitions for cryptographic one-way functions. Our presentation draws heavily from Goldreich [11] [12] and Bellare [13]. Our goal will be to formalize the ideas represented by the statement:

“A one-way function is a function which is easy to compute but hard to invert.”

Saying that a function f is easy to compute means that there exists a \mathbf{P} -time algorithm A which, given an input x , outputs $f(x)$. The notion of difficulty of inversion requires a more elaborate explanation. Saying that a function f is hard to invert means that every probabilistic \mathbf{P} -time algorithm A' trying on input y to find an inverse of y under f will succeed with only negligible probability. A probabilistic \mathbf{P} -time algorithm is one that is capable of making guesses. Negligible probability is a term that defines a robust notion of rareness. A rare event should occur rarely even if the experiment that generates the event is repeated a feasible number of times. Formally, a sequence $\{s_n\}$ is negligible in n if for every polynomial $p(\cdot)$ and all sufficiently large n , it holds that

$$s_n < \frac{1}{p(n)} \quad 2.1.1$$

Essentially, for some sufficiently large value of n , the members of the sequence $\{s_n\}$ are all smaller than $1/(p(n))$.

Finally, we recast the discussion above into a concise mathematical statement.

A function $f: \{0, 1\}^* \rightarrow \{0, 1\}^*$ is called *strongly one-way* if the following two conditions hold.

- *Easy to compute:* There exists a deterministic \mathbf{P} -time algorithm A such that on input x , A outputs $f(x)$ (that is, $A(x) = f(x)$)
- *Hard to invert:* For every probabilistic \mathbf{P} -time algorithm A' , every polynomial P , and all sufficiently large n

$$\Pr(A'(f(y)) \in f^{-1}(f(y))) < \frac{1}{p(n)} \quad 2.1.2$$

Therefore, the probability that algorithm A' will find an inverse of y under f is negligible. The essence of the second condition above is that the hardness to invert is specified as an upper bound on the probability of success of efficient inverting algorithms. In the definition above, two cases are important. If the size of the output, i.e., $f(x)$ is the same as that of the input x , then the function is called a *one-way permutation*. If the size of the output is always fixed, regardless of the size of the input, then the function f is called a *one-way hash function*.

Strong one-way functions above required that any efficient inverting algorithm has negligible success probability. Weak one-way functions require only that all efficient algorithms fail with some non-negligible probability. We will use these definitions as templates when we define physical one-way functions.

A *trapdoor one-way function* is a one-way function which can be inverted using a specific piece of information called the trapdoor. One-way functions are hard for everyone (legitimate users and adversaries) to invert, whereas trapdoor one-way functions can be efficiently inverted by legitimate users who possess the secret trapdoor.

Another beast in the cryptographer's zoo is the *one-way hash function*, also referred to as a Manipulation Detection Code (MDC). Formally, a hash function F is a transformation with the following properties:

- (1) *Variable input size*: F can be applied to an argument of any size.
- (2) *Fixed output size*: F produces a fixed-size output.
- (3) *Ease of computation*: $y = F(x)$ is easy to compute.
- (4) *Preimage resistance*: For any given y , the probability of finding x with $F(x) = y$ is negligible.
- (5) *2nd preimage resistance*: For any fixed x , the probability of finding $x' \neq x$ with $F(x') = F(x)$ is negligible.

Properties 3 and 4 are statements about the one-wayness of the transformation F . Property 5 has some subtlety to it. As stated, it means it is computationally infeasible to find another message which hashes to the same value. This is a statement about collision resistance, and the associated function is termed a weak one-way hash function. Property 5 may be strengthened by saying:

- (5') *Collision resistance*: It is computationally infeasible to find any two messages x_1 and x_2 , such that $F(x_1) = F(x_2)$.

Any transformation which satisfies the revised Property 5 is termed a strong one-way hash function. This distinction exists because the effort required to

invert is different in both cases, as will be demonstrated later.

(6) *High sensitivity*: A final interesting property of a one-way hash function is that if a single bit in the input is changed, approximately half the number of bits in the output are changed. This is sometimes referred to as the *avalanche effect*.

2.2 Authentication and digital signatures

2.2.1 Definitions

Authentication, as defined by Simmons [14], is the “*determination by the authorized receiver(s) or perhaps arbiter(s) that a particular message was most probably sent by the authorized transmitter under the existing authentication protocol and that it hasn’t subsequently been altered or substituted for.*” The authentication problem may be divided into the verification problem and the identification problem. Verification determines whether or not the message is altered or substituted, while identification determines whether the message originated at the transmitter. We point out that authentication *per se* has nothing to do with keeping the message secret. Secrecy and authentication are completely decoupled in the framework of modern cryptography. We note that authentication requires one-way functions, and secrecy requires trapdoor one-way functions in this framework.

Digital signatures are the electronic analog of written signatures. They are a pattern of bits that may be appended to the message or may be an integral part of it. In either case, the process of producing a digital signature is to input the message to an algorithm which produces the signed message. This is where one-way hash functions enter the picture. They form the heart of the algorithm that produces the digital signature. A key result from cryptographic literature is that one-way functions are necessary and sufficient for secure signatures [15].

2.2.2 How one-way hash functions are used in digital signatures

One-way hash functions play a critical role in information authentication and digital signature schemes. Two different protocols for message authentication are described here.

Verifying authenticity but not sender’s identity: Alice sends a message M to Bob, and both of them want to be certain that the message is intact. In order to achieve this, Alice computes a one-way hash function F with the message M as input to produce a hash value H , i.e.,

$$H = F(M) \tag{2.2.1}$$

She sends both M and H to Bob. Bob computes the same functions on the received message and obtains his own hash value H_B . If $H_B = H$, Bob can be sure that the message was not altered in transit. In this case, one-way hash functions are used to create “message digests” which can authenticate messages. There are several well-known hashing algorithms available for use in signature schemes. Among them are the Secure Hash Algorithm (SHA1) and the Message Digest 5 (MD5).

Verifying authenticity and sender's identity: In the previous case, an adversary could intercept both the message and the hash, alter the message and rehash it, and send it to Bob, who would then verify that it was authentic. The problem is that the message he verified as authentic never originated from Alice. In order to circumvent such incidents, one-way hash functions can be used in conjunction with symmetric or asymmetric cryptosystems to verify that the message is intact, and that it came from the person who claims to have sent it. In general, the elements of a scheme for unforgeable signatures requires that:

- each user has an efficient algorithm to produce his or her own signature.
- every user can efficiently verify that a certain string is the signature of another specific user
- nobody can efficiently produce signatures of other users to documents that they did not sign.

Here is a simple protocol that uses a symmetric cryptosystem. Alice and Bob both have access to the same secret key.

Alice:

- Create a hash of the message $H = F(M)$
- Create a digital signature by encrypting H with her secret key
- Append the digital signature to the message and send it to Bob

Bob:

- Create a hash H_B of the received message, i.e., $H_B = F(M)$
- Decrypt H_B with the secret key
- If H_B does not decrypt, then the message was not sent by Alice.
- If it decrypts, then compare the decrypted hash to the one created locally.
- If they are equal, then the message is unaltered. If not, then it was altered.

There are several protocols available for digital signature schemes, with and without encryption, and using either symmetric or asymmetric cryptosystems. A good review of digital signatures and protocols may be found in [17].

2.2.3 Attacks on one-way hash functions

Brute force attack: Assume a transformation F is a hash function with an n -bit output. Let x_1 be the first message which was hashed. We are looking for another message x_2 which produces the same hash value. Assuming the output of the hash function is random, any random message we choose has a 2^{-n} chance of hashing to the same value. If we try k random messages, then the probability of a match is $1 - (1 - 2^{-n})^k$. This is equal to $2^{-n}k \approx 10^{-n/3.3}k$. Therefore, in order to find a match with unity probability, an adversary would

have to evaluate the hash function approximately $10^{n/3.3}$ times. This is the brute-force approach to compromising one-way hash functions.

Birthday attack: The second attack, which is subtler, is so named because the problem of finding two random messages which hash to the same value is identical to finding two people in a group of people who share the same birthday with probability greater than a certain threshold. The analysis proceeds as follows.

Say we have one-way hash function which has m possible outputs. That is, if each output is l bits, then $m = 2^l$. We are interested in the probability of two random messages evaluating to the same value when we make k evaluations. The total number of ways in which k hash values can be distributed in m cells is m^k , i.e., each of the k hash values can be obtained in m ways. Now, for there to be no collision, the first of k hash values can take m values, the second one can take $m - 1$ values, and so on. Therefore the probability p of no collisions is

$$p = \frac{m(m-1)(m-2)(m-3)\dots\dots(m-k+1)}{m^k} \quad 2.2.2$$

Therefore, the probability of at least one collision is $P = 1 - p$ which is

$$P = 1 - \frac{m(m-1)(m-2)(m-3)\dots\dots(m-k+1)}{m^k} \quad 2.2.3$$

which is

$$P = 1 - \left(1 - \frac{1}{m}\right)\left(1 - \frac{2}{m}\right)\dots\dots\left(1 - \frac{(k-1)}{m}\right) \quad 2.2.4$$

It is possible to show that

$$\left(1 - \frac{1}{m}\right)\left(1 - \frac{2}{m}\right)\dots\dots\left(1 - \frac{(k-1)}{m}\right) < e^{\frac{-k(k-1)}{2m}} \quad 2.2.5$$

which leads to

$$P > 1 - e^{\frac{-k(k-1)}{2m}} \quad 2.2.6$$

Therefore, if we require that the probability of a collision be greater than 0.5, all we have to do is find that value of k in terms of m which makes it happen. This value of k is

$$k > (2m \ln 2)^{1/2} = 1.17m^{1/2} \quad 2.2.7$$

This is a surprising result. For an 80-bit hash function, this value of k is on the order of 10^{12} , which is the square root of the number of evaluations required for the previous case. When $m = 365$, $k > 22.35$, which means that in a room of more than 23 people, the probability of two people sharing the same birthday is greater than 0.5. Any hash function which makes the birthday attack computationally infeasible is called a strong one-way hash function. Essentially, the output of a strong one-way hash function has double the number of bits than the output of a weak one-way hash function.

2.3 Quantum money

In a classic paper entitled “Conjugate Coding”, Wiesner [1] introduced two important ideas. He showed how conjugate quantum variables could be used to produce banknotes that would be impossible to counterfeit and how to implement a “multiplexing channel”, wherein either but not both of two transmitted messages could be received. Our work is a direct intellectual descendant of the former idea, and quantum cryptography evolved from the later one

2.3.1 Photon polarization

Our discussion of quantum money requires the existence of a two-state quantum system. While several such systems are available, we use photon polarization in our example. The polarization states of a photon are represented as vectors in a two-dimensional Hilbert space H . H has several orthonormal bases. Three important ones are: (a) the quantum mechanical states of left- and right-circularly polarized photons (b) horizontally and vertically polarized photons and (c) linearly polarized photons at $\theta = \pi/4$ and $\theta = -\pi/4$ from the vertical. Any arbitrary polarization state may be represented as a linear combination of any of the above sets of states. These states are shown in figure 2.1.

Clearly, since each of the above sets of states comprise an orthonormal basis of H , all the sets may be represented in terms of each other, as shown below.

2.3.2 Quantum measurement of photons

Having seen how each set of photon polarization states may be represented in

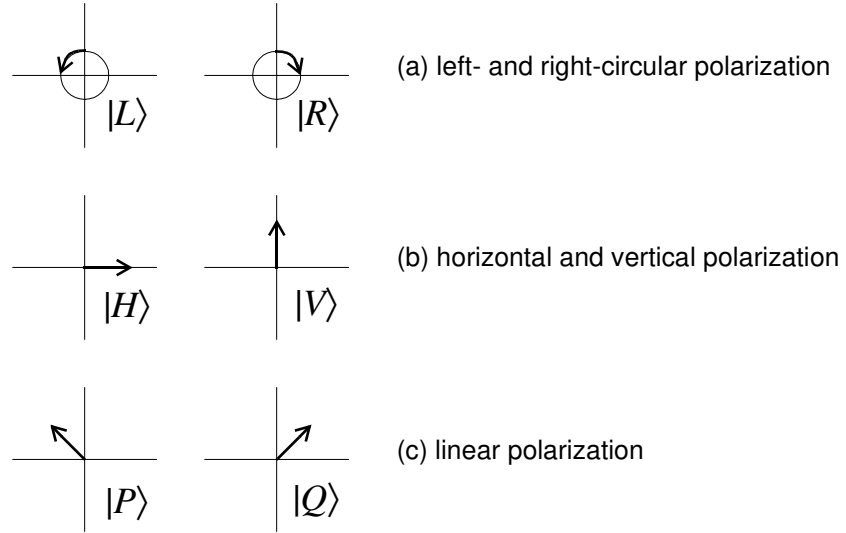


FIGURE 2.1 ORTHONORMAL BASES FOR PHOTON POLARIZATION. THE DIRAC NOTATION USED FOR EACH STATE IS ALSO INDICATED.

$ V\rangle = \frac{1}{\sqrt{2}}(Q\rangle + P\rangle)$ $ H\rangle = \frac{1}{\sqrt{2}}(Q\rangle - P\rangle)$	$ V\rangle = \frac{1}{\sqrt{2}}(R\rangle + L\rangle)$ $ H\rangle = \frac{i}{\sqrt{2}}(R\rangle - L\rangle)$
$ Q\rangle = \frac{1}{\sqrt{2}}(V\rangle + H\rangle)$ $ P\rangle = \frac{1}{\sqrt{2}}(V\rangle - H\rangle)$	$ Q\rangle = \frac{(1+i)}{2} R\rangle + \frac{(1-i)}{2} L\rangle$ $ P\rangle = \frac{(1-i)}{2} R\rangle + \frac{(1+i)}{2} L\rangle$
$ R\rangle = \frac{1}{\sqrt{2}}(V\rangle - i H\rangle)$ $ L\rangle = \frac{1}{\sqrt{2}}(V\rangle + i H\rangle)$	$ R\rangle = \frac{(1-i)}{2} Q\rangle + \frac{(1+i)}{2} P\rangle$ $ L\rangle = \frac{(1+i)}{2} Q\rangle + \frac{(1-i)}{2} P\rangle$

terms of another set, we turn our attention to measuring polarization states. Assume that we have a photon which has been prepared in state $|R\rangle$, i.e., right-circularly polarized. When this photon is passed through a vertical polarizer - one which lets $|V\rangle$ photons pass through with unity probability - the probability of seeing a photon is 0.5. This is an example of a quantum measurement and is shown in figure 2.2. This number is the squared magnitude of component of $|R\rangle$ in the direction of $|V\rangle$. The component is given by taking the inner-product

$$\|\langle V | \bullet | R \rangle\|^2 = \left\| \left(\frac{1}{\sqrt{2}}(\langle R | + \langle L |) \right) \bullet (|R\rangle) \right\|^2 = \left\| \frac{1}{\sqrt{2}} \right\|^2 = \frac{1}{2} \quad 2.3.1$$

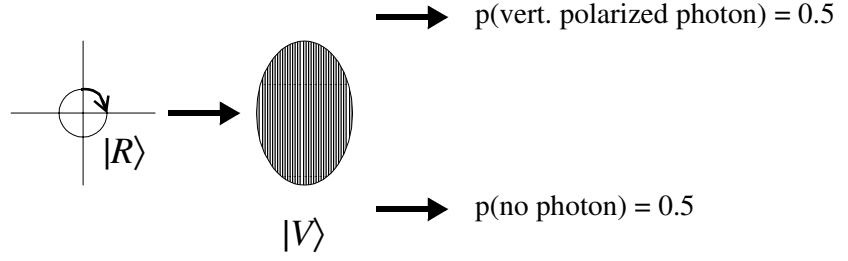


FIGURE 2.2 QUANTUM MEASUREMENT ON A PHOTON PREPARED IN A SPECIFIC STATE. MEASUREMENT PROBABILITIES ARE INDICATED.

which makes use of the fact that $|R\rangle$ and $|L\rangle$ are orthogonal.

The key point to note is that we could just have easily obtained the probability of 0.5 using a horizontal polarizer $|H\rangle$ in the above experiment. Given a 0.5 probability of seeing a photon at the output, there is no way of saying what the state of the input photon is. This observation plays a crucial role in the discussion of quantum money. We can make similar arguments for various other input photon states and measurement polarizers.

2.3.3 Preparing the quantum banknote

A quantum banknote contains a number, say n , of isolated two-state quantum systems such as spin 1/2 nuclei or photons with orthogonal polarizations. In the ensuing discussion, we use the photon polarization as our example. An important (but currently impractical) requirement is that the photons must be sufficiently isolated from the rest of the universe. Specifically, if a particular photon starts out in state $|H\rangle$ or $|Q\rangle$, then probability that a polarization measurement made on it during the lifetime of the quantum banknote will find it in a state $|V\rangle$ or $|P\rangle$ should be negligible. In other words, the photons should have a very long decoherence time.

In order to create a piece of quantum money, we need to encode the binary digits 0 and 1 using photon polarization states. This encoding is termed the quantum alphabet. Assume that the quantum banknote contains n photons. Generate two random binary sequences $M, N_i = \{0, 1\}$, ($i = 1, 2, \dots, n$). Each of the n photons is placed in one of the four states $|H\rangle, |V\rangle, |Q\rangle, |P\rangle$ depending on the concatenated sequence $M_i N_i$. Photon state preparation is depicted in the table below.

M_i	N_i	state
0	0	$ H\rangle$
0	1	$ V\rangle$
1	0	$ Q\rangle$

M_i	N_i	state
1	1	$ P\rangle$

The banknote is also given a serial number and the two sequences M_i, N_i are recorded along with the serial number. We now have a banknote with several isolated quantum systems whose state is determined by the two randomly generated binary sequences as shown in the table above. When the money returns to the bank, a measurement is made to see if the photons are still in their original state. Because the bank possesses the random sequences used to prepare the banknote, it also knows exactly how to carry out the measurement and obtain output photons with unity probability.

2.3.4 Forging a quantum banknote

We now consider how a potential forger would go about cloning the banknote. We assume that the quantum alphabets used in the encoding process are known.

All the forger has to do is prepare a counterfeit banknote in the same quantum state as the original. First, she has to make measurements on the photons of the original banknote and prepare photons in the same states and deposit them on the counterfeit. The latter process is assumed to be tractable, which leaves the issue of cloning the polarization states on the original banknote. We now show (non-rigorously) that this is impossible. The formal proof of impossibility [4] is provided by an amazingly simple quantum no-cloning theorem.

If the original note contains an $|R\rangle$ photon, then the probabilities of seeing an output photon with each of $|R\rangle, |L\rangle, |H\rangle, |V\rangle$ tuned polarizers are 1, 0, 0.5, 0.5 respectively. Therefore the forger has a (1/4) chance of measuring the original state. If the forger picks one of the resultant states and places it on the counterfeit banknote and repeats the experiment for each of the n states on the original, then the chance that the forgery will pass through undetected is $(1/4)^n \sim 10^{-(2n)/3.3}$. Therefore, the forger will have to prepare, on average, on the order of $10^{(2n)/3.3}$ counterfeits to produce one counterfeit which is indistinguishable from the original. For $n = 20$ this number of counterfeits is just a little over 10^{12} .

2.3.5 Discussion

The above gedanken experiment raises some interesting issues, which we point out here.

- By using random coin tosses to determine the original polarization state of the photons on the banknote, each banknote is associated with a unique signature that is dependent on a *physical structure*. When the number of photons is large (e.g., $n > 200$), the number of possible combinations of photon states is greater than the number of atoms in the universe, so we may assume that the probability that two randomly produced banknotes are identical is negligible.

- Unless one knows the original sequences (M_i, N_i) , the probability of cloning the physical structure is also very small, because cloning an unknown quantum state is impossible.
- In algorithmic cryptography, the attempts at forgery may be automated by using computer programs. A large number of combinations may be tried in a very short time. However, in systems based on physical structures, each attempt at forgery requires an experiment to be performed. Automating these experiments is much harder.

Quantum money hinges on the difficulty of reproducing a set of unknown quantum states. Practically speaking, however, quantum decoherence prevents any useful realization of the concept. In our work, we use physical structures to provide the signature, without the associated decoherence problem.

2.4 Algorithmic and computational complexity

We now turn our attention to the formal description of computational and algorithmic complexity. Modern cryptography is almost exclusively based on the gap between the difficulty of computation for legitimate users and adversaries. For example, in an encryption system, a legitimate user should easily be able to decipher the ciphertext by using some information known only to her. However, an adversary, who does not have access to the private deciphering key, should have a computationally infeasible task ahead of him. The formal study of algorithmic and computational complexity allow us to place notions of “easy”, “hard”, “infeasible” *et cetera* on a firm mathematical foundation. Excellent (and exhaustive) reviews of this material may be found in Papadimitriou [5], Greenlaw and Hoover [6], and Corman, Leiserson, and Rivest [7].

2.4.1 Problem size

An obvious question that arises is: how should we measure the size of a problem? Clearly, the time or resources used to perform a mathematical operation usually depends on the size of the inputs. How is the size of the input to be quantified? Because problem size depends heavily on the representation used in the problem, it is difficult, for example, to compare solutions to the same problem while using different representations (and hence, different size measures). Therefore, we seek a measure of problem size that is, for the most part, problem-independent.

One way to get around the problem-dependence is to assume a model of universal computation and declare the size problem to be the size of the input to this model. This is exactly what we choose to do. We use as our model of computation a Universal Turing Machine (UTM) and say that the size of our problem instance is the number of cells occupied by the input to the UTM [8]. So, for example, if we had a UTM that understood binary strings, and the input to the problem was the number 13, then the number of cells occupied by the input would be 4, which is the number of digits in the binary representation of 4 i.e., 1101.

In the ensuing discussion, we represent the size of the input by n .

2.4.2 Asymptotic notation

The lingua franca of algorithmic and computational complexity (hereafter ACC) is asymptotic notation. In general, asymptotics is a concept that has been developed to describe the growth rates of functions. Here, we are concerned with functions that measure the growth rates of computational resources such as time and memory. Asymptotic notation allows us to focus on the “big picture” of resource usage without getting bogged down by the messy details of a specific processor type or memory model.

The most prevalent asymptotic notation is the big-O notation. The basic idea of the big-O notation is to indicate that one function $f(n)$ is eventually bounded from above by another function $g(n)$. The formal definition of the big-O notation follows:

The function f is big-O of g , written $f(n) \in O(g(n))$, if and only if there exist constants $c > 0$ and $n_0 \in \mathbb{N}$ such that $f(n) \leq cg(n)$ for all natural numbers $n > n_0$. The set of all functions of growth rate order $g(n)$ is denoted by $O(g(n))$. In other words, $O(g(n))$ defines a family of functions.

As $g(n)$ takes on different characters, different names are given to $f(n)$. In the table below, we list the most common names given to $f(n)$.

complexity of $f(n)$	$g(n)$	O-notation
Constant	$g(n) = c$	$O(1)$
Linear	$g(n) = cn$	$O(n)$
Polynomial	$g(n) = cn^m$	$O(n^m)$
Exponential	$g(n) = m^{h(n)}$	$O(m^{h(n)})$

2.4.3 Complexity classes

The above notation is usually used to quantify the time or memory (space) resources required by algorithms. The same theoretical framework can be used to classify the hardness of problems, not just the algorithms used to solve them. The theory looks at the minimum time and space required to solve the hardest instance of a problem on a UTM. Problems that can be solved in polynomial time are called tractable. Problems which require greater than polynomial time are intractable. Figure 2.3 depicts the various problem complexity classes and the presumed relationships between these classes. Some of the relationships have not yet been strictly proved, but are widely believed to be true.

At the very bottom is the class **P** (hereafter, complexity classes will be denoted by a bold uppercase letter) of problems which are solvable in polynomial time. One level up from there are the class of **NP** problems, which

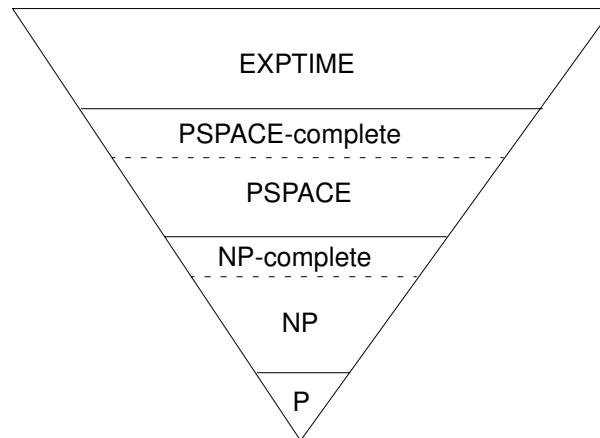


FIGURE 2.3 THE RELATIONSHIP BETWEEN COMPLEXITY CLASSES

are solvable in polynomial time on a nondeterministic Turing Machine. A nondeterministic Turing Machine is capable of guessing a solution and checking, in P -time, whether or not the solution is correct. The complexity class NP has special relevance to cryptography. Many cryptosystems can be cracked in NP -time by an adversary who guesses a solution and checks it in P -time. Clearly, NP includes P , but whether or not $P = NP$ is still an open question.

There is a subset of problems in every complexity class which are the hardest possible problems in that class. Formally, a problem X is said to be NP -complete if

- X is in NP , and
- every problem in NP can be transformed into an instance of X in P -time.

Notice that the second point embeds the notion of *polynomial reducibility*. Essentially, a problem's complexity class doesn't change if we are able to transform it into an instance of another problem in P -time. P -completeness is also amenable to polynomial reducibility.

Moving up the complexity food chain, we find the class of $PSPACE$ problems, which are problems solvable in polynomial space, but not polynomial time. $PSPACE$ -complete problems are problems with the property that if any one of them is in NP , then $PSPACE = NP$, and if any one of them is in P , then $PSPACE = P$. Finally, $EXPTIME$ is the class of problems solvable in exponential time.

Finally, we note that when we say feasible or tractable, we mean "solvable in P -time". Similarly, infeasible or intractable implies "solvable in greater than P -time".

2.5 Complexity in physical systems

2.5.1 Candidate metrics

Cryptography, as an intellectual endeavour, occurs at the confluence of several other disciplines: randomness, computation, information theory, communication theory, and computational complexity theory. Because cryptography draws from so many intellectual traditions, its study enables the discovery of interesting connections between them.

While computational complexity theory is concerned with the complexity of mathematical functions and algorithms, physical complexity is concerned with quantifying the complexity of physical systems. Although there is a large body of literature on the computational complexity of simulating physical systems, it is, in the words of Bennett [66], “*not immediately evident how a measure of the complexity of functions can be applied to states of physical models.*” We offer the idea that building physical cryptosystems will allow us to view the complexity of physical systems in a different light and, we hope, enable us to make better connections between physical complexity and computational complexity theory.

There are several approaches to quantifying the complexity of physical systems. In general, we seek a physical complexity metric that captures our intuitive beliefs of what is complicated while being rigorous enough to be mathematically formalized. This allows complexity-related questions to be posed well enough to be amenable to proof or refutation. Here we take a brief look at various candidate measures of physical complexity.

- *Thermodynamic potential* measures a physical system’s capacity for irreversible change but does not agree with our subjective notion of complexity. As an example, consider a supersaturated solution into which a seed crystal is introduced. The thermodynamic potential of the supersaturated solution is very high, but intuitively, its complexity is very low. On the other hand, the thermodynamic potential of the crystallized solution is low — there is no ability to change further irreversibly — but as viewed by an observer, its complexity is high.
- *Computational universality* is the ability of a physical system, programmed through its initial conditions, to simulate any digital computation. It is not entirely clear whether computational universality alone is a useful measure of physical complexity. One reason is that this definition does not distinguish between a system that is capable of universal computation and one in which computation has actually occurred.
- *Computational space/time complexity* is the asymptotic difficulty of simulating the physical system. However, defining physical complexity solely in terms of the complexity of simulating the underlying physical mechanism does not completely encapsulate the physical complexity of the system. To be more specific, it does not address the complexity of corporeally constructing the physical system.

- *Long-range order* is the existence of appreciable correlations between arbitrarily remote parts of the system. There are several examples of why this definition falls short of the mark. For example, consider a perfect crystal, which consists of a specific molecular unit repeated endlessly in three dimensions. To our intuition, this is a simple structure. However, the long-range order measure of complexity defined above regards a crystal as being extremely complicated, because the correlation between arbitrarily remote parts of the system is unity.
- *Thermodynamic depth* is the amount of entropy produced during the actual evolution of a physical system. It is easy to find physical systems that which arrive at very simple states through a large amount of dissipation and conversely, arrive at (subjectively) complicated states through very little dissipation. This definition of physical complexity is very system dependent.

Summarizing, each of the quantities described above captures one facet of a complicated physical system. However, it is easy to find physical systems which conform to the definitions above while violating our intuitive notions of complexity. In the next section we will look at a physical complexity metric which is avoiding this problem.

2.5.2 Kolmogorov complexity

Kolmogorov Complexity (also know variously as *algorithmic information content*, *algorithmic randomness*, and *algorithmic entropy*) is a definition of physical complexity which quantifies physical complexity in terms of the randomness in the physical system. It was introduced independently by Solomonoff [67][68], Kolmogorov [69], and Chaitin [70] in the early 1960s.

Kolmogorov Complexity (KC) is defined as *the size of the smallest computer program (in bits) required to generate the object in question to some degree of accuracy*.

To see what this means, let us consider this (often-used) example: we have two binary strings $x = 0101010101010101$ and $y = 10011010010110110010$. We are required to write a computer program which prints out each of the strings. The algorithm for the program which generates the first string might be simply "*Print 01 ten times.*" If the series were extended, by the same rule, the algorithm would have to be modified only slightly. It could, for example, now read "*Print 01 one million times.*" The program length has increased only very slightly in the second case, but the length of the output has increased considerably. In essence, the rate at which the program size increases is much smaller than the rate of increase of output size.

For the second sequence, it is not immediately evident what the algorithm should be. A potential algorithm might just be "*Print 10011010010110110010*". Notice here that the program has to essentially enumerate every bit in the string — there is no shortcut. Consequently, the size of the program is on the same order as that of the string. This example contains the definition of algorithmic randomness: *a sequence is random if the smallest algorithm*

capable of specifying it to a computer has approximately the same number of bits as the series itself. We now proceed to formalize this definition.

The Kolmogorov complexity $K_U(s)$ of a binary string s is defined as the length, in the number of digits, of the shortest program P that will produce output s and halt when used as input of a Universal Turing Machine U . Formally,

$$K_U(s) = |P| \quad 2.5.1$$

Clearly, the length of the program depends on the choice of symbol encoding, and the machine. However, by the definition of a universal computer, any program executable on computer U will also be executable, and yield the same output, on another U' , provided that it is preceded by a prefix program $T_{UU'}$ which allows the second computer to translate the first computer's program. Therefore, the algorithmic information content of a sequence may be considered, up to an additive constant, independent of the actual computer used, as long as it is a universal computer. Therefore,

$$K_{U'}(s) = T_{UU'} + K_U(s) \quad 2.5.2$$

The prefix program is, of course, independent of the string s . In the ensuing discussion we will assume that the computer is always a universal computer and omit the subscripts U and UU' .

It is instructive to reflect on what the definition of Kolmogorov complexity really means. One may think of the program as an *explanation* of the *observed data* which is the string. It is in this context, of treating programs as theories which explain strings, that Solomonoff discovered algorithmic complexity. The shortest program, he then declared, must represent the simplest explanation of the data — a statement very similar to Occam's Razor.

Another crucial point to note about KC is that, in contrast to the traditional Shannon entropy, it allows measurement of disorder without any need for probabilities. This is an important point and we will spend a little time discussing it here.

In general the entropy of a single state of a continuous system is not defined. Rather, one has to consider an ensemble of systems and define the entropy H as

$$H = \ln W \quad 2.5.3$$

where W is the number of possible macroscopically indistinguishable

microscopic configurations of the physical system. In quantum mechanics, the analog of entropy is given by von Neumann as

$$H = -Tr(\rho \ln \rho) \quad 2.5.4$$

where ρ is the density matrix of the system.

Shannon defined the entropy of a sequence of symbols $\{x_i\}$ each occurring with probabilities $\{p_i\}$, where $\sum p_i = 1$, as

$$H = -\sum_{-i} p_i \log_2(p_i) \quad 2.5.5$$

Each of these definitions relies on the probability density function of the ensemble in order to calculate the entropy. The entropy of any specific, completely-known physical state is always zero.

In stark contrast, KC does not require knowledge (or, indeed, the existence) of a probability density function for the ensemble of physical states. However, the two definitions are not all that dissimilar at least for thermodynamic ensembles. Bennett [71] has pointed out that, for a thermodynamic ensemble, the average Kolmogorov complexity is equal to the statistical ensemble entropy.

We now briefly outline some of the properties, without proof, of Kolmogorov complexity, focusing on those which we will find useful later in this dissertation.

- The Kolmogorov complexity of a typical string s is approximately equal to its length in bits, i.e.,

$$K(s) = |s| \quad 2.5.6$$

- If s is interpreted as a binary integer, then equation 2.5.6 implies that

$$K(s) = \log_2(s) \quad 2.5.7$$

- The *joint Kolmogorov complexity* of two strings s and t — the shortest program that generates each of the two strings in sequence — is given by

$$K(s, t) \leq K(s) + K(t) + O(1) \quad 2.5.8$$

where $O(1)$ is a constant. This may be also written as

$$K(s, t) = K(t) + K(s|t, K(t)) + O(1) \quad 2.5.9$$

- The mutual Kolmogorov complexity is given by

$$K(s;t) = K(s) + K(t) - K(s, t) \quad 2.5.10$$

This is a measure of the independence of the two strings. It measure how many more bits a program needs to calculate s and t separately rather than jointly. If, from equation 2.5.9, we determine that the joint Kolmogorov complexity is simply $K(s) + K(t)$, then the mutual complexity is 0. In such a case, we declare the two strings to be independent.

- Almost all strings of a specific length require programs of that length to generate them. In other words, most strings are algorithmically random and, therefore, equally likely. In such a case, we refer to the strings as *typical strings* or *typical sequences*.

2.5.3 Summary

In this dissertation, we are interested in physical complexity for two reasons. We are interested in determining the effort required to simulate the interaction of a physical probe with a physical system. This effort is measured by the familiar space/time computational complexity. We are also interested in quantifying the randomness present in any given instance of a physical system. We use Kolmogorov complexity for this purpose.

3 Related work

In this chapter, we will look at prior art and related work in the domain of optical document security. A comprehensive review of optical document security is too voluminous to include here, but we will review key techniques and patents. We also briefly look at recent work by Smith [24] who is using the texture of paper fibers to derive unique identity and at a patent issued to Amer, DiVincenzo, and Gershenfeld [25] that proposes tamper detection by bulk multiple scattering. These two pieces of work were starting points in our own investigation.

3.1 Prior art in physical authentication

We divide prior art into two categories. First, a representative selection of Optically Variable Devices (OVDs) is examined. OVDs may be thought of as physical structures which modulate incident light in a characteristic way that is dependent on the angle of incidence. The resulting light may be either human-readable or machine-readable. A search of the literature [18] has revealed that most OVDs are regular, 2D structures, with no variability from one device to another. An example is the hologram commonly found on all credit cards. The other category of prior art we will examine are systems which use random—as opposed to regular—2D physical features to authenticate objects. These random features may either be an intrinsic part of the object being authenticated or may be externally introduced. The clear distinction between these systems and our work is the use of three-dimensional, inhomogeneous microstructures and the use of coherent radiation to interrogate them.

3.1.1 Optically variable devices

In our visual world, the colors of objects are generally invariant to viewing position. Objects usually scatter light equally in all directions, a phenomenon called diffuse reflection. This homogeneous scatter of incident radiation is brought about by the highly irregular structure of matter on a microscopic scale. No wavelengths are preferred over others. In addition to this invariance with respect to angle of observation, the phenomenon of color constancy ensures that we perceive objects to be the same color almost independent of ambient light level. A piece of paper appears to be the same shade of white both in blazing sunlight and in a fluorescently illuminated laboratory. These two invariances collude to make objects appear invariable under normal illumination. The story changes dramatically when order is imposed on microscopic structures. The changing colors of an oil film on water (caused by interference) and the rainbow produced by a compact disc (caused by diffraction) are examples of microscopic order giving rise to optical variability.

The principal phenomena available for use in optical methods of authentication are: transmission, reflection, absorption, and scattering. These may be classified as shown in the figure 3.1. The two modes of operation of most optically variable devices (OVDs) are either reflection or transmission, as shown in the first and third quadrants of figure 3.1. A highly reflective structure is extremely conspicuous and lends itself easily to human

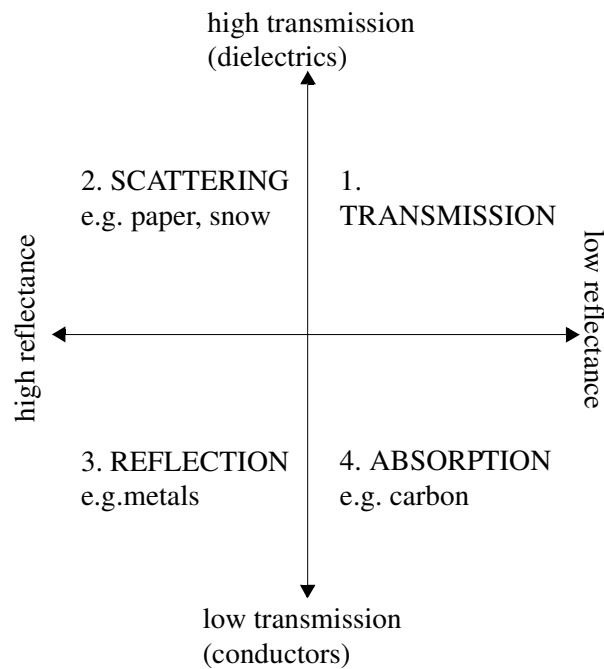


FIGURE 3.1 OPTICAL PHENOMENA AVAILABLE FOR USE IN OVDS

identification and verification. A transparent structure is useful for overlays and it is possible to construct them so that they exhibit iridescence. Absorption, in the fourth quadrant, is not a very useful optical phenomenon in the authentication business.

Another useful axis of classification of OVDs is shown in figure 3.2. Regular structures are viable for machine-readable as well as human-readable authentication systems, but usually they don't allow for unique identifiers. Random structures must necessarily form the basis of a machine-readable authentication system, and allow for the generation of unique identifiers. The gamut of iridescent OVDs is shown in figure 3.3 on the following page.

3.1.2 Authentication using random features

There are several patents in the literature that concern themselves with authentication and/or tamper resistance using *two-dimensional* random structures. We observe that in all cases, the authentication token is two-dimensional and no attempt is made to use cryptographic concepts in the description and analysis of the systems.

The first system [19] uses magnetic fibers randomly sprinkled and embedded in a thin substrate. To read the identity of the token, a magnetic read head is passed along the substrate and the return signal is logically combined, using the AND operator with a clock sequence. This produces a digital signal that is the identifier. The second patent [20] uses the variable translucency when a sheet of paper is illuminated with a light source. The data from the optical

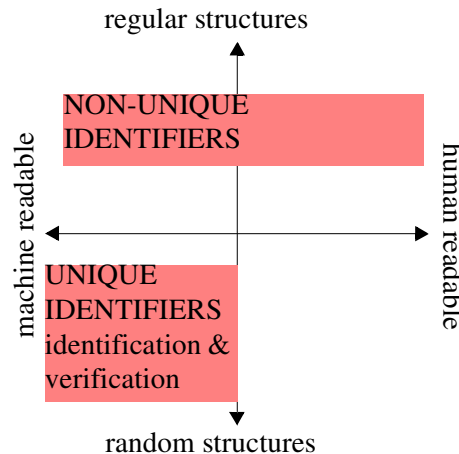


FIGURE 3.2 CLASSIFICATION OF REGULAR AND RANDOM STRUCTURES

reader is logically combined with a clock to produce the identifier. A third patent [21] uses small conducting particles embedded in an insulating substrate and uses microwaves to read the unique identifier. A fourth patent [22] uses a video microscope to view a small area of a painting at several magnifications and correlates these image with previously stored images.

The system that is most interesting to us is the 3D structure authentication system (3DAS) proposed by van Renesse [23] and currently being commercialized by Unicate [86]. In this system, a piece of cloth made from nonwoven 40 micron diameter polymer fibers is illuminated by two infrared LEDs in transmission mode, as shown in figure 3.4 below. In the identification case, only one of the LEDs is on, and the shadow of the fibers is projected onto the detector. The intersection of fibers, when projected onto the detector, produces convex polygonal shapes. The ten largest shapes are detected and their centers of gravity - twenty coordinates in all - are used as a 20 byte identifier. These identifiers are enrolled in a database and when a candidate token is presented, its identifier is computed and compared with all the members of the database.

In addition to identification, if verification is also desired, both LEDs are switched on in sequence, and one image is subtracted from another to produce an image which is sensitive to the parallax between both images. The security assumption here is that it is hard to spoof the parallax image, since it is hard to reproduce the fiber pattern. This may be viewed as a simple challenge-response protocol wherein the fiber structure is interrogated twice, and the interrogator knows what response to expect.

In the same vein, Smith [24] has used the texture of paper to derive identity information. Specifically, the problem addressed by this work is to prevent double spending that occurs when a postage stamp that is downloaded is photocopied and used as actual postage. The approach is to use a "texture hash string" derived from a specific location on *the envelope* on which the

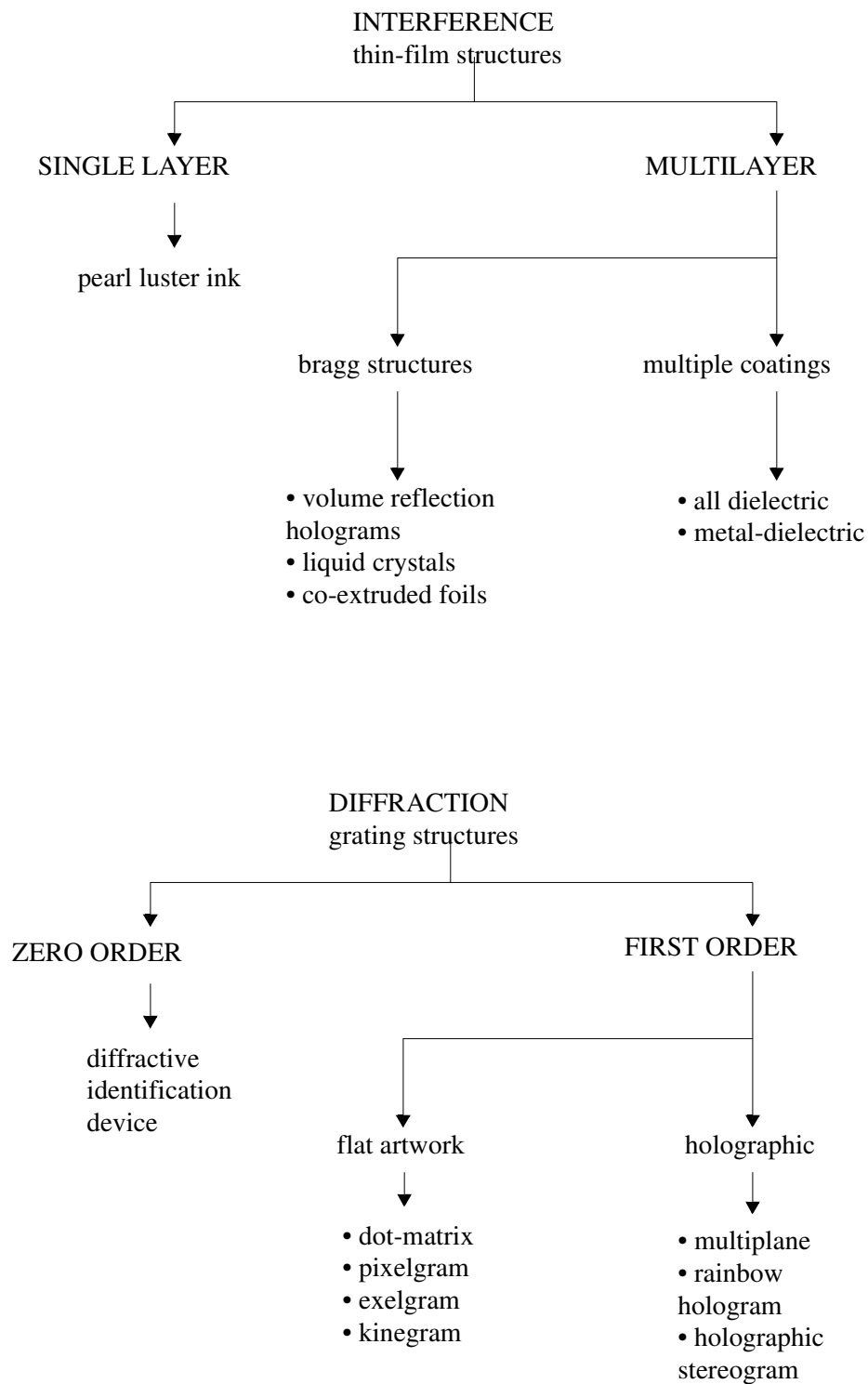


FIGURE 3.3 THE GAMUT OF IRIDESCENT OVDS

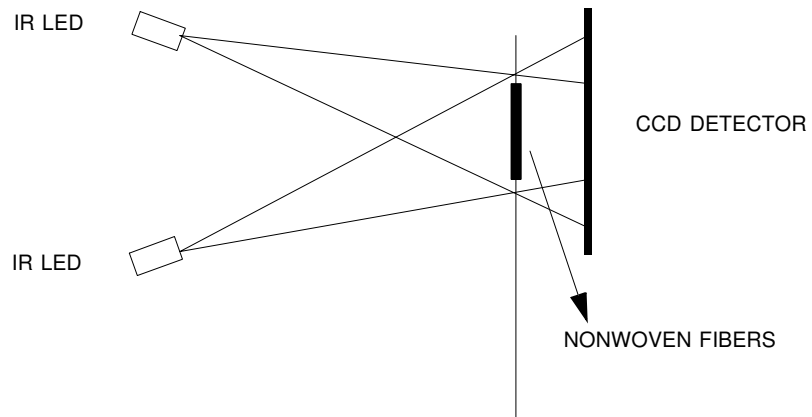


FIGURE 3.4 3DAS SETUP

postage is to be printed. This string is printed on the envelope in machine readable form and concatenated with a digital signature of the string. This combined string is referred to as an indicium. The digital signature is performed by an authorized agent of the postal department.

In order to validate a piece of postage, the two parts of the machine readable indicium are read. First, the signature is checked against the hash string. If there is a match, then a new texture hash string is obtained from the envelope and correlated against the texture hash string already printed on the envelope. The validating reader then makes an accept/reject decision based on whether or not the correlation score is below a threshold.

Finally, we briefly look at the patent which was the starting point for our own investigation. The patent, titled *Tamper detection using bulk multiple scattering* [25], disclosed a method of detecting intrusion into a protected area or package. The area to be protected is enclosed by an inhomogeneous medium. The extreme sensitivity of scattered light to changes in the structure was used as a sign that the package has been intruded into. The authors of the patent suggest that the response of the medium can also be used to provide a unique identity key. However, no exemplary embodiment was constructed [26], which we took on as our initial goal.

3.2 Summary

We conclude this chapter by presenting a few key points gleaned from our search of the literature in the various fields of inquiry discussed in the preceding sections and the previous chapter.

- Our work is inspired by the notion of Quantum Money - money which is tamper-evident and unforgeable.
- One-way functions are necessary and sufficient for secure signatures.

- The study of physically-based cryptosystems, in general, and physical authentication, in particular, are interesting ways to understand the complexity of physical systems.
- Prior art in physical authentication has focused almost exclusively on two-dimensional structures as the source of authentication information.
- In all but one of the cases, incoherent radiation was used to probe the physical structure.
- Previous work in physical authentication has not made an explicit connection with algorithmic cryptography.

4 Concept, design choices, and problem formulation

In the introductory chapter, we demonstrated the need for non-silicon-based inexpensive authentication systems and looked at the available methods of physical authentication and verification in chapter 3. We concluded the literature and prior-art search by observing that almost all previous methods of physical authentication relied on using *two-dimensional* structures which were probed by *incoherent* radiation. We also observed that previous work in physical authentication has not made an explicit connection with algorithmic cryptography.

In this chapter, we present the general concept of a physical authentication system whose security properties may be examined in a cryptographic light. In section 4.1 we present the system concept and lay out the data pipeline from the physical system to the unique identifier derived from it. We then state the ideal requirements for each component of the system. In section 4.3, we declare the choices we make in order to implement an exemplary embodiment of the system. Finally, in section 4.4 we present a set of questions that must be asked of any physical authentication system, and that we endeavour to answer in later chapters.

4.1 System concept and data pipeline:

A general physical authentication system consists of a physical system S encapsulated in a token T . Physical probe P and detector D together comprise the reader R . The probe P acts on the system S to produce an output O that is recorded by the detector D . Then an algorithm A acts on the received signal to produce the unique identifier U . This process is diagrammatically represented in figure 4.1.

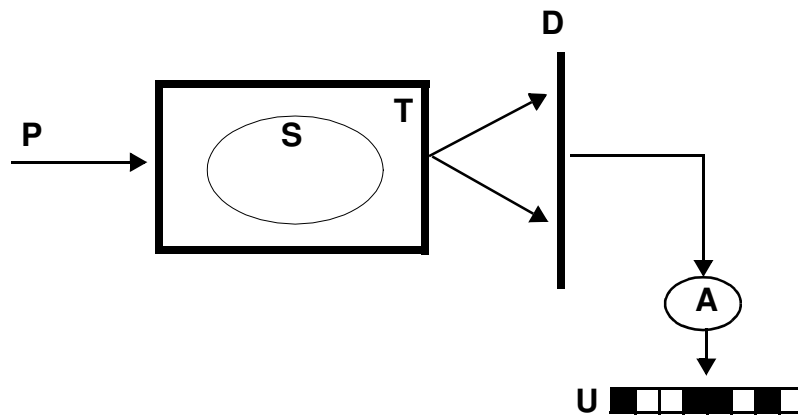


FIGURE 4.1 CONCEPTUAL AUTHENTICATION SYSTEM

Clearly, the choice of each component of the system determines the

configuration and the performance of the authentication system. The physical system and desired performance together determine the probe, and the relationship between P and S determines electro-mechanical design of the token reader. The characteristics of the detector play an important role in determining the quality and robustness of the unique identifier U . Given these interdependences, it is important to qualitatively prescribe the requirements of each component, which we do here.

4.2 Requirements of each system component

Our exposition of the characteristics of each component in the system is driven by the desirable properties of the resulting authentication system. We take up each requirement in turn.

4.2.1 Physical system requirements

(1) *Easy to fabricate*: Our prototypical physical system P must be easy and inexpensive to make. This is a requirement because we anticipate that a very large number of these systems will be deployed in the real world. For this reason, it must be possible to mass-produce the tokens used in the authentication system inexpensively.

(2) *Easy to probe*: The system must be easy to probe. There must be a simple way to set up the probe and obtain the response of the physical system. If this phase were complicated, it would limit the practical utility of physical authentication by increasing the cost and complexity of the reader as well as make the identifier less robust to small changes.

(3) *Hard to clone*: The physical system must be hard to refabricate. Another way of saying this would be: it is difficult to build a machine which, given one token, produces another token with exactly the same structural configuration. Note that this condition is independent of the interaction between the structure and the probe, it merely requires a certain amount of hardness in cloning the physical system.

(4) *Structurally stable*: Because we expect the token to have a long lifetime (on the scale of years) the physical system must remain dimensionally stable over time. We are interested in systems whose mechanical and electromagnetic properties remain stable over time.

4.2.2 Requirements for the probe

(5) *Easy to generate*: The physical system must be capable of being interrogated by a probe which must be easy and inexpensive to generate. This requirement originates from the fact that the probe must be replicated in every reader and we expect several readers to be deployed at any given time.

(6) *Easy to reproduce a specific state*: The probe must be capable of presenting the same query to the physical system regardless of the specific instantiation of the reader. The readers might be in spatially disparate locations but the probes must be capable of being instantiated in a specific state. Specifically, every characteristic of the probe must be reproducible to an accuracy that depends on the interaction between the probe and the physical

system.

4.2.3 Detector requirements

(7) *Identical response*: Since each reader contains a detector, we require that the detector has an identical response to identical input incident on it. In other words, detectors must be interchangeable without any performance penalty.

In addition to the requirements specified above, interaction between various components of the system place further demands that must be met in an engineering implementation of a physical authentication system. We discuss these here.

4.2.4 Interaction between physical system and probe

(8) *Impractical or infeasible to simulate*: The interaction between the probe and the system must be computationally impractical or infeasible to simulate. We require this in order to circumvent the possibility of an adversary simulating the response of the system to a specific probe configuration. Ideally, we require the simulation of the response to be in complexity class greater than \mathbf{P} .

(9) *Output very sensitive to changes in probe or system*: We want the output O to be extremely sensitive to changes in the system or the probe. This condition allows for tamper-resistance. Any changes in the system configuration are easily detectable. This is both a blessing and a curse because tamper-resistance is obtained without cost, but requires careful engineering of the token reader R and has a bearing on the design of the algorithm A .

(10) *Hard to invert*: Finally, we require that it must be hard to infer the exact configuration of the system given knowledge of the probe and access to the output of the detector. In (3) above, we were concerned with the difficulty of cloning the physical system *independent of the probe*, while here we have access to the probe, and are interested in the difficulty of inferring the configuration of the physical system.

Finally, we note that definition of “hard” in (3) and (10) is not (yet) a mathematically formal one. (3) is a statement about the intuitive notion of the difficulty of three-dimensional microfabrication of arbitrary structures and (10) is a statement about the computational difficulty of inverse problems and depends very intimately on the relationship between the probe and the system.

4.3 Design choices

Here we present our choices for a physical authentication system which fulfill, for the most part, the requirements outlined above.

- *Physical system*: The physical system we used was a three-dimensional, inhomogeneous microstructure implemented by curing micron-scale glass spheres in optical-grade epoxy. These tokens are easy to make, very hard to clone, and dimensionally stable over the lifetime of the token.
- *Probe*: A Helium-Neon laser beam at a wavelength of 632.8 nm. For our

purpose this beam may be treated as being at a single wavelength. Laser light is easy to generate, and the laser starts up in the same state each time.

- *Detector*: We use a garden-variety charge coupled device (CCD) detector which has 320x240 pixels.

The physical phenomenon underlying the interaction between the system and probe to produce the output is termed *coherent multiple scattering*. We will have a lot more to say about this phenomenon in a later chapter. Thus far, we have avoided any reference to the algorithm which takes raw detector output and generates a unique identifier from it. Discussion of the algorithm is also deferred to a later chapter.

4.4 Problem formulation

We are now in a position to formulate the problems that we tackle in this dissertation. We do this by asking a series of questions that lead us from the concept through the engineering and theory to future work.

4.4.1 System concept

- Given a physical system, probe, and detector, is it possible to design and implement a physical authentication system that allows the reliable and repeatable production of an identifier that uniquely distinguishes the physical system from other similarly produced systems?

The engineering part of this dissertation, presented in chapter 7, tackles the above question, and answers it in the affirmative.

4.4.2 System theory and performance

- Given that we can build the physical authentication system, what are the parameters and tradeoffs which govern its performance?

This question is related to:

- How robust is the identifier to changes in the token, probe, and environment?
- What are the probabilities of “false accept” and “false reject”?
- How does the system performance scale with the size of the physical system?
- How does it scale with the number of tokens in circulation?
- For a given token size, what is the maximum size of the identifier (in bits) possible?

4.4.3 Attacks and spoofing

Another class of questions deals with attempts to spoof the physical authentication system. These are:

- What are possible attacks on a physical authentication system?
- How can they be baffled?
- How hard is it to clone an inhomogeneous 3D microstructure without access to the exact probe used?
- How hard is it if the probe is available?

4.4.4 Cryptographic framework and future work

Finally:

- Is it possible to view physical authentication systems in the same framework as algorithmic authentication systems?
- Is it possible to build full-fledged cryptosystems by using concepts presented in this dissertation?
- If yes, what form would they take?
- If not, how do the above concepts need to be changed?

The preceding questions take the basic notion of a physical authentication system and turn it into a well-posed research plan. The rest of this dissertation is devoted to executing this research plan.

5 Light transport through disordered media

The physical mechanism used to implement physical one-way functions is the transport of coherent radiation through disordered media. Coherent waves propagating through a disordered medium will emerge from that medium with a phase that varies randomly along the wavefront. The characterization of the complicated emerging wavefront — dubbed the *speckle pattern* — in terms of the physical parameters of the random medium and the incoming radiation is fascinating, and has led to several surprising conclusions recently. It is the purpose of this chapter to present key results related to speckle patterns. Section 5.1 presents the assumptions and notation used through out this dissertation. Section 5.2 looks at the various length scales and scattering regimes which are relevant to our work. Sections 5.3, 5.4, and 5.5 present relevant theory and results related to coherent multiple scattering, light transport through nonlinear media, and optical localization respectively. Finally, we present a summary of key ideas and results in section 5.6.

5.1 Assumptions and notation

Before we proceed, however, we briefly look at the notation used here (and in the rest of this work) and define key terms. The diagram in figure 5.1 depicts the standard geometry used in the study of transmission speckle patterns. A

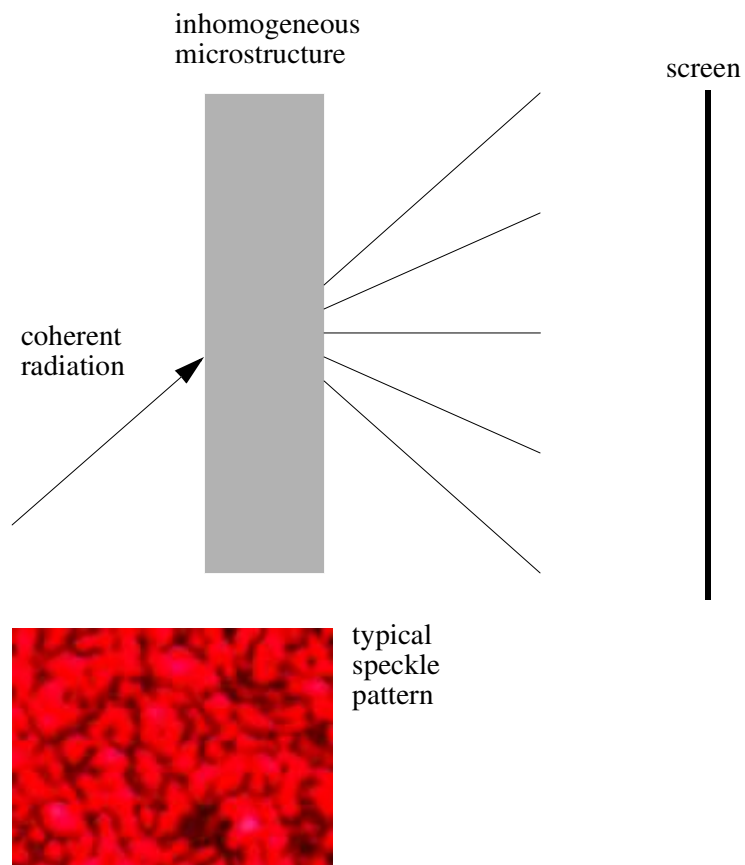


FIGURE 5.1 STANDARD GEOMETRY USED TO STUDY TRANSMISSION SPECKLE PATTERNS. A TYPICAL SPECKLE PATTERN IS ALSO SHOWN.

coherent wave of wavelength λ and wavenumber $k = (2\pi)/\lambda$ is incident on the left side of the slab. The incident beam is assumed to have a spot size $W \gg \lambda$. The thickness of the slab is assumed to be L . The slab is assumed to have random inhomogeneities so as to scatter the wave *elastically*, which means that the scattered wave has a definite phase relation with respect to the incident wave. Further, an elastic scattering event is entirely reversible. In contrast, an inelastic scattering event does not preserve phase the phase relation between the incident and scattered waves. Because the speckle pattern is an interference pattern, definite phase information improves the contrast of the pattern, and maximizes correlation effects (which are crucial to the systems discussed in this dissertation). We will restrict our analysis to elastic scattering and assume that absorption effects are negligible.

We will denote the *elastic mean free path* by l . The mean free path is the average distance between scattering events. As the density of scatterers increases, the mean free path decreases. The mean free path may also be defined as the distance in the slab at which the coherent incident beam intensity has fallen to $1/e$ of its value before entering the slab. We will also assume that the slab is dimensionally stable and the medium is static, i.e., the inhomogeneities which give rise to the speckle pattern do not fluctuate in time.

5.2 Length scales and scattering regimes

There are generally four length scales we need to consider when we look at coherent scattering. The first is, of course, the wavelength λ of the radiation being used to interrogate the structure. The second is the mean free path l . Third, we have the thickness of the disordered structure L , followed by the lesser of either the absorption length L_{abs} or the coherence length of the radiation L_{cl} . We assume that the coherence length is much longer than the absorption length in our work, and will neglect it from consideration. As we will see below, the relationship between each of these lengths determines the regime in which the scattering occurs, and guides experiments.

If the slab thickness L is smaller than the mean free path l , then, on average, the incident beam suffers only one or no scattering event before exiting the slab. This case is called the *Born regime*, since the well-known Born approximation [43] from quantum mechanics may be employed to study this regime. It is also generally referred to as the *single scattering* regime. If, however, $l \ll L \ll L_{abs}$, then the incident wave will suffer *multiple scattering* events before exiting the sample. Our work crucially depends on multiple scattering, and we will focus our attention on this regime throughout this dissertation. Our work will, therefore, be governed by the inequality $\lambda \ll l \ll L \ll L_{abs}$. The first inequality ensures that localization effects (see below) are small, the second ensures that multiple scattering occurs, and the last one ensures that not all radiation is absorbed.

The propagation of light through a disordered medium may also be described as a diffusion process characterized by a diffusion coefficient D . This leads to an Ohm's Law description: the conductance of light through the sample decreases linearly with increasing sample thickness. However, the diffusion

approach completely neglects any interference effects inherent in wave propagation. It accounts only for the average intensity transmitted through the sample. However, under certain conditions, it is possible to completely stop diffusion through the slab and effectively trap or *localize* light in the medium.

This phenomenon, aptly termed *optical localization*, was predicted by Anderson [44] while studying electron transport through disordered metals. Optical localization is governed by the Ioffe-Regel [45] criterion $\lambda/l \sim 1$, i.e., the mean free path is on the same order as the wavelength of the radiation. The precursor to complete localization is termed *optical weak localization*, and is governed by $\lambda/l \ll 1$. Optical weak localization is accompanied by enhanced backscattering of light [46] [47]. In the localized state, there is an exponential decrease in the conductance of light as the sample thickness increases [48]. Therefore, effectively, the medium undergoes a phase transition as it passes from the macroscopic diffusion regime to the localized regime. We summarize the governing equations and properties of the various regimes in the table below.

single scattering	$L \ll l$
multiple scattering	$L \gg l \gg \lambda$
optical weak localization	$l \gg \lambda$ + enhanced backscattering
optical localization	$l \sim \lambda$ + no transmission

We point out that in all the cases above, the medium is linear. In a later section we will consider light transport through a disordered nonlinear medium i.e., one that is governed by a nonlinear differential equation. Although our work will not include any nonlinear media, we will look at linear media as a limiting case of nonlinear media.

In our work, we are interested in the multiple scattering and optical weak localization regimes. Clearly, a localized system is of no use to us, since we rely on the scattered light to derive authentication information. Therefore, we will focus our attention on the characterization and properties of the former two regimes in the ensuing discussion.

5.3 Coherent multiple scattering

5.3.1 Classical speckle theory

The classical theory of speckle patterns resulting from coherent multiple scattering was developed in a series of papers by Joseph Goodman and is summarized in [49] and [50]. The approach followed by Goodman is physically plausible and intuitive and proceeds as follows. Given an incident beam of unit amplitude in the direction a , the complex scattered wave amplitude in the direction b is a coherent superposition of a great many Huygens' wavelets, each coming from their last scattering event in the sample. Since the sample is assumed to be thick enough to permit multiple scattering, it is reasonable to assume that the phases of the emerging wavelets vary greatly (compared to 2π) and randomly. This may be mathematically

treated as a random walk of wavelet amplitudes on the complex plane. Specifically, we can write the total scattered complex amplitude in direction b , after N scattering events, as

$$A_{ab} = \sum_{k=1}^N a_k e^{i\phi_k} \quad 5.3.1$$

Since several scattering events occur, it is reasonable to assume that a_k and ϕ_k are *uncorrelated* random numbers. From this description, the density function of the transmitted intensity $T_{ab} = |A_{ab}|^2$ may be obtained by using the random walk analysis in [49].

$$P(T_{ab}) = \frac{1}{\langle T_{ab} \rangle} e^{-T_{ab}/\langle T_{ab} \rangle} \quad 5.3.2$$

where $\langle T_{ab} \rangle$ refers to the ensemble averaged intensity in direction b . This is a simple negative exponential density function. The variance of this density, defined by

$$\delta T_{ab}^2 = \langle (T_{ab} - \langle T_{ab} \rangle)^2 \rangle \quad 5.3.3$$

can be shown to be

$$\delta T_{ab}^2 = (\langle T_{ab} \rangle)^2 \quad 5.3.4$$

This implies that the standard deviation is equal to the mean value, which means that the typical variations of the intensity about the mean are equal to the value of the mean, i.e., the speckle contrast is unity. This is the origin of the extremely grainy look of a speckle pattern.

This picture is perfectly valid and correctly predicts the first (and higher) order statistics of the speckle intensity. However, because of the assumption that all a_k and ϕ_k are uncorrelated, the analysis is unable to account for any correlations of the speckle intensity variation.

5.3.2 Born again: the memory effect

In the last sentence of the previous section, we claimed that the classical theory of speckle patterns is unable to account for any correlations of the speckle intensity pattern. This, of course, implies that there *are* correlations present in the speckle intensity pattern. In this section, we look at both a

gedanken experiment and experimental evidence to support this claim.

Gedanken experiment: Suppose the slab is very thin ($L \ll l$) so that we are in the Born regime. Light comes in the a direction and we are looking at the speckle pattern from the b direction. If the incident beam is rotated up by an infinitesimal angle $\delta\theta$, then it seems plausible that the speckle pattern should shift downward by the same angle. This is illustrated in figure 5.2. If the

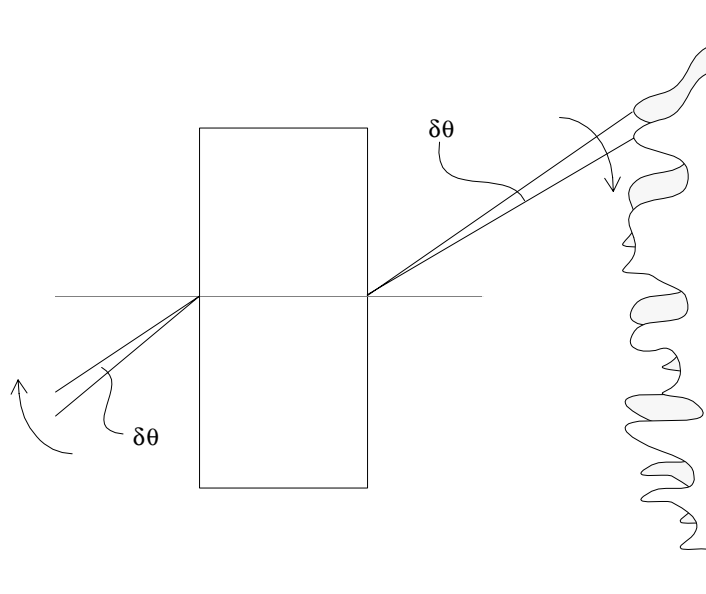


FIGURE 5.2 AS THE INCIDENT BEAM IS ROTATED BY A SMALL ANGLE, THE SPECKLE PATTERN ALSO SHIFTS BY THE SAME ANGLE.

phases ϕ_k are indeed uncorrelated, then we would expect a completely different speckle pattern to be formed when the incident beam is rotated by a small angle. Instead, the exiting scattered light “remembers” that the incident beam has been rotated. This *memory effect* [52] suggests that the phases are not quite uncorrelated as we assumed in section 5.3.1. Rather, they appear to be complicated functions of the incident angle.

One might argue, intuitively, that this effect might be observed in the Born regime, but would vanish in the multiple scattering regime. It might be supposed that as the thickness of the slab L becomes much greater than l , the typical path followed by a scattered wave is so complicated that the resulting speckle pattern retains no knowledge of the incident direction a . However, even in the multiple scattering case, the memory effect is still present, although it is weaker than in the Born regime.

5.3.3 Experimental observation of the memory effect

A simple experiment was carried out to verify the memory effect in the multiple scattering case. The memory effect is extremely useful to us as it allows us to relax the mechanical registration requirements on the authentication system. We will have a lot to say about this in a subsequent

chapter.

A disordered inhomogeneous structure was made by stirring 450 to 650 micron glass microspheres into optically clear epoxy which was allowed to set into a cavity of size 10mm x 10mm x 2.54mm. When the epoxy set, it had a milky appearance, indicating that incident light was being multiply scattered. This token was mounted on a rotation stage and a Helium-Neon laser ($\lambda = 632.8nm$) with a beam width of a few mm was set up so as to be normally incident on it. Speckle patterns were recorded at 0° incidence and at intervals of 1/100th of a degree.

Figure 5.3 shows the results of this experiment. The white line is a reference line and the white circles delineate a feature of the speckle pattern. Clearly, as the token is rotated, the speckle pattern does not instantly change into a completely different pattern. The structural changes appear after the token is rotated by more than three-hundredths of a degree. This experiment, performed in the multiple scattering domain, unambiguously verifies the memory effect and demonstrates the existence of correlations in the speckle intensity pattern. We now quantify these correlations.

5.3.4 The C_1 , C_2 , and C_3 correlations

This section draws heavily on the work done by Feng, Kane, Lee, and Stone [52]. We omit the complete (and complicated) mathematical treatment in favor of intuitive understanding and working formulae. The complete derivations are available in [52]. We will use the waveguide geometry depicted in figure 5.4 as a reference geometry. The waveguide geometry is chosen for simplicity. In an open geometry, the incident wave from a laser is not a plane wave, but a Gaussian wave packet with a width W . The cross-section area of the slab is $A = W^2$ in three dimensions. In a waveguide geometry, the incident and scattered beams are quantized (i.e., waveguide modes exist) and the correlations can be calculated precisely. The incoming beams are denoted by a and a' and the exiting beams are denoted by b and b' .

We begin by noting that there are three possible intensity transmission functions.

- The first, T_{ab} , is simply the angular transmission coefficient and measure the intensity exiting the medium in direction b as a result of incident light in direction a . Its correlation function is called the C_1 correlation and leads to the familiar high contrast speckle pattern.
- The second intensity transmission function is all the outgoing light as a result of incident light in direction a . This is given by

$$T_a = \sum_b T_{ab} \tag{5.3.5}$$

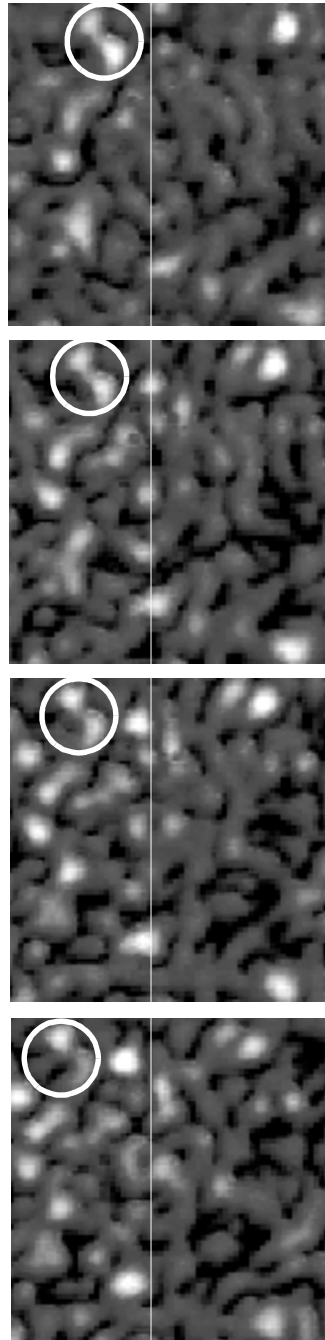


FIGURE 5.3 FOUR SPECKLE PATTERNS. THE TOKEN WAS ROTATED BY $1/100^{\text{TH}}$ DEGREE BETWEEN EACH RECORDING. THE WHITE CIRCLES DELINEATE A FEATURE THAT SHIFTS ACROSS THE RECORDING PLANE AS THE TOKEN IS ROTATED.

Its correlation function is denoted by C_2 .

- Finally, we have the sum of speckle intensity over all the input and output angles, given by

$$T = \sum_a T_a = \sum_a \sum_b T_{ab} \quad 5.3.6$$

Its correlation function is denoted by C_3

Thus, we are interested in three correlation functions. Specifically, we are interested in the functional form and orders of magnitude of the three functions.

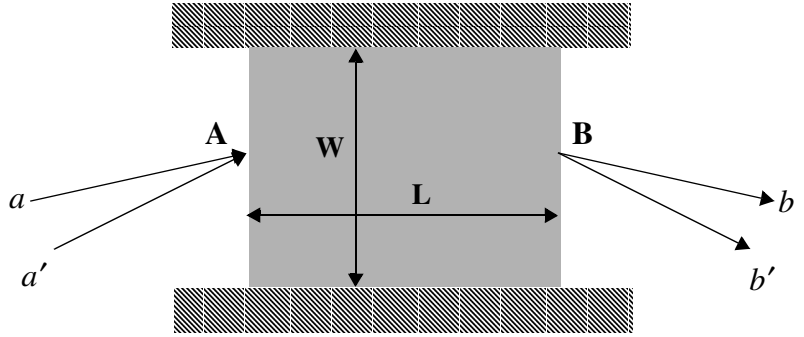


FIGURE 5.4 WAVEGUIDE GEOMETRY OF LIGHT PROPAGATION THROUGH DISORDERED MEDIA. LIGHT TRAVELS FROM A TO B THROUGH THE STRUCTURE.

The total number N of waveguide modes is related to the area and wavelength as follows.

$$N = k^2 A = \frac{4\pi^2}{\lambda^2} A \quad 5.3.7$$

We define the first correlation function C_1 as

$$C_1 = \langle \delta T_{ab} \delta T_{a'b'} \rangle \quad 5.3.8$$

where $\delta T_{ab} = T_{ab} - \langle T_{ab} \rangle$. Therefore we need to compute the ensemble average $\langle T_{ab} \rangle$ and use it to compute C_1 . The easiest way to compute the ensemble average is to use the standard diagram technique [53]. This yields

$$\langle T_{ab} \rangle = \frac{c_a c_b}{c^2} \left(\frac{l}{NL} \right) \quad 5.3.9$$

where c_a and c_b are the velocities in the a and b directions respectively. The prefactor $(c_a c_b)/c^2$ is of order unity, which leaves us with

$$\langle T_{ab} \rangle \sim l/(NL) \quad 5.3.10$$

Equation 5.3.14 makes intuitive sense and is a statement of the Ohm's law for the slab. The average intensity is inversely proportional to the slab thickness, and the number of "channels" supported in the waveguide.

The C_1 correlation may then be written as

$$C_1 = \langle T_{ab} \rangle \langle T_{a'b'} \rangle \delta_{\Delta q_a \Delta q_b} F_1(\Delta q_a L) \quad 5.3.11$$

where $q_a = a/(\pi W)$, $q_b = b/(\pi W)$, $\Delta q_a = (\Delta a)/(\pi W)$, and $\Delta q_b = (\Delta b)/(\pi W)$.

The shape of the correlation function is governed by the function

$$F_1(x) = \frac{x^2}{\sinh^2(x)} \quad 5.3.12$$

$\delta_{\Delta q_a \Delta q_b}$ is the Kronecker delta function and is equal to unity only when $\Delta q_a = \Delta q_b$.

A little reflection is sufficient to determine that our intuition about the memory effect developed in sections 5.3.2 and 5.3.3 is built into equation 5.3.15. The Kronecker δ ensures that C_1 is *non-zero only when* $\Delta a = \Delta b$, i.e., when the shift in incident and exit angles is identical. This is exactly what we observed in figure 5.3. Let us consider the case when the shift in angles is identical. We then have a correlation function governed by

$$C_1 = \langle T_{ab} \rangle \langle T_{a'b'} \rangle F_1(\Delta q_a L) \quad 5.3.13$$

A plot of $F_1(x)$ reveals that it falls off exponentially as the argument increases. Further, the peak becomes sharper as the argument increases. This is shown in figure 5.5. When $\Delta q_a = 0$, i.e., there's no change in the input angle,

$$C_1 = (\langle T_{ab} \rangle)^2 \quad 5.3.14$$

which is identical to equation 5.3.4. Therefore, C_1 reduces to the variance predicted by the classical theory of speckle when there is no change in the input and output angles, and has the memory effect built into it via the Kronecker δ function. We also note that C_1 decreases to zero much faster as the thickness L of the slab increases or as the incident angle changes.

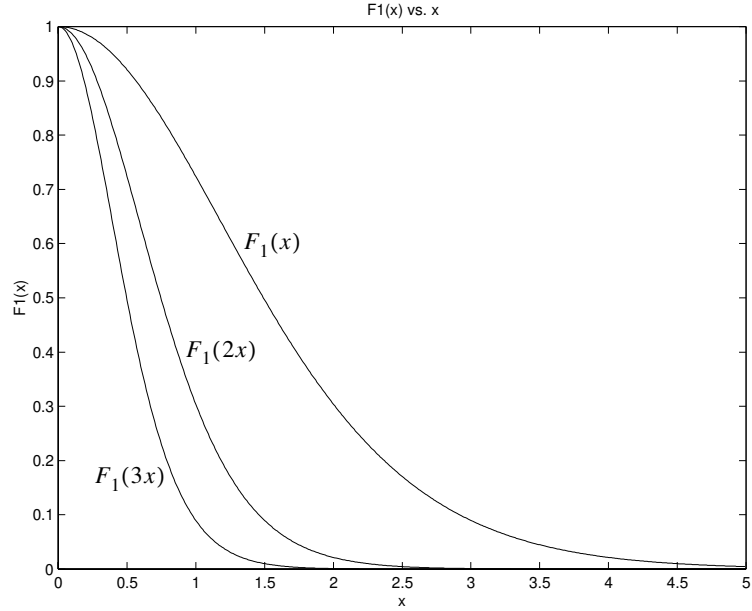


FIGURE 5.5 THE FUNCTIONS $F_1(x)$, $F_2(x)$, AND $F_3(x)$ RESPECTIVELY.

The effect of the C_1 correlation decays exponentially for $\Delta q_a L > 1$. The C_2 correlation may be written as

$$C_2 = \left(\frac{L}{Nl} \right) \langle T_{ab} \rangle \langle T_{a'b'} \rangle (F_1(\Delta q_a L) + F_2(\Delta q_b L)) \quad 5.3.15$$

where $F_2(x)$ is given by

$$F_2(x) = \frac{2}{x} \left[\coth(x) - \frac{x}{\sinh^2(x)} \right] \quad 5.3.16$$

While C_1 is non-zero only when $\Delta q_a = \Delta q_b$, C_2 is non-zero when *either* $a = a'$ *or* $b = b'$. If we assume that $a = a'$, we obtain

$$C_2 = \left(\frac{L}{Nl} \right) (\langle T_{ab} \rangle)^2 \sim \frac{l}{N^3 L} \quad 5.3.17$$

Thus, the intensities at different spots of a speckle pattern are uniformly and positively correlated by C_2 . Note that C_2 is $(Nl)/L$ times smaller than C_1 . Similarly, C_3 is given by

$$C_3 = \left(\frac{L}{Nl} \right)^2 (\langle T_{ab} \rangle \langle T_{a'b'} \rangle) \quad 5.3.18$$

This is smaller still, but it is a positive correlation. The intensity-intensity correlation of the speckle pattern may then be written as the sum of the three correlations as

$$C_{aba'b'} = C_1 + C_2 + C_3 \quad 5.3.19$$

5.3.5 C_1 , C_2 , and C_3 : an engineering view

From an engineering viewpoint, the existence of C_1 allows us to relax mechanical registration requirements on our token reader. To see why this is so, consider that we have an authentication token of thickness L and we have used a portion of its speckle pattern as its signature. When the token is presented to the reader again, we need to be able to match the new speckle pattern to the old one. If the C_1 correlation were simply a δ function, as predicted by the classical speckle theory, then the token would have to be presented to the interrogating light beam with a *variation of no more than the wavelength of light in spatial position*. Achieving this kind of repeatability in mechanical position is complicated and expensive, and translates directly into bulky and expensive readers.

We can relax the mechanical registration requirements and use the fact that C_1 exists to do a search of nearby positions *in software* to match the new speckle

pattern to the old one. This allows the readers to be less complicated and expensive while achieving the same authentication performance. We will have more to say about this in a later chapter. The C_1 correlation also dictates by how much the input angle must be rotated in order to obtain a new speckle pattern which is statistically uncorrelated with the old one.

However, the C_2 and the C_3 correlations, which are uniform and positive guarantee that any amount of rotation of the input angle will not produce a new speckle pattern which is completely uncorrelated with the old one. In particular, the C_3 is a long-range correlation which means that any two speckle patterns will be correlated by an amount equal to C_3 . The net effect of the latter two correlations is to reduce the total number of available identifiers.

5.3.6 Speckle sensitivity

In this section, we discuss the question: how sensitive is the speckle pattern to a small change in the slab? The answer is surprising: in 1D and 2D even the motion of a single scatterer in the structure can cause a strong change in the conductance of light through the slab. In 3D, only a small fraction of the scatterers has to be moved before the sample can be treated as a completely new sample.

The reasoning for these results is as follows. The conductance through the sample is proportional to the transmission probability through the sample, which can be understood in terms of the interference between light taking different paths through the structure. For a disordered structure, the paths are random walks of step-size l . In a random-walk, the total number of steps required to travel a distance D , given a step size l , is $N = (D/l)^2$. Therefore, the number of scatterers that each path encounters is $(L/l)^2$. Alternately, a finite fraction of all paths visit a specific scattering site. Assuming the total number of scatterers is on the order of $(L/l)^3$, the fraction of scatterers visited by a given path is proportional to l/L .

To probe a little deeper in to the sensitivity of the speckle pattern, assume there are N_p paths through the structure from A to B in figure 5.4. The total amplitude arriving at B in the output plane is given by summing the complex contributions arriving from all paths. This may be written as

$$A = \sum_{k=1}^{N_p} (|a|/\sqrt{N_p}) e^{i\phi_k} \quad 5.3.20$$

where it is assumed that each path contributes $(1/\sqrt{N_p})$ of the wave amplitude and ϕ_k is a uniform random variable in $[0, 2\pi]$. Clearly, this equation may be written as

$$A = \left(\frac{|a|}{\sqrt{N_p}} \right) \sum_{k=1}^{N_p} e^{i\phi_k} \quad 5.3.21$$

from which it is clear that the sum term may be viewed as a random walk of N_p steps in the complex plane. The sum is equal to $\sqrt{N_p}$ on average. This means that

$$\langle A \rangle = \left(\frac{\langle |a| \rangle}{\sqrt{N_p}} \right) \sqrt{N_p} = \langle |a| \rangle \quad 5.3.22$$

independent of the number of paths.

Now let us allow *one* scatterer in the volume to be moved. This means that $(N_p l / L)$ paths are affected. The change in the wave amplitude arriving at B is then given by the same equation as 5.3.22 except that the sum is now taken over just the affected paths. This may be written as

$$\Delta A = \left(\frac{|a|}{\sqrt{N_p}} \right) \sum_{k=1}^{(N_p l) / L} e^{i\phi_k} = \left(\frac{|a|}{\sqrt{N_p}} \right) \left(\sqrt{\frac{N_p l}{L}} \right) = |a| \sqrt{\frac{l}{L}} \quad 5.3.23$$

Therefore the fractional change in the arriving wave amplitude is

$$\frac{\Delta A}{A} = \sqrt{\frac{l}{L}} \quad 5.3.24$$

Therefore, since each site is visited by multiple paths, it is clear that the motion of a single scattering site is sufficient to affect a large change the speckle pattern. This is a key feature of coherent multiple scattering that is missing from other flavors of physical authentication: the extreme sensitivity of the speckle pattern to changes in the physical structure. In any system which does not rely on coherent interference between paths, the sensitivity is usually on the order of $(l/L)^3$, which is much smaller. A more formal approach to determining sensitivity of speckle patterns due to the motion of a single scatterer was developed by Berkovits [57] in which he arrived at the

same conclusions.

Of course, the high sensitivity of the speckle pattern is both a blessing and a curse. High sensitivity makes tamper resistance possible while it also amplifies the effects of wear-and-tear on the authentication token. This trade-off must be considered in the design of the features used to derive the unique identifier.

5.3.7 The random matrix formalism

There are several different approaches to determining the statistical properties of coherent radiation propagating through disordered media. The results in section 5.3.4 were obtained through the use of the Feynman diagram technique, which takes into account the sum of all possible paths light could take in getting from the input plane to the output plane of the disordered medium. An equivalent technique is the Green's function formalism, which allows one to calculate the amplitude of the electric field at any point in space given the electric field at any other point. Of course, the form of the Green's function is motivated by the physics of the problem at hand. Each of these techniques has a regime of applicability depending on the complexity of the problem or the presence of sources in the disordered medium. Suffice it to say that the Green's function approach is the most general, and can be applied in any situation. Other equivalent formalisms, which we will not address here, are the Kubo formalism, based on fluctuation-dissipation concepts, and the more familiar Hamiltonian formalism.

If, however, we are interested *only* in the relationship between incoming and outgoing radiation, the intervening disordered medium is linear and does not contain any sources of radiation, we can make use of a matrix formalism to describe radiation transport between input and output planes. The matrix formalism has several advantages. First, it is more intuitive than other formalisms. It is easier to think about light transport through the medium as a matrix multiplication. Second, the matrix approach allows us to easily calculate the computational complexity of simulating the passage of light through the structure. A complicated physical problem is now represented as a matrix multiplication.

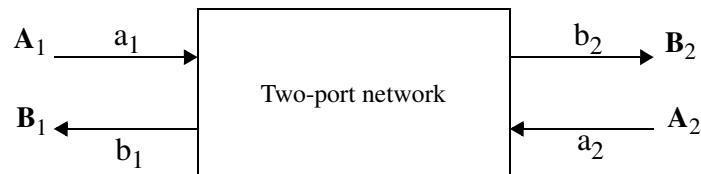


FIGURE 5.6 A STANDARD TWO-PORT NETWORK

Consider the familiar two-port network shown in figure 5.6. The ports are designated by uppercase letters. Incoming wave amplitudes are represented by two vectors a_1 and a_2 , and outgoing waves are represented by b_1 and b_2 .

In the general case, we may represent the relationship between the incoming and outgoing waves by a *scattering matrix* S as follows.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = S \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad 5.3.25$$

Each s_{ij} is a (usually complex) scattering parameter and equation 5.3.25 is a set of linear equations relating the inputs and the outputs via the scattering parameters. If, instead of a single wave, n waves are incident on a particular port and n waves exit at a different port, then the a s and b s in equation 5.3.25 are replaced by n -vectors. We write this as follows.

$$\mathbf{b} = S\mathbf{a} \quad 5.3.26$$

where S is, in general, a $2n \times 2n$ matrix, representing n waves for each of the two input ports, and similarly for the two output ports. We are only concerned with the incoming and outgoing waves on ports A_1 and B_2 , so our S will be an $n \times n$ matrix. After the diagram in figure 5.6 has been pruned to reflect this situation, we note that it bears a remarkable similarity to the diagram in figure 5.4.

In our case, the size of the matrix is determined by the number of modes N defined in equation 5.3.11. Because energy must be conserved, we find that the $N \times N$ element matrix S has to be unitary.

$$\sum_{m=1}^N |a_m|^2 = \sum_{m=1}^N |b_m|^2 \quad 5.3.27$$

where $\{a_m\}$ and $\{b_m\}$ are the input and output mode amplitudes respectively. This is the same as

$$\mathbf{a}^\dagger \mathbf{a} = \mathbf{b}^\dagger \mathbf{b} \quad 5.3.28$$

where the † represents a conjugate-transpose operation. This leads to

$$\mathbf{a}^\dagger \mathbf{a} = \mathbf{a}^\dagger \mathbf{S}^\dagger \mathbf{S} \mathbf{a} \quad 5.3.29$$

Therefore,

$$\mathbf{S}^\dagger \mathbf{S} = \mathbf{1} \quad 5.3.30$$

In terms of the elements of \mathbf{S} , we have

$$\sum_{m=1}^N |s_{mn}|^2 = 1 = \sum_{n=1}^N |s_{mn}|^2 \quad 5.3.31$$

That is, the sum of all elements in a single column of \mathbf{S} must equal to one - because all the energy in an input mode must be distributed amongst all the other modes. The photon has to go *somewhere*. The second part of equation 5.3.31 is less intuitive. It states that the sum of all energies entering a specific *output* mode must equal one. There is no simple reason why this should be true, other than the fact that this follows from the unitarity of the \mathbf{S} matrix.

Finally, our \mathbf{S} matrix must contain random elements because we know that the disordered medium mixes modes with no preference for any specific mode. This leads us to conclude that we are dealing with a special class of matrices usually referred to as *random unitary matrices*.

The scattering matrix formalism can be used in the mathematical analysis of the propagation of light through a disordered medium. Kogan and Kaveh [58] describe how a random-matrix formalism may be applied to derive all the relevant quantities, such as distribution functions for the total transmission coefficient and the angular transmission coefficient, may be derived in a random-matrix framework. For our purpose, it suffices to recognize that the scattering matrix is of size $N \times N$, where N is as defined in equation 5.3.7. For example, for a 10×10 mm slab and incident radiation of wavelength 0.5 microns $N \sim 10^{10}$, which means the scattering matrix has 10^{20} elements.

The transmission probability T_{mn} , indicating the probability that a wave in input mode m ends up in output mode n is given by

$$T_{mn} = |S_{mn}|^2 \quad 5.3.32$$

We may also interpret this equation as saying the probability that light incident on the slab at angle a exits the slab at angle b is given by T_{ab} . Each input and output mode essentially corresponds to an input and output angle.

5.4 Light transport through nonlinear media

In section 5.3, we looked at the transport of light through linear disordered media, and established that the sensitivity of the speckle pattern to small changes in the medium was very high, but finite. Here we will look at some very recent work by Spivak and Zyuzin [54] and Skipetrov and Maynard [55] that examines the sensitivity of speckle patterns to small structural changes in a *nonlinear* medium.

5.4.1 The approach

The approach to determining the sensitivity of the speckle pattern to structural changes is as follows: given an incident plane wave, and a disordered structure, [54] demonstrates that the number of possible speckle patterns increases exponentially with sample size. This implies that, given a speckle pattern, it is exponentially harder to determine the configuration of the structure which caused it as the size of the structure increases. Specifically, if we have a plane wave $\phi_0(\mathbf{r})$ incident on a disordered nonlinear medium with a scattering potential $u(\mathbf{r})$, then the propagation of light through this medium is governed by the nonlinear Schrodinger equation given by

$$\left(-\frac{\partial^2}{\partial r^2} - E + u(\mathbf{r}) + \beta n(\mathbf{r})\right)\phi(\mathbf{r}) = 0 \quad 5.4.1$$

where E is the incident wave energy, β is a constant, $\phi(\mathbf{r})$ is the electric field amplitude of the speckle pattern and $n(\mathbf{r}) = |\phi(\mathbf{r})|^2$ is its intensity.

[54] shows that the number of solutions of equation 5.4.1 for a given scattering potential, wave energy, and β is given by

$$N \sim e^{a\gamma^{3/4}} \quad 5.4.2$$

where $a \sim 1$ and

$$y = \left(\frac{3n_0\beta}{2E}\right)^2 \left(\frac{L}{l}\right)^3 \quad 5.4.3$$

where $\sqrt{n_0}$ is the electric field amplitude of the incident plane wave.

Equations 5.4.2 and 5.4.3 may be interpreted as saying that given microstructure pattern and an incident plane wave, the sensitivity of the speckle pattern to infinitesimal changes in the structure or in the input wave increases exponentially with sample size. The exponent is proportional to the cube of the ratio of the sample size to the mean free path.

5.4.2 Engineering issues

The engineering ramifications a nonlinear medium are interesting. First, the sensitivity of the speckle pattern to incident angle increases exponentially. Thus, if the incident angle θ changes by $e^{-\alpha r^{3/4}}$, the speckle pattern changes completely. This places extreme mechanical constraints on the token reader. Also, it is difficult to practically use a nonlinear medium in an authentication system of the kind we want to implement because the same structure produces (possibly) a different speckle pattern each time it is interrogated. This poses a problem in an authentication situation as we are interested in using the speckle pattern to derive a unique signature for the structure. The solution, then, is to use a very weak nonlinearity in the disordered structure to obtain the benefits, while keeping the complexity of the system unchanged. This weak nonlinearity may be achieved by mixing in a small amount of a nonlinear optical material into the structure. Such materials might be doped glasses, semiconductors, or organic materials. Recent work has demonstrated that glass doped with Cadmium Selenide nanoparticles has a large optical nonlinearity. It is conceivable that glass microspheres doped with CdSe might be used as scatterers in a physical authentication system.

The very strong complexity-theoretic analogy — the number of possible speckle patterns increases exponentially with increasing structure size — helps strengthen the case for physical one way functions.

5.5 Optical localization

We now turn our attention to the phenomena of optical localization (OL) and its precursor optical weak localization (OWL). Optical localization is the condition under which the diffusion of the wave through the inhomogeneous structure would grind to a halt — the diffusion constant vanishes. This condition was predicted for electrons being scattered through metals with inhomogeneous structures by Anderson [48] and later extended to the optical regime by Akkermans *et al.* [56]. When OL occurs, light is frozen in the structure, it does not escape.

The physical picture behind localization is quite simple. Suppose a wave propagates in a random medium from A to B, as shown in figure 5.7. The total probability for arriving at B is given by summing all possible paths and squaring the result. The final result would contain the sum of the squares of each individual path — the incoherent contribution — and the interference terms. Given the random phases of the interference terms, it is reasonable to assume that, on average, the interference terms vanish leaving only the sum of the squares of each individual path. However, this analysis does not take into account the case when A and B are one and the same. In this case, regardless of how long a particular path is, it always has a mate which has travelled exactly the same distance in the opposite direction. The probability of a wave arriving at A is not simply the sum of the squares of the individual paths, but four times the probability associated with a path. This interference is always constructive, and must not be neglected.

In the optical weak localization regime, there is a simply measurable effect that demonstrates this increased probability of return. The effect is called the

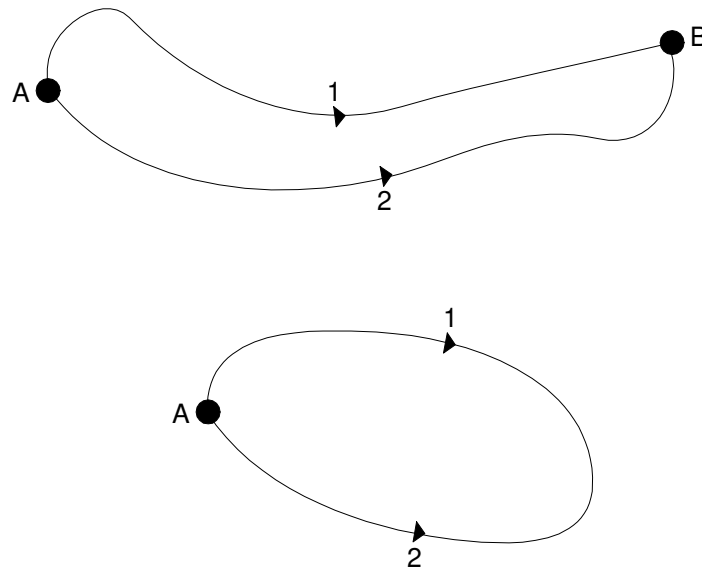


FIGURE 5.7 TWO RANDOMLY CHOSEN PATHS FROM A TO B IN A DISORDERED MEDIUM. A LOOP FROM A TO A IS SHOWN IN THE BOTTOM HALF OF THE FIGURE.

enhanced backscatter cone. If light is incident on a sample whose disorder is characterized by $l/\lambda \gg 1$, the intensity of the backscattered light (measured as a function of angle) will show a peak in the direction of the incident light. The width of the enhanced backscatter cone is approximately λ/l . This enhanced backscatter is a precursor to strong localization or, simply, localization.

If the scattering is very strong, i.e. $l \sim \lambda$, the diffusion constant tends to zero, and the optical localization regime is reached. The same scattering process that causes the enhanced backscatter also contributes to a drastic reduction in the diffusion constant. As the scattering becomes stronger, the contribution of the loops becomes stronger, the return probability of intensity increases, the diffusion constant reduces, and eventually all light is trapped in the structure. The transition from a non-zero diffusion constant to a zero-diffusion constant is called Anderson Localization.

In the diffusion regime, extended states are responsible for diffusion because they have infinite extent. In the Anderson localization regime, there exist only localized states: no extended states exist. the localized states decay exponentially to zero over one localization length. Thus, the diffusion regime and the localized regime are two distinct phases in which light behaves very differently. Interestingly, phase-transitions of this kind are associated with dramatic increases in the physical complexity of the system [73][74].

5.6 A summary of key ideas

Finally, we summarize some key ideas presented in this chapter

- Speckle patterns are complicated fingerprints of the internal structure of disordered media.
- Speckle patterns are extremely sensitive to extremely small structural changes in the medium. This is a consequence of the fact that there are a multitude of paths through the structure, and motion of a single scatterer affects a significant fraction of them.
- Classical speckle theory predicts that the mean intensity of a speckle pattern is equal to its variance, which results in extremely high contrast speckle patterns. However, classical theory completely neglects any correlation effects.
- Speckle patterns can be characterized by three correlation functions: C_1 , C_2 , and C_3 . C_1 is responsible for the memory effect, wherein a speckle pattern “remembers” the direction of the incident beam. C_2 and C_3 cause uniform positive correlations in intensity.
- In the case of a nonlinear medium (even with weak nonlinearity) each structure produces one of many speckle patterns when coherent light is incident on it. This is a consequence of the fact that the nonlinear Schrodinger equation has multiple solutions for the speckle pattern intensity. The number of solutions increases exponentially with increasing sample size.
- Although a nonlinear medium is not very practical for authentication purposes, there is a very strong analogy with cryptographic one-way functions. Weak optical nonlinearities may be produced fairly easily.
- Optical weak localization is accompanied by enhanced backscattering, while optical localization causes the light to stop diffusing. There is a phase transition as the medium goes from the diffusion regime to a localized regime. These kinds of phase transitions are usually accompanied by a dramatic increase in the physical complexity of the system.

6 Theory of physical one-way (hash) functions

The purpose of this chapter is to define physical one-way functions and physical one-way hash functions, present their properties, and compare them to their cryptographic counterparts. Like their cryptographic cousins, POWFs and POHFs have a strong asymmetry built into their definitions. However, unlike their cryptographic analogs, which convert strings of bits to other strings of bits, POWFs and POHFs operate on physical systems to produce strings of bits.

In section 6.1, we recall the definition for computational one-way functions proposed by Goldreich. In section 6.2, we provide a more general definition of physical one-way functions independent of any specific realization. We then show that coherent multiple scattering implements a physical one-way (hash) function in sections 6.3 and 6.4.

6.1 Computational one-way functions

We recall Goldreich's definition of one-way functions here (see section 2.1.2). A function $f: \{0, 1\}^* \rightarrow \{0, 1\}^*$ is called *strongly one-way* if the following two conditions hold.

- *Easy to compute:* There exists a deterministic \mathbf{P} -time algorithm A such that on input x , A outputs $f(x)$ (that is, $A(x) = f(x)$)
- *Hard to invert:* For every probabilistic \mathbf{P} -time algorithm A' , every positive polynomial p , and all sufficiently large n

$$Pr(A'(f(U)) \in f^{-1}(f(U))) < \frac{1}{p(n)} \quad 6.1.1$$

where U is a uniformly drawn input and both occurrences of U refer to the same value. This condition is referred to as collision-resistance.

The principal elements of this definition are:

- an input x , an output $f(x)$, and a \mathbf{P} -time algorithm A that, given x , outputs $f(x)$,
- an arbitrary algorithm A' which has a negligible probability of success in finding the inverse of $f(y)$ when y is chosen from a uniform density and,
- a robust notion of rareness.

If the function f produces a fixed-length output regardless of the length of the input, it is called a *one-way hash function* (OWHF). A OWHF has the following additional properties.

- *Variable input size*: f can be applied to an argument of any length.
- *Fixed output size*: f produces a fixed-length output.
- *High sensitivity*: Approximately half the bits in the output change when one bit changes in the input. This is the avalanche effect.

6.2 General definition of physical one-way functions

We provide general definitions of physical one way functions here. Our goal for this section is to define physical one-way functions without regard to any specific implementation.

Let Σ be a physical system in an unknown state $X \in \{0, 1\}^l$. X could also be some property of the physical system. l is a polynomial function of some physical resource such as volume, energy, space, matter *et cetera*.

Let $z \in \{0, 1\}^k$ be a specific state of a physical probe P such that k is a polynomial function of some physical resource. Henceforth, a probe P in state z will be denoted by P_z .

Let $y = f(X, P_z) \in \{0, 1\}^n$ be the output of the interaction between system Σ containing unknown state X and probe P_z .

6.2.1 Definitions

$f: \{0, 1\}^l \times \{0, 1\}^k \rightarrow \{0, 1\}^n$ is a *physical one-way function* if

- \exists a deterministic physical interaction between P and Σ which outputs y in $O(1)$, i.e. constant, time.
- Inverting f using either computational or physical means requires $\Omega(\exp(l))$ queries to the system Σ .

This may be restated in the following way. The probability that any probabilistic polynomial time algorithm or physical procedure A' acting on $y = f(X, P_r)$, where $y \in \{0, 1\}^n$ is drawn from a uniform distribution, is able to output X or P_r is negligible. Mathematically,

$$Pr[A'(f(X, P_r)) \text{ outputs } X \text{ or } P_r] < \frac{1}{p(l)} \quad 6.2.1$$

where $p(\cdot)$ is any positive polynomial. The probability is taken over several realizations of r

We also stipulate that for any physical one-way function f

- Simulating y , given X and P , requires either $O(\text{poly}(l))$ or $O(\exp(l))$ in time/space resources depending on whether f is a *weak* or *strong* physical one-way function
- Materially constructing a distinct physical system Σ' such that *its* unknown state $X' = X$ is hard.

6.2.2 Discussion of the definition

The above definition has two parts which closely parallels the two-part definition of algorithmic one-way functions. However, the meaning of the words “input”, “output”, and “function” should be carefully considered. To avoid any confusion, we specify what we mean by these words.

- *Input*: In the definition, “input” refers to the physical system *and* the probe which can be used to interrogate it. This is reflected in the two arguments of the function $f(\cdot)$. The physical system is represented by the unknown state X and the probe is represented by P_z .
- *Output*: The “output” is a set of measurements of the interaction between the physical system and the probe.
- *Function*: The “function” is the procedure by which the interaction takes place and the arrangement of the input and output with respect to each other.

Let us now look at each part of the definition.

- The first part of the definition posits a deterministic physical interaction between the probe and the system which produces the output in constant time. Why do we require this? This is the same as “easy to compute” in the definition of an OWF. Further, we show later in our embodiment of a POWF-based authentication system that this is indeed *possible*.
- The second component defines the “one-wayness” of POWFs. We require that there be no efficient algorithmic or physical procedure that, given the output of the function, is able to discover the unknown state X or probe P . In principle, an adversary should not be able to discover these two inputs even after running the procedure A' a feasible number of times.
- We then partition POWFs into two classes, *weak* and *strong*, depending on the difficulty of computing y given both a description of Σ and P_z . If simulating y is a polynomial time computation in the size of the unknown state, i.e., $O(\text{poly}(l))$, then we say that the function f is a weak physical one-way function. If this effort is exponential in the size of the unknown state, i.e., $O(\exp(l))$, then we say that the function is a strong one-way function.

We make this distinction because we would like an adversary neither to be able to discover X nor be able to discover any specific state of the

probe. If the simulation in the forward direction were $O(\text{poly}(l))$, then the adversary could potentially mount a brute-force attack to discover the probe state by simulating the interaction between the all possible probes and the given system. A strong POWF avoids this possibility.

- In part four, as in sections 4.2.1 and 4.2.4, we are faced with the problem of quantifying the difficulty of materially constructing a physical system which contains a specific unknown state. Henceforth, we will refer to this difficulty as *fabrication complexity* and will discuss it a subsequent chapter.

6.3 Coherent multiple scattering implements a POWF

The purpose of this section is to show that the interaction between a inhomogeneous 3D microstructure and coherent radiation can implement a physical one-way (hash) function.

6.3.1 Notation

Consider the schematic diagram in figure 6.1 which pertains to coherent multiple scattering from disordered microstructures.

The shaded triangle represents the range of input angles, wavelengths, and any modulation of the incident coherent radiation. We notationally represent the set of input wavelengths as $\Lambda = \{\lambda\}$ and the set of three-dimensional input angles as $\Theta = \{\theta\}$. The set of complex spatial light modulation patterns, which we assume to be bitmaps, are denoted by $B = \{I(p, q)\}$, where each element of B is a bitmap I of size $\max(p) \times \max(q)$.

Then, the sets of possible input angles, wavelengths, and spatial modulations may be represented by

$$P = \{(I, \lambda, \theta) | \lambda \in \Lambda, \theta \in \Theta, I \in B\} \quad 6.3.1$$

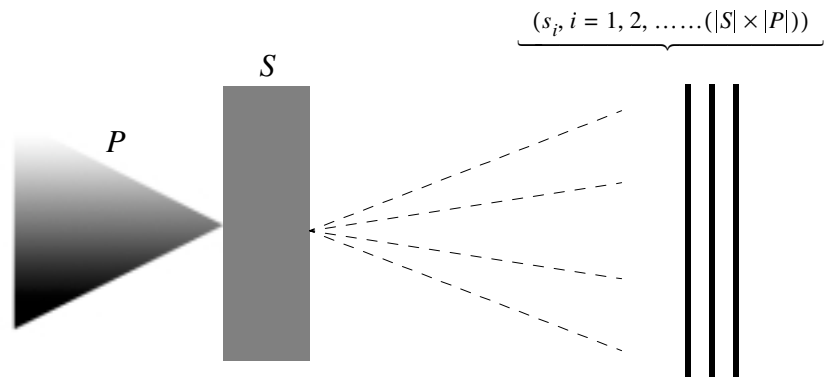


FIGURE 6.1 SCHEMATIC DIAGRAM USED TO DEFINE PHYSICAL ONE-WAY FUNCTIONS.

Each element of the set P is a distinct angle, wavelength, and modulation triad. We also assume that the elements of Λ and Θ are chosen so as to

produce uncorrelated speckle patterns. In other words, the elements of these two sets are spaced far enough apart from each other such that the memory effect (see section 5.3.2) has no influence on the corresponding speckle patterns.

The set of all inhomogeneous microstructures — of a given volume $V = L^3$ and mean free path between scatterers l — by S_T . Henceforth, we will abbreviate this to S unless there is a potential for confusion.

The set of all speckle patterns produced by all possible interactions between the elements of S and P as s , where $|s| \geq |S| \times |P|$. In the figure, we represent the set of speckle patterns by thick vertical lines.

We define:

- the *input* is the set of structures S and the set of probes P
- the *output* is the set of speckle patterns s
- the *function* is the interaction between P and S , i.e., the physical processes of coherent multiple scattering and wave propagation.

6.3.2 POHFs as sampled speckle patterns

A simple route to physical one-way hash functions is to sample the speckle fields $\{s\}$ on a regular grid. This is easily accomplished by recording the speckle patterns on a charge-coupled device (CCD) camera or a CMOS image sensor. It is important to ensure that the spatial frequency of the sampling grid is at least a factor of two greater than the maximum spatial frequency of the speckle patterns. This condition is necessary to ensure that no information is lost in the sampling process and no aliasing occurs. Assuming an ideal image sensor (no noise and large dynamic range), we now have several *fixed size* speckle patterns regardless of the size (in bits) of the 3D structure that gave rise to them. This reduction, from a variable sized 3D structure to a fixed size array of speckle intensities, may be regarded as a hash function. Hashing, therefore, is equivalent to sampling the speckle intensity patterns.

6.4 Heuristic arguments

We now show that coherent multiple scattering implements physical one-way hash functions as defined in section 6.2.1.

6.4.1 Easy to “compute”

It is obvious that there exists a deterministic physical interaction between the system and probe which produces an output speckle pattern in (almost) constant time. In practice, for disordered media, the physical probe produces output in almost constant time. This is because the size of the structure is such a small fraction of the distance light travels in a given time. For example, for a structure whose longitudinal dimension is 10 mm, the time taken to produce output is 3×10^{-11} seconds. As a matter of fact, given that we are working with length scales well below the absorption length (see section 5.2), we can

reasonably assume that the output is produced in constant time.

6.4.2 Hard to invert

We are now interested in answering the question: how hard is it to invert a speckle pattern to determine either the unknown state X or the probe P ?

Let us consider a few factors which determine the difficulty of inversion.

First, the exiting wavefront is spread out over a large solid angle but our implementation samples only a small fraction of this angle. In our experimental setup, the angle subtended by the CCD detector at the 3D microstructure is approximately 1° , but the speckle pattern is available over several tens of degrees. In sampling such a small solid angle we lose a lot of information about the wavefront, or, equivalently, about the structure.

Second, when we detect the speckle pattern, all phase information is destroyed. Even if one were to record the entire wavefront, not having access to the phase information makes inversion non-unique.

Third, we know that the speckle pattern is extremely sensitive to the configuration of scatterers in the structure. Even the motion of a single scatterer affects the speckle pattern drastically. We can quantify this by determining the number of possible structures that can be distinguished by a probe of a given wavelength.

The input space is the product space of all possible probes and all possible structures. Let us begin by enumerating the size of the input probe space and the number of possible structures.

Input probe space: The space of all possible probes depends on the number of probe angles, wavelengths, and complex amplitude modulations. We already know that, in the linear case, the sensitivity of the speckle pattern to changes in input angle is inversely proportional to the size of the structure and directly proportional to incident wavelength. In other words, the minimum angle that incident probe must be deviated by in order to produce an uncorrelated speckle pattern is inversely proportional to L . This suggests that the space of all probes is some polynomial function of structure size and wavelength. Thus,

$$|P| = \text{poly}(L, l, \lambda) \quad 6.4.1$$

Number of possible structures: How many different possible structures are there? We tackle this question here.

Assumptions

Size of the structure = $L \times L \times L$

Mean free path = l

Wavelength = λ

Number of scatterers = $N = (L/l)^3$.

This last statement assumes that the scatterers are uniformly distributed in the structure.

Argument

We proceed as follows.

- (1) Determine the total number of possible structures S_T
- (2) Determine the subset S_p of these which cannot be distinguished from each other by a probe

The probability that two structures produce the same speckle pattern is then given by $P \leq (S_p/S_T)$

We will show that P is negligible.

Total number of structures

First, let's consider (1) above.

If the structure were truly inhomogeneous, then we would have to describe the structure at the scale of a voxel of volume λ^3 . The total number of voxels is given by $b = (L/\lambda)^3$, and therefore the total number of structures is

$$S_T = 2^b \quad 6.4.2$$

where we have assumed that each voxel can be represented by one bit.

We can arrive at S_T for spherical particles by a different route.

Assume that the structure is divided up into cells of size $l \times l \times l$ and, with high probability, each of these cells contains no more than one scatterer. For the (reasonable) geometric approximation ($l \gg \lambda$), we may replace each sphere by a point scatterer located in a cubical box of volume l^3 . Pictorially, we have the situation described in figure 6.2.

In 3D, the number of ways of populating a cell with a point scatterer is $(l/\lambda)^3$ and there are $(L/l)^3$ such cells. Therefore the total number of possible structures is

$$S_T = \left(\frac{l}{\lambda}\right)^{3(L/l)^3} \quad 6.4.3$$

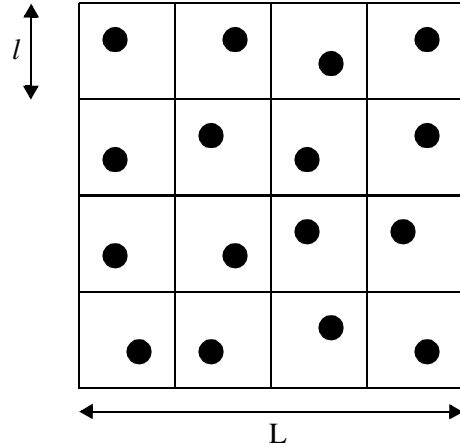


FIGURE 6.2 REPRESENTATION OF SCATTERERS IN 2D

Adopting the abbreviations $r = l/\lambda$ and $k = L/l$, we have

$$S_T = (r^3)^{k^3} \quad 6.4.4$$

It may be easily shown that the value of S_T in equation 6.4.4 is smaller than the one from equation 6.4.2. We will use the latter value hereafter.

Number of structures which are not distinguishable by a given probe

According to Berkovits [57], if more than k scatterers are moved from their original locations, the resulting speckle pattern is *almost uncorrelated* with the original speckle pattern. Ideally, we would like the two resulting speckle patterns to be *completely uncorrelated*, but the long-range correlation C_3 (see equation 5.3.18) precludes this situation. Even if we were to discretize the probes so as to make the C_1 and C_2 correlations irrelevant, the C_3 correlation persists.

The number of ways of selecting less than k scatterers from a collection of $N = k^3$ scatterers may be shown to be less than $e^k k^{2k+1}$.

This is true because the total number of ways w is given by

$$w = \sum_{i=1}^k \binom{k^3}{i} \leq k \binom{k^3}{k} \sim e^k k^{2k+1} \quad 6.4.5$$

Therefore, the number of possible ways to move scatterers such that there is no change in the speckle pattern is w . Equivalently, this is the number of structures that produce the same speckle patterns upon irradiation with the same probe. Therefore $S_p = w$.

Therefore, the probability we are looking for is

$$P = \frac{S_P}{S_T} = \frac{e^k k^{2k+1}}{(r^3)^{k^3}} \leq \frac{e^k k^{2k+1}}{(k^3)^{k^3}} \quad 6.4.6$$

Therefore,

$$P \leq \frac{e^k}{k^{3k^3 - 2k - 1}} \quad 6.4.7$$

Since P is smaller than any polynomial in k , it is negligible.

For $k = 5$ this probability is $P \sim 10^{-250}$ and for $k = 10$ it is $P \sim 10^{-2750}$. Thus, a microstructure of a given volume and containing a given density of scatterers can potentially “support” a number of structural configurations. This number of configurations is exponential in the size of the structure. Another way of looking at this is to say a given probe is able to distinguish a very large number of different configurations of scatterers in a given volume of material.

We can conclude from equation 6.4.7 that the probability P is *exponentially small* in the structural parameter (L/l) . This parameter may be increased by increasing L or decreasing l . This conclusion also satisfies the requirements of our intuition: as the size of the structure grows, the probability that two randomly produced structures produce the same speckle pattern decreases exponentially.

The above calculations show that the size of the input space, including the space of probes and the space of possible structures, is exponential in the size of the structure.

Another case for preimage resistance: We can also guess that the probability of finding two different structures S and S' with identical speckle patterns is low. We provide a heuristic argument here. We know that the number of scattering events per path in a structure is $(L/l)^2$, and the number of paths through the structure increases exponentially as (L/l) increases. Further, an appreciable fraction (l/L) of all paths pass through a given scattering site.

Assume that one such scatterer σ is now moved by a small distance from its original location. This affects the accumulated phase on an a fraction (l/L) of all paths through the structure. We can imagine adjusting the position of a single scatterer along each of those affected paths to nullify the effect of changing the location of σ and thereby producing an unchanged speckle pattern. However, since each of those scatterers also lies in the same (l/L) fraction of (possibly distinct) paths through the structure, this adjustment would cause further changes in the speckle pattern.

From this argument, it is clear that creating a second structure by modifying the positions of the scatterers in a given structure to produce the same speckle pattern is a very difficult problem. We infer, therefore, that physical one-way hash functions exhibit 2nd preimage resistance.

The above three points lead us to conclude that this specific implementation of a physical one-way hash function is *preimage-resistant* and *collision-resistant*.

6.4.3 Simulating the output

Given the unknown state X and all the probes, how hard is it to simulate y ? In the case of a linear structure, this is equivalent to a matrix multiplication where the $N \times N$ scattering matrix, the adversary's computational complexity is $O(N^2)$, which is the number of elements in the scattering matrix and the number of multiply operations required to obtain the *full* output. We note in passing that $N^2 \propto (1/\lambda^4) \sim 10^{24}$, a large number. The computational complexity of producing a *single* response to a challenge is in time $O(N)$.

Therefore, in our implementation, using a linear structure, we have built a weak POWF system. The complexity of simulation is a polynomial function of the size of the unknown state X , not an exponential one.

6.4.4 High-sensitivity

Because each scattering event affects an appreciable number of paths through the structure, and because each path encounters a large number of scattering events, each scatterer in the structure has an influence on the speckle pattern which is much larger than its own spatial dimension. Consequently, moving a single scatterer causes a large change in the speckle pattern. This corresponds to the avalanche effect exhibited by cryptographic one-way hash functions. We will experimentally demonstrate this avalanche in section 8.3.

6.4.5 Cloning the structure

We defer this discussion to a later chapter.

6.5 Summary

As we delve deeper into the simple phenomenon of coherent multiple scattering from inhomogeneous 3D microstructures, we begin to see the remarkable similarities between this physical mechanism and algorithmic one-way functions. We have seen that the physical process of coherent multiple scattering from three-dimensional inhomogeneous microstructures implements a physical one-way function as defined above. Further, we have discovered very strong parallels between the properties of physical one-way hash functions and algorithmic one-way hash functions.

7 System design and engineering

In this chapter, we focus our attention on the design and engineering of the entire physical authentication system, starting with the tokens and ending with the unique identifier.

7.1 What needs to be designed?

The following components of the physical authentication need to be designed.

- The *token* which encapsulates the physical system
- The *physical probe*
- The *reader* which encapsulates the probe and the detector
- The *algorithm* which converts raw data into the unique identifier

We recognize that each of these elements may take a significantly different form depending on the specific application context. In this dissertation, our archetypal token will assume the form of a credit-card. This is not a random choice. As we mentioned in the introductory chapter, we initiated the study of physical authentication systems based on the practical problem of providing unique, tamper-resistant, and unforgeable identifiers for smart cards.

7.2 Token design

7.2.1 Creating the microstructure

The physical system we used is a three-dimensional, inhomogeneous microstructure. We made this microstructure by curing micron-scale glass spheres in optical-grade epoxy. Specifically, we used precision solid-glass spheres as the scatterers. The glass spheres, manufactured by Cataphote, ranged in size from 500 microns to 650 microns and have a refractive index of approximately 1.5. On average about 90% of the spheres were true spheres. The procedure by which we made the microstructures is fairly simple. A small batch of optical-grade, transparent epoxy was mixed up and a small quantity of glass sphere was carefully stirred into the epoxy. Care was taken not to cause any systematic patterns in the epoxy due to stirring. The appearance of the final composite was milky white, indicating that multiple scattering was indeed taking place.

7.2.2 Making the token

The token was created from a sheet of plexiglass about 2.54 mm thick. We used a laser-cutter to produce a credit-card-sized form with a centered square aperture 10mm on a side. Thus the total volume of the aperture was 254 cubic mm. Additionally, three circular apertures of diameter 4.5mm were also cut out of the token. These apertures provide a means for registering the token in the reader.

The epoxy mixture was poured into the central aperture of the token and was allowed to set for a few hours. The final result of this process is depicted in figure 7.1. We note that these tokens are very easy to make, and it is not hard to see how the production process might be automated, in the event that a large number of these tokens are needed. In figure 7.2, we depict an earlier version of the tokens.



FIGURE 7.1 TOKENS USED IN THE PHYSICAL AUTHENTICATION SYSTEM. NOTE THE EPOXY SYSTEM IN THE CENTER AND THE REGISTRATION HOLES ARRANGED IN A TRIANGLE.



FIGURE 7.2 EARLIER TOKENS - NOTE THAT EACH ONE BEARS EVIDENCE OF A DIFFERENT REGISTRATION SYSTEM

7.3 Probe design

Obviously, there are a large number of candidate physical probes that may be used to probe the 3D microstructure. Each method of probing the structure has its own pros and cons. In this section we take a look at two such methods of peering inside the structure with a view to demonstrating their inadequacy for our purpose. Then we show how a simple laser beam can achieve the

performance we desire.

7.3.1 Optical coherence tomography (OCT)

The insight behind OCT is extremely simple: light with a very short coherence length that is sent down two distinct paths and later recombined will produce a strong interference signal if, and only if, the path lengths are identical. Consider the diagram in figure 7.3.

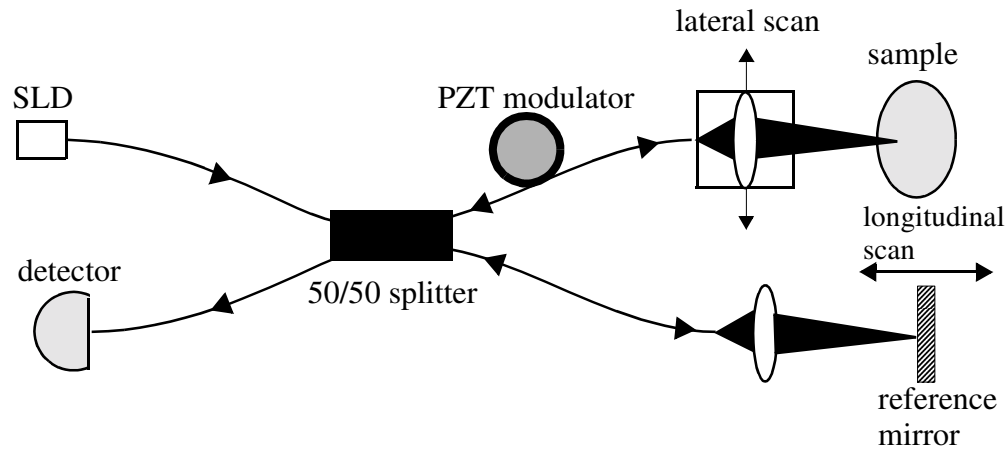


FIGURE 7.3 SETUP OF AN OPTICAL COHERENCE TOMOGRAPHY SYSTEM

Light from a source with a very short coherence length, e.g., a superluminescent diode, enters a standard Michelson Interferometer implemented with optical fibers. One beam reflects off a mirror on a longitudinal scanning platform and the other beam scatters off the sample and is modulated at some frequency determined by the piezoelectric transducer (PZT). As the mirror is scanned, different path lengths become equalized and scattered light from a specific depth in the sample adds coherently with the light from the mirror. Light from all other depths in the sample adds incoherently, producing an incoherent background. As the mirror is scanned, it is possible to acquire several 2D slices - hence the use of the word tomography - demodulate them, and computationally assemble them into a 3D image.

Huang et al. [27][28] report a longitudinal spatial resolution in air of 17 microns, a transverse spatial resolution of 10-30 microns, and for image acquisition time for 150 scans within a depth of 2mm to be approximately 190 seconds. Current high-resolution OCT systems have a longitudinal spatial resolution between 4 and 20 microns. The depth limit of OCT is determined by the regime where scattering predominates absorption, and the image quality decreases as the amount of multiply scattered light increases. Essentially, each scattering event “uses up” phase coherence and limits the

ability of the returned light to produce a strong interference pattern. Speckle limits the image quality as well.

OCT was our earliest choice for an imaging method to look inside the 3D microstructure. It has several advantages: it is a simple, inexpensive method of collecting 3D structural information from translucent structures. The light source required is a superluminescent diode and, because the returned light is detected interferometrically, the diode does not need to have a high luminous output. A typical image obtained from an OCT scanner is shown in figure 7.4.

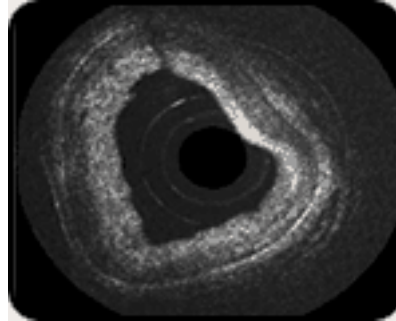


FIGURE 7.4 A TYPICAL OCT IMAGE. THIS IS AN INSIDE VIEW OF A HUMAN CORONARY ARTERY

However, as we examined OCT in the light of desired properties of the physical probe, it came up short in many respects. First, the longitudinal spatial resolution of OCT is on the 10 micron scale, and we were interested in exploiting smaller structural features. Next, each scattering center in the token, viewed in an OCT system, does not have any effect on the image beyond its own spatial dimension. This implies that changing a small volume of the token does not have a large effect on the image, and that the complexity class of simulating the effect of the OCT probe on the structure is **P**-time. Properties 8 and 9 of section 4.2.4 are not fulfilled when we employ OCT as the physical probe.

7.3.2 Magnetic resonance imaging (MRI)

MRI is an imaging technique used primarily in medical settings to produce high quality images of the inside of the human body. MRI is based on the principles of nuclear magnetic resonance (NMR), a spectroscopic technique used by scientists to obtain microscopic chemical and physical information about molecules [REF]. Because MRI images nuclear spin density, we assume that we have proton-filled glass spheres in the token instead of solid glass spheres. When a particle with net spin is placed in a magnetic field of strength B , it aligns itself in the direction of the field. This is the lowest energy configuration. However, there is another possible configuration in which the particle is aligned in exactly the opposite direction. This is shown in the figure 7.5. A particle may be knocked out from the lowest energy configuration into the higher-energy one by absorbing a photon at a precise frequency given by

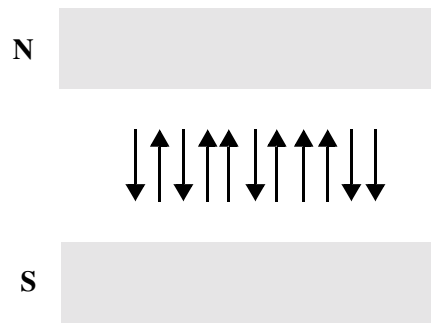


FIGURE 7.5 PARTICLES WITH SPIN ALIGN THEMSELVES IN ONE OF TWO WAYS. N AND S REPRESENT AN EXTERNAL MAGNETIC FIELD, AND THE ARROWS REPRESENT NUCLEAR SPIN ALIGNED EITHER WITH OR AGAINST THE FIELD

$\nu = \gamma B$, where γ is the gyromagnetic ratio and has units of frequency per unit magnetic field strength. For hydrogen, $\gamma = 42.58 \text{ MHz/Tesla}$.

If the volume containing the particles is in the magnetic field, and irradiated with electromagnetic radiation of frequency ν , all spins of the same species respond identically regardless of their spatial location. If, however, a linear magnetic field gradient is imposed on the volume, the spatial location of a set of spins may be encoded in the irradiating frequency. To see why this is so, consider the volume of spins shown in figure 7.6. We assume the field in the

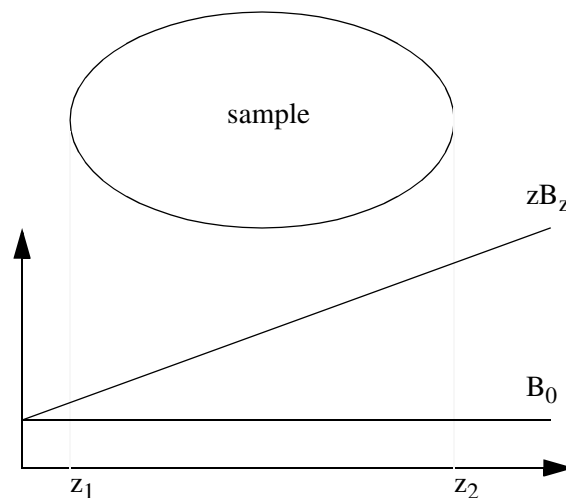


FIGURE 7.6 FREQUENCY ENCODING OF SPATIAL LOCATION. DIFFERENT SLICES OF THE SAMPLE CAN BE IMAGED BY IRRADIATING THE SAMPLE WITH A SPECIFIC FREQUENCY.

center of the magnet is B_0 with a corresponding frequency $\nu_0 = \gamma B_0$. In addition, a linear gradient is imposed along the z -axis as shown. This gradient may be represented as $G_z = zB_z$. Thus, the equation describing the response of the volume to incident radiation may be written as

$$\nu = \gamma(B_0 + zB_z) = \gamma B_0 + \gamma z B_z \quad 7.3.1$$

This gives us

$$z = \frac{(\nu - \nu_0)}{\gamma B_z} \quad 7.3.2$$

Thus, the spatial location of the set of spins is determined by the static magnetic field, the gradient field, and the gyromagnetic ratio of the particles. As in OCT, a 3D image is assembled from several 2D slices. Although this explanation is pedagogically useful, imaging is almost never carried out in this way in the real world. We will not concern ourselves with the details of real-world MRI systems, but move on to point out why MRI is not an attractive option as far as physical authentication systems are concerned.

First, MRI machines are expensive, and although there is an active effort to make tabletop MRI machines, they are still likely to cost several hundred dollars. At this cost, it would be a tremendous expense to deploy them wherever transactions are carried out. Second, spatial resolution of MRI scanners depends linearly on the strength of the magnetic field. Existing high-resolution MRI scanners, which use a static field of 4 Tesla, and a field gradient of 0.1 Tesla/meter, have a spatial resolution in the range 150-300 microns and slice thicknesses in the range of 300-600 microns. However, achieving these resolutions is extremely time-consuming, with each scan lasting several hours. These resolution and time constraints rule out MRI as a candidate probe. Finally, each element of the volume has a one-to-one correspondence with the image. As we noted earlier, this means that a small change in the token leads to a corresponding small change in the image - a feature which violates requirements 8 and 9 of section 4.2.4.

7.3.3 Laser beam

We now turn to the probe actually used in our physical authentication system - a collimated laser beam of diameter approximately 1 mm of wavelength 632.8 nm. The beam originates from a commercially available 30mW Helium-Neon laser manufactured by Melles-Griot. All we do is shine the laser beam at the 3D microstructure and detect the emerging wavefront, known as a *speckle pattern*. This is conceptually depicted in figure 5.1.

We will spend the whole of chapter 5 discussing interaction between the laser beam and structure, so we will not dwell on the details here. However, we

walk through the desirable qualities of using a laser beam as the probe. First, lasers are extremely inexpensive, and start up in the same state each time. As we will see, we do not require the laser beam to have the same *absolute* phase each time it starts up, we merely require that the *relative* phase between a beam which is scattered along multiple paths in the structure be predictable. This is always true if the distance the scattered beams travel is always less than the coherence length of the laser, and we will impose this condition on our system.

Lasers can be mass-produced with ease - indeed, every compact disc player has a semiconductor laser in it that was manufactured for less than a dollar. Lasers are now an indispensable part of everyday life, and the physical authentication systems we propose in this dissertation make use of garden-variety lasers.

The speckle pattern is the result of coherent interference between light that has taken multiple paths through the inhomogeneous microstructure, and we show in the next chapter that it is extremely sensitive to infinitesimal changes in any single path. This is primarily because we are detecting output that is intimately dependent on *phase coherence* of the radiation. Any property not related to phase coherence (such as total capacitance) would be far less sensitive to changes in the structural configuration of the token. This notion is mathematically formalized in the next chapter.

In summary, when compared to both OCT and MRI, a system wherein a simple laser beam probes the structure incorporates all the properties we require in a physical authentication system.

7.4 Reader design

We now turn our attention to the reader, which is the object that encapsulates the laser and the detector and allows us to present the token for interrogation. We begin by prescribing the mechanical requirements of an ideal reader, and then describe the reader we implemented. Finally, we take a look at the performance of the reader.

7.4.1 Mechanical requirements

The reader has to accomplish three things. First, it has to allow for accurate and repeatable positioning of the laser beam. For reasons that will become clear later, we are also interested in interrogating the structure from multiple angles. The laser positioning requirements extend to this case as well. Next, the reader has to allow us to present the token to the laser beam with a minimum of misregistration. Finally, it has to provide a detector for the speckle pattern whose position is invariant with respect to the rest of the system.

- *Laser positioning*: The limits on laser positioning performance are set by the laser itself and the hardware which controls the angular travel of the beam across the structure. The limits on the position of the structure with respect to the beam are set by the C_1 , C_2 , and C_3 correlations discussed in the next chapter. Essentially, when the laser beam is rotated with respect

to the token by more than $\Delta\theta_{max} = 1/(kL) = \lambda/(2\pi L)$ we obtain an independent speckle pattern. In our case, $L = 2.54$ mm and $\lambda = 632.8 \times 10^{-6}$ mm, which means that $\Delta\theta_{max} = 40 \times 10^{-6}$ radians. Therefore, in order to obtain the same speckle pattern each time, the laser beam has to be incident on the microstructure at an angle that never deviates from its previous value by more than $\Delta\theta_{max}$.

- *Token registration:* For the reasons discussed in the foregoing paragraph, the token must also be repeatably placed in the same position each time the token is presented to the system. Any changes in the spatial location of the token are tantamount to changes in the incident angle of the laser beam, and cause the same deleterious effect of producing an independent speckle pattern.
- *Detector positioning:* In our implementation, this is the easiest to achieve. The CCD camera is simply mounted in a fixed location with respect to the token. Because the detector is not a moving part this is a one-time alignment effort.

7.4.2 Implementation

The implementation of the reader is an extremely crucial determinant of system performance. We built several readers which led to sub-optimal performance. In each of these cases, it was the token registration system that did not perform as expected. We begin by providing pictorial evidence (figures 7.7 and 7.8) of these attempts and then discuss the final implementation in detail.

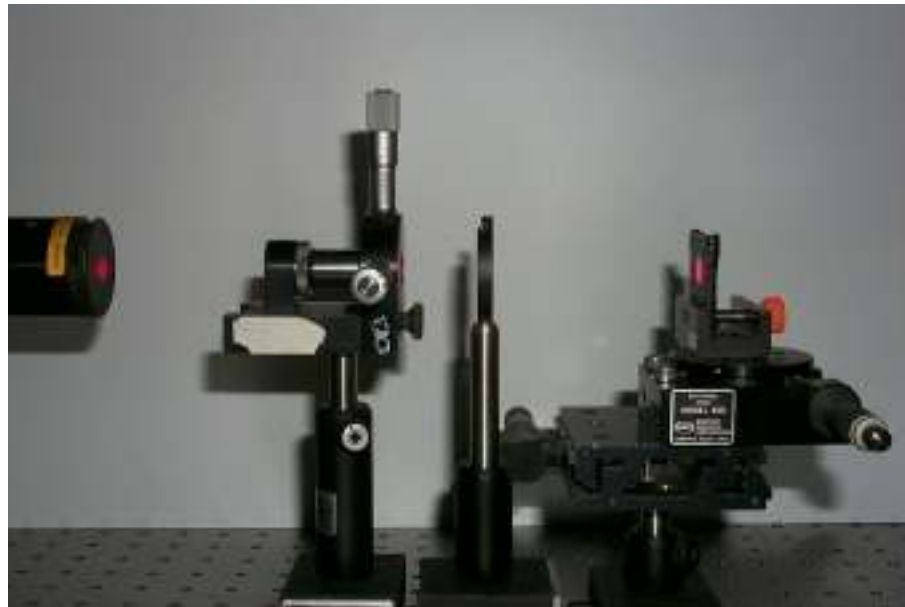


FIGURE 7.7 THE EARLIEST PHYSICAL AUTHENTICATION SYSTEM

Laser positioning: Our choice of laser positioning hardware consisted of a front-surface mirror mounted on a precision gimbal mount which was driven



FIGURE 7.8 ANOTHER INSTANTIATION OF THE SYSTEM

by a two high-precision actuators. Essentially, a laser beam from a He.Ne laser reflects off the mirror and is incident on the token. As the actuators travel, the location of the laser beam on the token scans the area of the structure. An image is shown in figure 7.9.

The Newport 850G actuators are capable of 1 micron bidirectional repeatability and are both driven by the Newport ESP300 motion controller, which is in turn controlled by a personal computer. The motion controllers have an extensive command set and can be directly controlled via an RS232 serial port. We found that this setup has excellent repeatability, and all the data we present in this dissertation was obtained with it.

Token registration: As we said earlier, token registration was a thorny problem and several attempts were made to solve it. The early attempts are visible in figures 7.7 and 7.8. We reiterate the goal of the token registration here: *each time the token is placed in the reader, it must be as close as possible to the same absolute position in 3D space.*

We achieved this goal by using ideas drawn from Thorlabs' kinematic

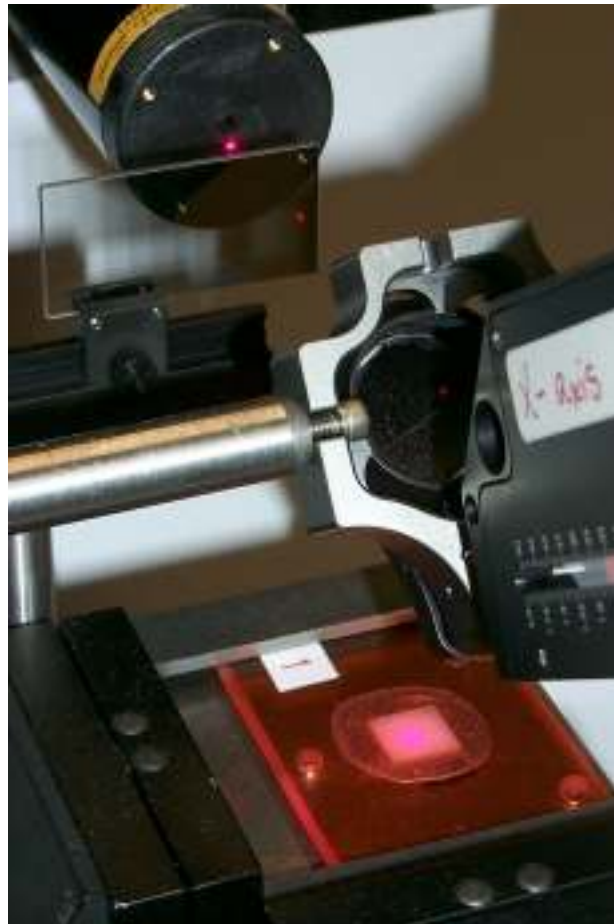


FIGURE 7.9 TWO VIEWS OF THE LASER POSITIONING SYSTEM

mounts, which use the cone-groove-flat mechanism to constrain 3D position. Kinematic mounts consist of two plates: a top mounting plate and a bottom base plate which are coupled by an extremely strong magnet. When the bottom plate is fixed, the top plate can be removed and replaced, repositioning to the same exact position with the repeatability of a microradian. We used two such mounts, and mounted their bottom plates vertically at right angles to each other so that they formed a single unit. This is shown in figure 7.10.



FIGURE 7.10 KINEMATIC MOUNTING UNIT

The two top plates were mounted on a token mount that was produced in-house. The right-angle kinematic unit allows us to place the token at very precise height above the camera. It is the job of the token mount to provide very accurate and repeatable x-y positioning. This was achieved by making a steel mount with three spring-loaded titanium steel balls arranged in a triangle. In this arrangement, the token, which has circular apertures in exactly the same triangular arrangement, simply slides over the balls till the circular apertures are spatially co-located with the balls. At this point, the balls push the token up against two retaining brackets. This is seen in figure 7.11. We used titanium-steel balls because of their very small deviation from sphericity, and used high force-constant springs with flat ends in order to support the balls perfectly. The channels in the token mounts were also drilled out using high-precision bits in order to ensure that there was no eccentricity as the balls moved up and down.

7.4.3 Performance

The performance of the token-reader is best ascertained by performing an experiment wherein a token is placed in the reader and a reference speckle pattern is obtained. Then the token is removed and replaced in the reader several times - obtaining a new speckle pattern each time. The angle of illumination is kept constant during this process. All the subsequent speckle patterns are compared with the reference speckle pattern to see if there are any



FIGURE 7.11 TOKEN MOUNT - WITH AND WITHOUT A TOKEN PRESENT.

systematic deviations.

We performed the experiment with two different tokens and obtained six different speckle patterns for each token. In order to examine if there were any systematic misregistration problems, we plotted the intensity along a single row (row 120) and a single column (column 120) of all six related speckle patterns. These plots were overlaid on the same axis. The results of this experiment are shown in figures 7.12 and 7.13.

Inspection of the four plots reveals that the general shape of each set of plots

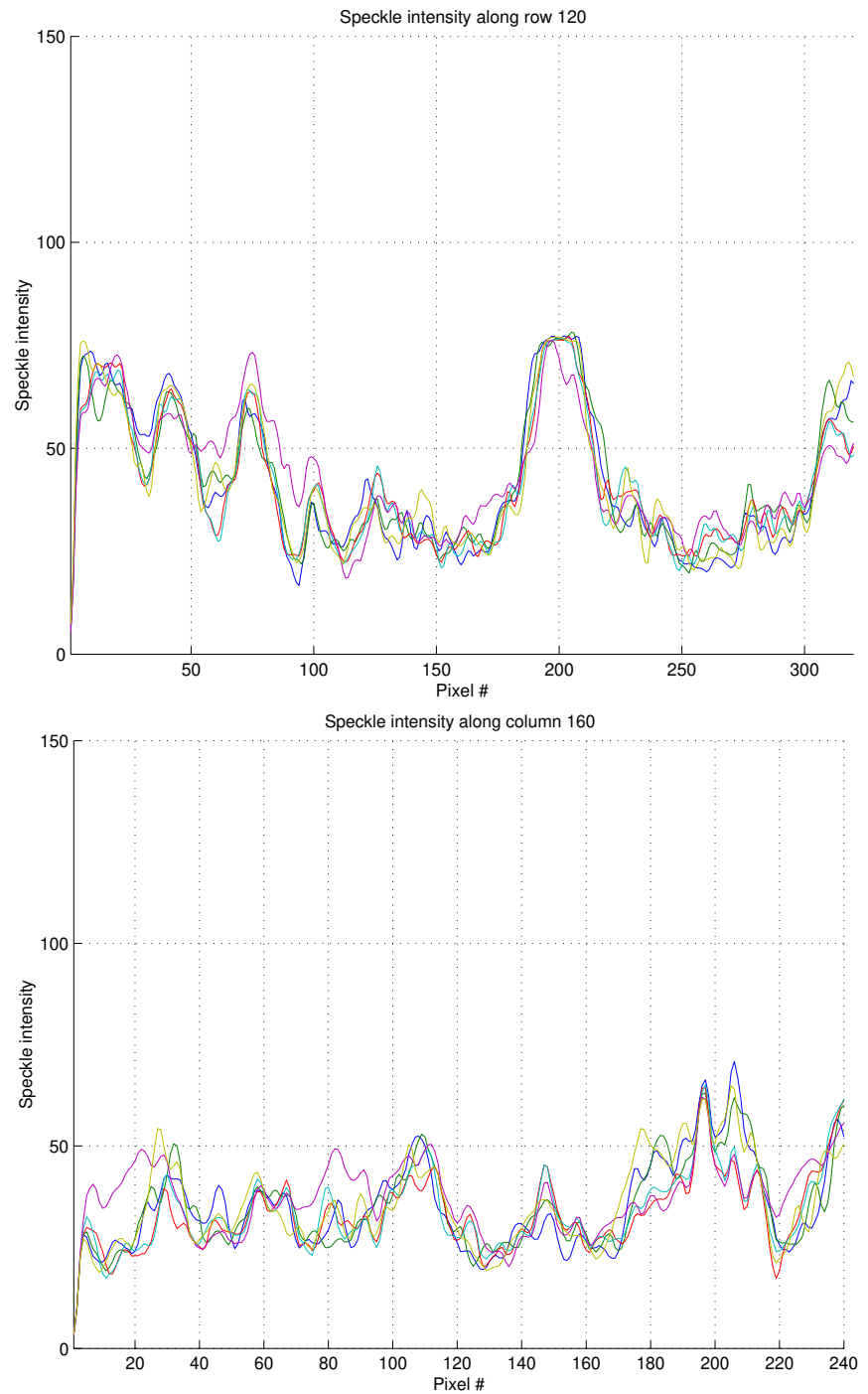


FIGURE 7.12 INTENSITY PLOTS ALONG A SINGLE ROW AND COLUMN OF SIX SPECKLE PATTERNS OBTAINED FROM THE SAME MICROSTRUCTURE

is the same, although there are local differences in speckle intensity. It is noteworthy that there does not appear to be any systematic offset of features between members of each set of plots. If there were such offsets, either in the

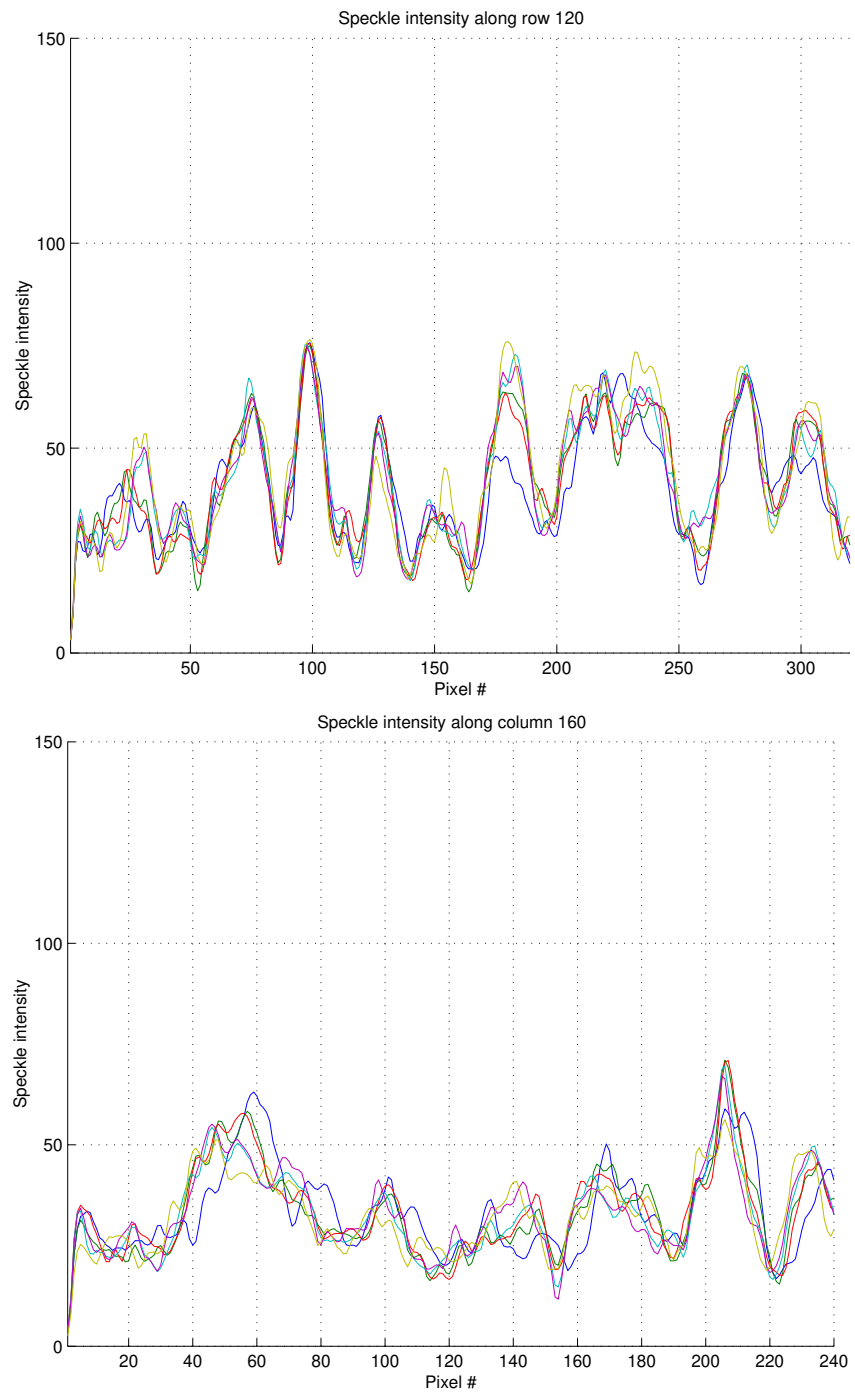


FIGURE 7.13 THE SAME PLOTS FOR A DIFFERENT MICROSTRUCTURE

row plots or the column ones, this would be cause for concern. However, the plots indicate that the token registration system performs very well, given the sensitive dependence of the speckle pattern on the relative positions of the token and the laser beam.

The local variations in speckle intensity are attributable to change in ambient illumination, fluctuations in laser power, photon noise, and noise in the CCD detector. Interposing a chopper in the beam path, using a light-tight enclosure for the entire system, and using a low-noise CCD detector will alleviate this problem.

7.5 The Gabor hash algorithm

We now focus on the hash algorithm *A* (see figure 4.1) whose function is to take the raw speckle data and boil it down to string of bits. Here we will first prescribe the desired qualifications of such an algorithm, present the theoretical background for our choice - the Gabor Transform - and show how we used it to produce a unique identifier from a speckle pattern. Hereafter, we refer to the algorithm as the *Gabor Hash Algorithm* and the unique identifier as the *Gabor Hash*. We end the section by presenting an inventory of the issues which must be considered in our implementation.

7.5.1 Desired features

What are the desired qualities of the hash algorithm? First we look at the qualities that any algorithm must possess regardless of the implementation, and then we look at the qualities specific to our implementation.

Implementation-independent qualities

- *Efficiency*: The algorithm must have a computationally efficient implementation. For reasons that will become clear later, we expect to run this algorithm several times during a single authentication session, and a fast algorithm that produces identifiers from speckle patterns is absolutely essential.
- *Distinguishability*: The algorithm must be able to distinguish two speckle patterns based on their features. In the ideal case, the algorithm is such that some distance metric between two identifiers is maximized when they are derived from two distinct speckle patterns and zero when they are from the same speckle pattern.
- *Analytic expression*: The algorithm must be amenable to mathematical analysis. This allows us to explore its properties and characterize its performance analytically.

Implementation-dependent qualities

- *Insensitive to changes in ambient light level*: The algorithm should be insensitive to changes in global ambient illumination. Mathematically speaking, the algorithm should have no dc response. This is essential because there is no guarantee that the average light level in our system will remain constant over time. The algorithm must take this variation into account.
- *Insensitive to token misregistration*: Ideally, we would have a perfect token registration system which allows us to reproduce the same exact speckle pattern each time the same token is inserted into the system. However, given the sensitivity of the speckle pattern to changes in the

relative position of the laser beam and the token, it is likely that the speckle pattern suffers small spatial transformations (translations or rotations in the horizontal plane). We will require our algorithm to be insensitive to these small changes.

- *Scale selection*: Finally, we would like the algorithm to be flexible enough to accommodate changes in the position, size, and orientation of the features of the speckle pattern. The position, size, and orientation of the features depend intimately on the optical system that produces them, and clearly, depending on the application context, we will use very different optical systems to produce the speckle patterns. Our algorithm must allow us to make changes in the optical system without an attending performance penalty.

7.5.2 Theory of Gabor Transforms

In the light of the requirements above, and after trying out several other methods, we decided to use the Gabor Transform as the algorithm. Gabor functions have historically been used in a wide variety of applications: image enhancement, coding, and compression [29][30], texture analysis [31], and motion analysis [32] to name a few. Another big area of use has been in the field of multiscale image representation in the visual cortex, primarily because their basis functions bear a strong resemblance to the receptive fields of simple cortical cells [34].

In our work, we use the 2D Gabor Transform proposed by Daugman [35] which itself is an extension of the 1D transform proposed by Gabor [36]. The elemental Gabor Function (GF) has the functional form

$$g(x_0, y_0, f, \theta) = \underbrace{e^{-[\pi(a^2(x-x_0)^2 + b^2(y-y_0)^2)]}}_{\text{Gaussian}} \underbrace{e^{[i2\pi f(x\cos\theta + y\sin\theta)]}}_{\text{Sinusoid}} \quad 7.5.1$$

The first grouped term is simply an elliptical 2D Gaussian function located at (x_0, y_0) where a determines the effective width along the x -axis and b determines it along the y -axis. The second grouped term is a complex 2D sinusoid of frequency f and an orientation defined by θ . Clearly, the GFs can be freely tuned to a continuum of spatial locations, spatial frequencies, and orientations by varying the parameters. This enables a GF to select features from an image at a specific location, scale, and orientation.

The GT is a specific case of a more general image processing technique usually referred to as *multiresolution image decomposition* or *pyramidal decomposition* [37][38]. In a pyramidal decomposition scheme local operators at several scales but with identical shape serve as the basis functions. Usually the operators (of which $g(x, y, f, \theta)$ in equation 7.5.1 above is an example) are localized both in space and spatial frequency. The basic method by which such a decomposition, in an image-encoding context, takes place is as

follows.

Consider an image I_0 , and a low-pass filter L . We apply L to the image to produce a low-pass filtered version of the image I_1 . We then have a residual defined by $R = I_1 - LI_0$. Instead of encoding the raw image I_0 , it is much more efficient to encode the residual and the low-pass filtered image. The pixels of the residual have much smaller dynamic range and entropy than the raw image, while I_1 may be encoded at a lower sample rate. This process may be iterated by applying the filter to a resampled I_1 and storing the residual. As the process is iterated, we are left with a small image which represents the repeated low-passed and resampled image, and a set of residuals of decreasing size. This representation is called the multiresolution pyramid.

In practice, the same filters are applied to low-passed and downsampled images for reasons of computational efficacy. If the local basis functions are from an orthogonal family of wavelets, the decomposition and subsequent are computationally efficient. In our case, the Gabor Functions are non-orthogonal, which leads to problems in reconstruction. However, we are not concerned with reconstituting the speckle from the unique identifier. All we are interested in is to go from the speckle pattern to the identifier. The coefficients of the GT are obtained by simply convolving the raw speckle pattern with filters which are obtained by translating the *mother wavelet* in equation 7.5.1 across all locations of the pattern.

Prior work [35] has determined that 2D quadrature filters, such as the one in equation 7.5.1, are jointly optimal in providing the maximum possible resolution for information about the orientation and spatial frequency content of local image structure (“*what?*”), simultaneously with 2D location (“*where?*”). This property is very useful when studying image texture, as we do in this dissertation.

7.5.3 Implementation to derive unique identifier

Because the mother wavelet is a complex function, the transform has both real and imaginary parts. In our work we focus exclusively on the imaginary part of the transform because the basis functions are odd functions, and therefore do not respond to changes on the ambient light level in the speckle image. This is a very useful property because the ambient light level is usually prone to small fluctuations due to either laser power variation or changes in the lighting of the environment.

Our implementation of the GT closely parallels the method proposed by Nestares et al. [39]. They developed an optimized spatial-domain representation using 1D masks which are reproduced in figure 7.14. The 2D transform is computed as the outer product of two fast 1D convolutions. Although the orientation parameter is continuously tunable, we chose to use four orientations given by $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$. This lets us select structure in the horizontal and vertical directions and the two diagonal directions. We use the same pair of even-odd 1D masks for the horizontal and vertical directions, and a single pair for the diagonal directions.

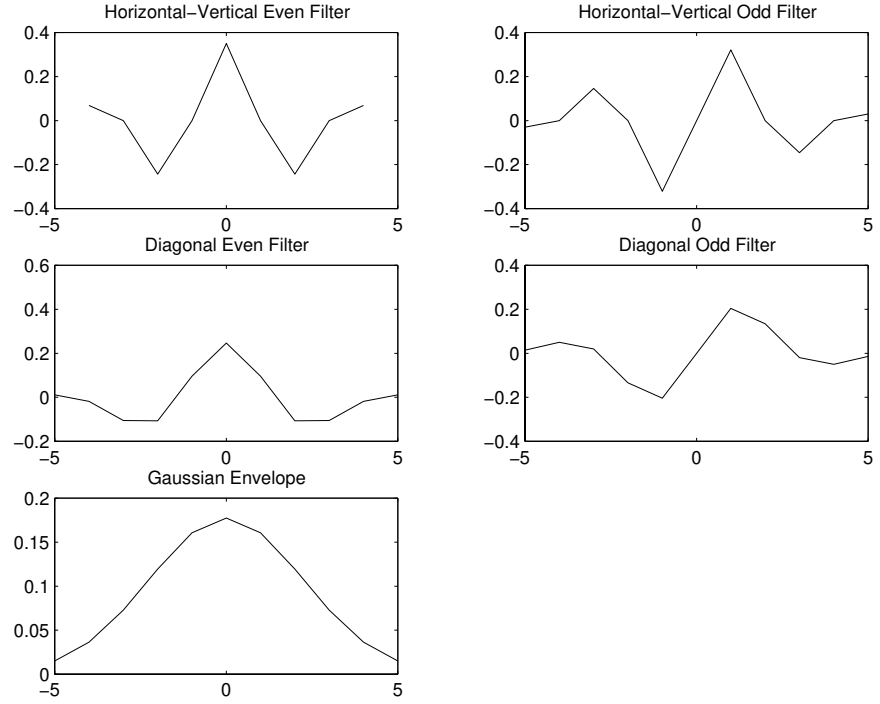


FIGURE 7.14 1D MASKS USED IN THE PYRAMIDAL GABOR DECOMPOSITION

Our procedure for generating the bit string from the raw speckle pattern proceeds as follows. First, we compute the imaginary part of the GT for four orientations and several levels. This is accomplished by convolving the image with the kernel. The analytical expression for this convolution is:

$$\iint I(x, y) g((x_0 - x), (y_0 - y), f, \theta) dx dy \quad 7.5.2$$

where θ variation is done independently, and the variation in f comes from image subsampling. It is worth noting that although we talk of scaling the basis functions in the spatial frequency domain, in practice we scale the image instead. The results are identical but the latter case is computationally much more efficient.

We then choose the imaginary part of equation 7.5.2 and threshold at zero to produce a new binary image. We repeat this procedure for all orientations and levels.

$$B(x, y) = 1 \text{ if } \text{Im}(\iint I(x, y) g((x_0 - x), (y_0 - y), f, \theta) dx dy) \geq 0 \quad 7.5.3$$

$$B(x, y) = 0 \text{ if } \text{Im}\left(\iint I(x, y)g((x_0 - x), (y_0 - y), f, \theta)dx dy\right) < 0 \quad 7.5.4$$

Assume we start with a 240×320 speckle pattern. At the N th level, we have 4 images each of size $(240/N) \times (320/N)$. The total number of available bits after transforming and thresholding is given by

$$\sum_{\theta=1}^4 \left(\sum_{k=1}^N \left(\frac{240 \times 320}{k^2} \right) \right) \quad 7.5.5$$

Two questions remain. First, what subset of these bits comprise the identifier? Second, how many orientations should be used?

In our implementation, only the two diagonal orientations are used. This is primarily because the values of the Gabor Transform along the diagonals is much less sensitive to small changes in the $x - y$ positioning of the token. However, we point out that the performance of the registration system makes this choice unnecessary.

We also use the coefficients only from the 4th level of the transform. At this level, each orientation has 30×40 bits, and since we use two orientations, we have a total of 2400 bits available to contribute to the identifier. Our choice of the level is driven by two competing issues. At level 1, we have as many coefficients as there are pixels in the image. However, the coefficients are sensitive to intensity variations on the scale of a single pixel, which can be quite high, given that we are using a garden-variety CCD camera. At a high level, the single pixel variations get averaged out, and allow for a much more robust identifier. However, this limits the number of available bits. As we see from equation 7.5.5, the number of bits decreases as the square of the level. We will demonstrate this tradeoff presently. The entire data pipeline is depicted in figure 7.15.

7.5.4 An example

In this section we provide an example of the functionality of the data pipeline with a view to demonstrating the tradeoff between number of bits and robustness.

We start with a speckle pattern P obtained from a token (figure 7.16). This pattern is Gabor transformed to level 4 and for 4 orientations, and the imaginary part of the transform is retained (figure 7.17).

The images are then thresholded and only the two diagonal images at level 4

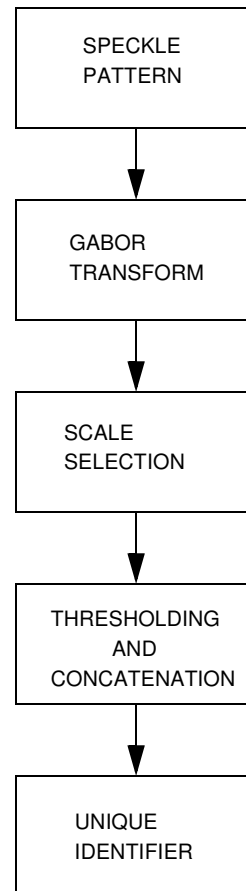


FIGURE 7.15 DATA PIPELINE: FROM SPECKLE PATTERN TO UNIQUE IDENTIFIER

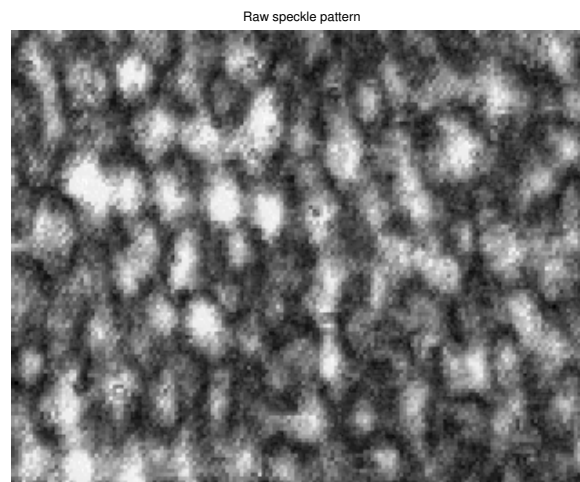


FIGURE 7.16 RAW SPECKLE PATTERN

are retained (figure 7.18). These two 30x40 retained images are treated as a long string of 1200 bits each and concatenated to produce a 2400 bit identifier.

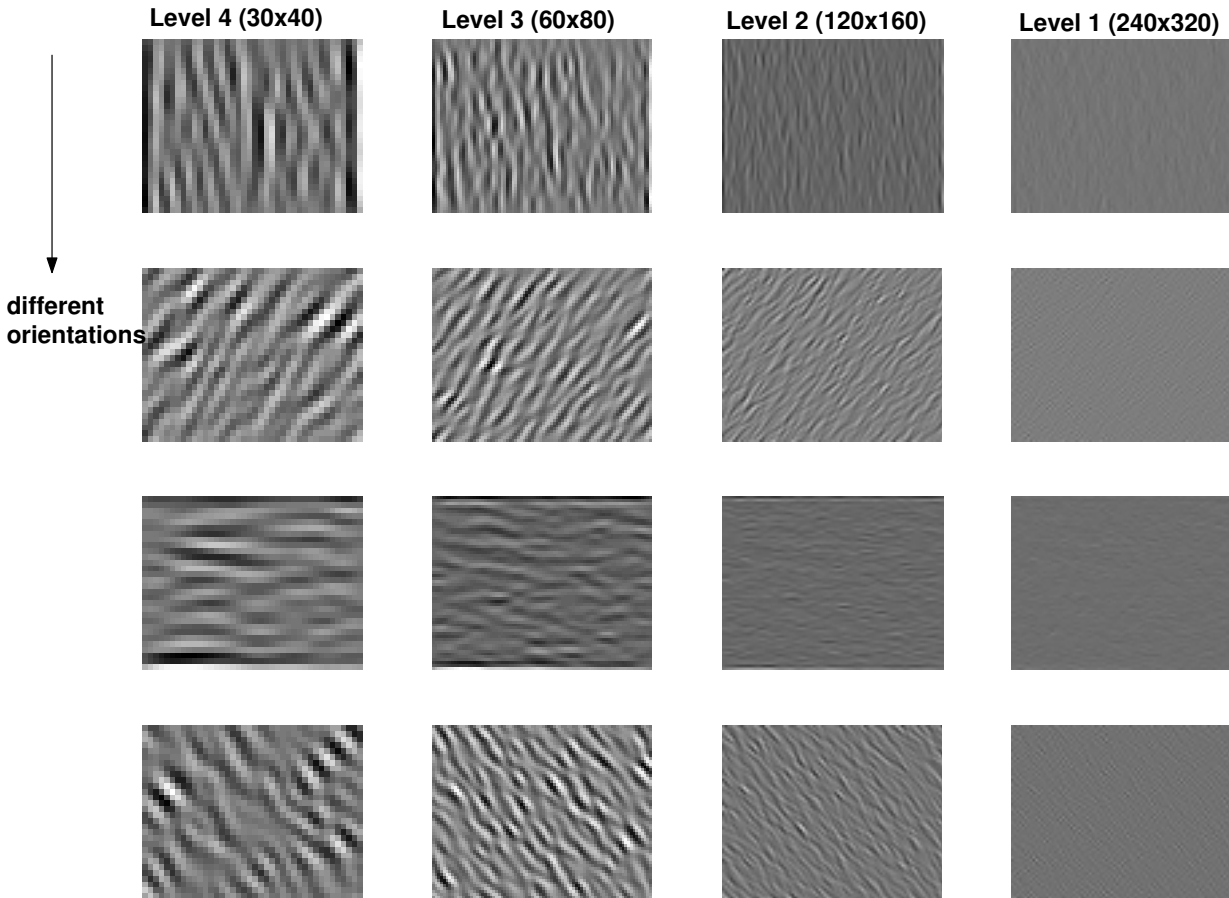


FIGURE 7.17 THE GABOR TRANSFORM - FOUR LEVELS AND FOUR ORIENTATIONS, IMAGINARY PART ONLY

7.5.5 Tradeoffs

We now proceed to demonstrate the tradeoff between level of analysis, number of bits in the identifier, and robustness. The experiment proceeds as follows. A reference speckle pattern is obtained and a bit string is derived from it at *all four levels*. Thus, we now have 4 strings of lengths 153600, 38400, 9600, and 2400. Let us denote these strings by R_1 , R_2 , R_3 , and R_4 respectively. The token is removed from the reader and replaced on five subsequent occasions, and 4 identifiers are derived from each speckle pattern. We denote the set of five identifiers at a particular level by D_{ij} where $i = \{1, 2, 3, 4, 5\}$ denotes a new speckle pattern and $j = \{1, 2, 3, 4\}$ denotes the level.

We then determine the fraction of bits that disagree (the *Hamming Distance*) between a reference identifier (one of R_i , where $i = \{1, 2, 3, 4\}$) and all the identifiers of the same length from the set D_{ij} . In the ideal case, with a perfect token registration system and no changes in the illumination or environment, we expect the fraction of disagreeing bits to be identically zero. The table below summarizes the results of the experiment. At level 1, the Fractional Hamming Distance (FHD) is in the region of 0.5, which means that approximately 50% of the 153600 bits disagree between the reference

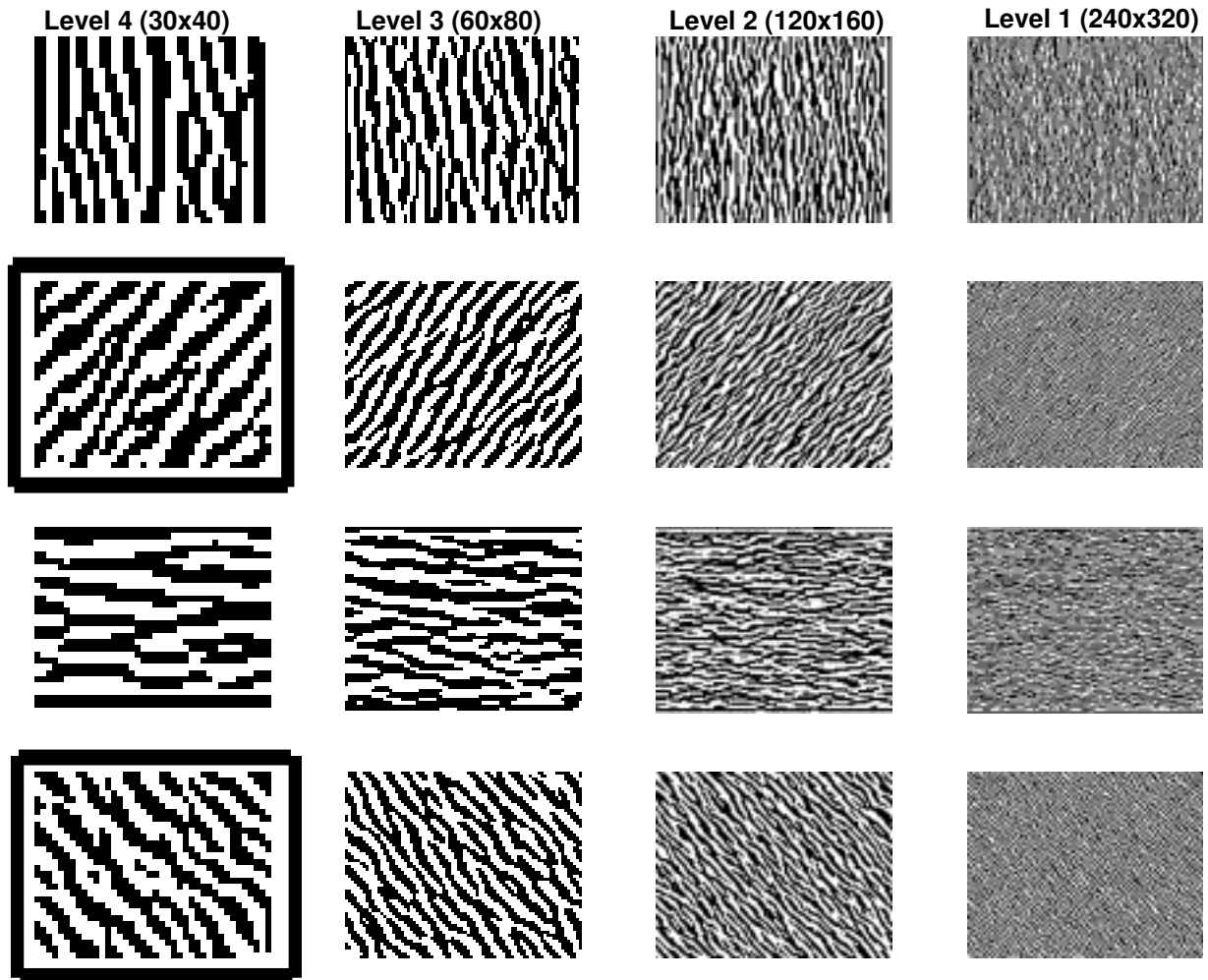


FIGURE 7.18 THE TWO THRESHOLDED IMAGES SELECTED ARE SHOWN HERE

FRACTIONAL HAMMING DISTANCES AT FOUR LEVELS

	LEVEL1	LEVEL 2	LEVEL 3	LEVEL 4
INSTANCE 1	0.4913	0.4518	0.1922	0.1008
INSTANCE 2	0.4361	0.4716	0.1919	0.0938
INSTANCE 3	0.4955	0.4410	0.2241	0.1300
INSTANCE 4	0.5136	0.4272	0.2309	0.1333
INSTANCE 5	0.5142	0.4094	0.2484	0.1833

identifier R_1 and the set of five identifiers D_{i1} . In other words, the results of repeating the experiment are the same as populating the identifier with random bits obtained by a coin-flipping experiment. Clearly, there is no way to distinguish such a random bit string from the identifier derived from the

speckle pattern.

We observe that the FHD steadily decreases as the level increases. At level 4, the FHD is, on average, about 0.1 for a 2400 bit string. This means that approximately 240 bits disagree between the reference identifier and identifiers derived from subsequent instances. We will see later that the average FHD between identifiers derived from distinct speckle patterns is 0.5.

Clearly, there is a tradeoff here between the level of analysis, the number of bits in the identifier, and robustness of the identifier. For an increase in the analysis level by one, the number of bits in the identifier decreases by a factor of 4. However, the FHD decreases rapidly from a maximum value of about 0.5 to 0.1. We also observe here that the benefits of increasing the level of analysis do not accrue indefinitely. As the level continues to increase, the number of bits decreases, and several speckle patterns map to the same configuration of bits. Thus, the level of analysis must be chosen so as to provide the maximum number of bits while maintaining statistical distinguishability.

7.6 Final system

In the preceding section, we have described each subsystem in detail. Here we briefly look at how they all fit together. In summary, we have a token which is mechanically positioned in a token reader. The token reader consists of an accurate laser positioning system, a token registration system and a CCD detector. The output of the token reader is a raw speckle pattern which is processed by an algorithm to derive a unique identifier. The algorithm is based on a multiresolution Gabor pyramid decomposition of the speckle pattern. This multiresolution decomposition is extremely flexible and allows complete control over the analysis of the speckle pattern. This flexibility is especially important if the optical configuration changes. The full system is shown in figure 7.19.

We remind the readers of an important point: there are two levels of hashing going on. The first is from the 3D microstructure to the speckle pattern. Physics, or more specifically, coherent multiple scattering allows this to happen. The second level of hashing is from the speckle pattern to the Gabor Hash. This procedure is reversible and may be regarded as merely a thresholding scheme. All the hard work is done by nature in the first hash, and the second hash may be replaced by any other threshold scheme.

7.7 Potential improvements of the system

In this final section of the chapter, we look at improvements that could be made to the reader to render it more effective.

One area where significant improvement is possible is to replace the rather slow motor controllers with a faster laser positioning system. One approach which immediately comes to mind is to use a digital micromirror device (DMD) produced by Texas Instruments. A DMD is a thumbnail-sized silicon chip that contains thousands of individual square mirrors, each about 10 microns on a side, which can be switched digitally. It is accurately

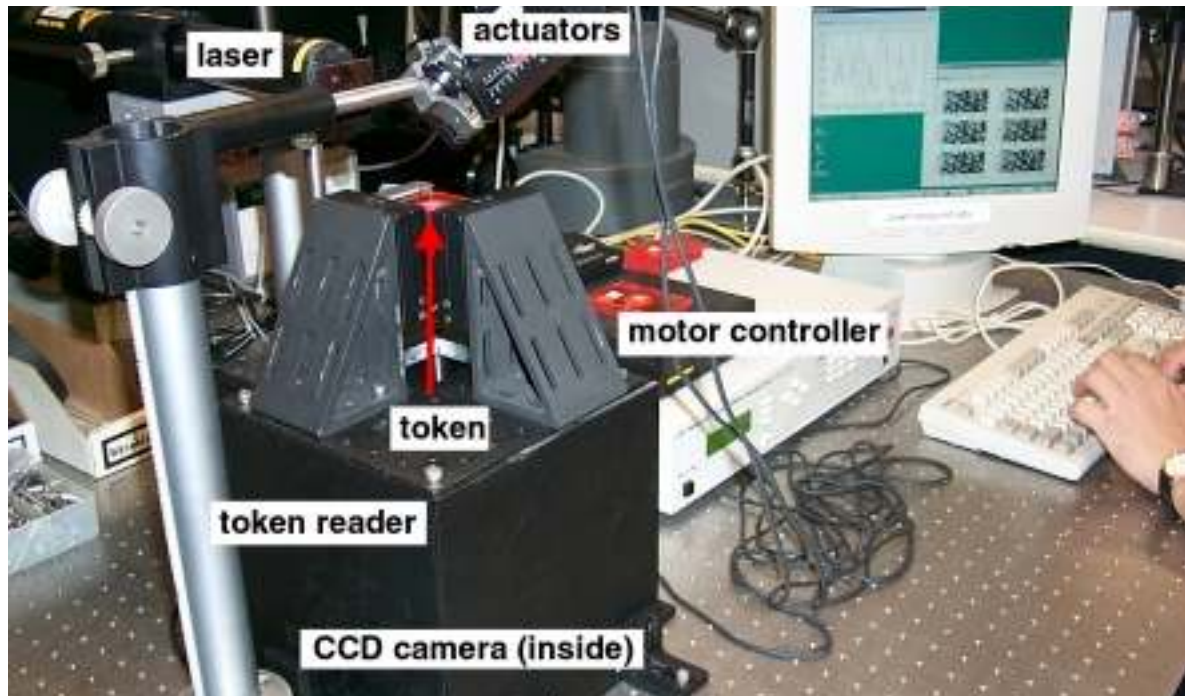


FIGURE 7.19 IMAGE OF THE FINAL SYSTEM

characterized as a reflective array of fast digital light switches that are monolithically integrated onto a silicon address chip [40][41][42]. The basic mode of operation is binary. Light shining onto a DMD is either reflected into a particular direction or it is reflected out of the optical system.

A DMD is interesting in our application for several reasons. First, it is an all-digital device which is extremely fast. Second, it is a reflective device, and thus does not affect the coherence properties of the laser light. All it does is switch the direction of the incoming laser light.

The most important reason, however, is that a DMD provides a mechanism by which the space of challenges to the microstructure is extremely large. Assume we have an $M \times N$ array of mirrors. In practice, M and N are in the region of 1000. This implies that the number of distinct bitmaps displayable on the DMD is $2^{M \times N}$, an extremely large number. Thus, by replacing our actuated mirror with a DMD, we increase the challenge space significantly.

8 Experiments and results

This chapter is devoted to describing the experiments performed in order to determine that our intuition about physical one-way functions is justified, and to demonstrate that it is indeed possible to design and implement a physical authentication system based on inhomogeneous 3D microstructures that allows the reliable and repeatable derivation of an identifier that uniquely distinguishes the structure from other similarly produced structures.

The first experiment is a proof-of-principle experiment. We are primarily interested in showing that a unique identifier can be obtained from an inhomogeneous 3D microstructure repeatably by probing it with a laser beam. The second experiment asks questions related to the statistics of the identifiers. Here we deal with a large number of speckle patterns and look at how distinguishable they are from one another. In the final experiment, we focus on determining the effect of small change in the microstructure on the identifier.

8.1 Proof-of-concept experiment

8.1.1 The setup

Our first experiment was geared towards demonstrating the proof of concept of physical authentication. We considered a system with a database in which a small number of tokens - four in this case - were initially enrolled. This was done by obtaining speckle patterns from the tokens by illuminating them from the same angle. Each of these patterns was reduced to a 2400 bit Gabor hash string via the Gabor Hash Algorithm (see section 7.5).

The goal of the experiment is to determine that the Gabor Hash is indeed a statistically significant determinant of token identity, which in turn is intimately dependent upon the structural configuration of the 3D microstructure. We achieve this goal by presenting one of the tokens to the system and determining the Fractional Hamming Distance (FHD) between its Gabor hash and those of all the tokens stored in the database. We also present a token that was not enrolled in the database to the reader.

We expect to see the following two results. The FHD between the Gabor Hash of a token and the value of the Gabor Hash of the same token stored in the database should be close to zero, while the FHD between the token and all other tokens should be closer to 0.5. Further, if we present a token that was not enrolled in the database, we expect to see a uniform FHD between the new token and the ones in the database of approximately 0.5.

8.1.2 Results

We implemented a simple graphical user interface in Matlab to demonstrate the results of this experiment. In figure 8.1, we depict a bar graph showing the fraction of bits which agree between the subsequent token and the tokens in the database. The x -axis is the token number and the y -axis is the percentage agreement, i.e., $100(1 - FHD)$. From the graph, it is clear that the new token agrees most with the stored token 1 (~95%), while the agreement with all the other stored tokens hovers around 50%. In this case, the token that was

presented to the system can be declared to be token 1. We observed similar results for all the other tokens. One point to note: in order for one of the other tokens to be mistakenly identified as token 1, approximately 45% of its bits would have to be flipped. This is equal to 1080 bits. As we will see in the next section, the probability that this occurs by chance is very low.

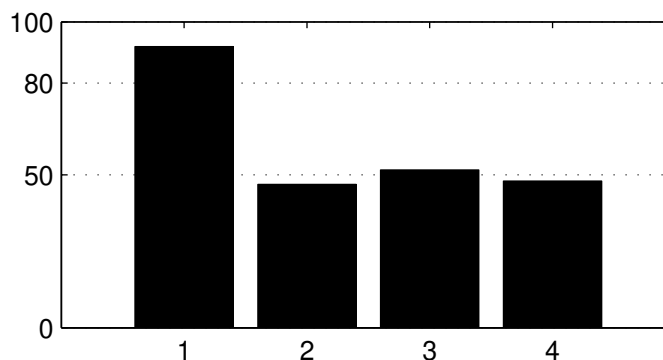


FIGURE 8.1 BAR GRAPH DEPICTING FRACTION OF BITS THAT AGREE BETWEEN NEW TOKEN AND STORED TOKENS. CLEARLY, THE NEW TOKEN MAY BE DECLARED TO BE TOKEN 1.

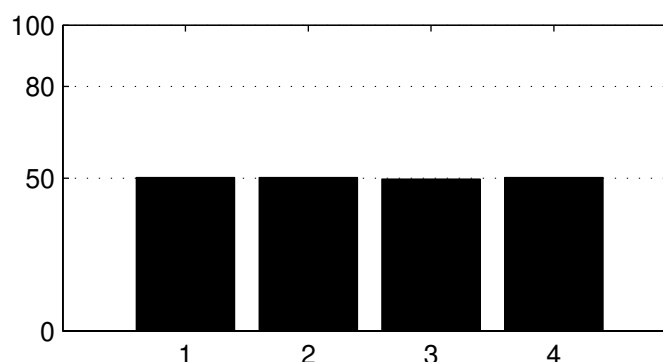


FIGURE 8.2 THIS GRAPH IS FOR A TOKEN NOT INITIALLY ENROLLED

In the other part of the experiment, a token that was not initially enrolled in the database was presented to the system and the resulting FHDs were determined. As expected, the FHDs are all very close to 0.5. This is shown in figure 8.2.

All the resulting FHDs are presented in the table below. The columns represent the tokens stored in the database, and each row represents the FHD between the stored in the database and subsequently re-presented tokens. Of particular interest are the highlighted values, which are very close to zero, as expected. Also note that the FHD between a token (D7) not enrolled in the database and all the tokens is very close to 0.5, which is what we expect.

Finally, we present the results of the two experiments above in pictorial form. In figure 8.3, the first column of images represent 1200 bits of the 2400 bit

FHD BETWEEN TOKENS IN THE DATABASE AND SUBSEQUENT TOKENS

	TOKEN 1	TOKEN 2	TOKEN 3	TOKEN 4
TOKEN 1	0.0540	0.5371	0.4854	0.5256
TOKEN 2	0.5154	0.1594	0.5152	0.5033
TOKEN 3	0.4783	0.5073	0.1698	0.4904
TOKEN 4	0.5317	0.5156	0.4894	0.1583
TOKEN D7	0.4981	0.4979	0.5029	0.4981

identifier of tokens stored in the database. The image in the second column is the corresponding set of 1200 bits from a candidate token. The third column represents the bitwise XOR between the image in the second column and one of the images in the first column. In this case, a black pixel represents no difference between the corresponding bit locations. Once again, the agreement with stored token 1 is very high, and we can clearly observe that there about an equal number of black and white pixels in the resulting images for all other stored tokens. In figure 8.4, we show the results for a token not enrolled in the database.

We conclude that it is indeed possible to identify inhomogeneous 3D microstructures by examining their speckle patterns.

8.2 Statistics of Gabor Hash strings

8.2.1 The setup

In this experiment, we are interested in the statistics of a large number of Gabor hash strings. Our primary goal will be to gather sufficient data and use it to characterize the statistics of the hash strings. More specifically, we are interested in determining a threshold for the FHD below which we can declare a presented token to be one of many stored in the database. We are also seeking to gain some intuition into the scaling properties of the system - does the present method scale to a large number of tokens?

We acquired 144 distinct speckle patterns from each of four tokens - a total of 576 speckle patterns. Each token was interrogated from 144 different angles, taking care to ensure that the incident angle changed by greater than the minimum deviation required to remove significant correlations between speckle patterns. The 2400 bit Gabor hash string was computed for each speckle pattern.

8.2.2 Statistical results

We look at the statistics of the 576 speckle patterns in two different ways. First, we plot the probability of a bit being set in a specific location. Procedurally, this is equivalent to bitwise mean of 576 hash strings. A 100 bit subset of this plot is shown in figure 8.5 below. The graph shows that the probability of a particular bit being set in the Gabor hash is very close to 0.5. In fact, the average value of the bitwise mean is 0.5002. This implies that,

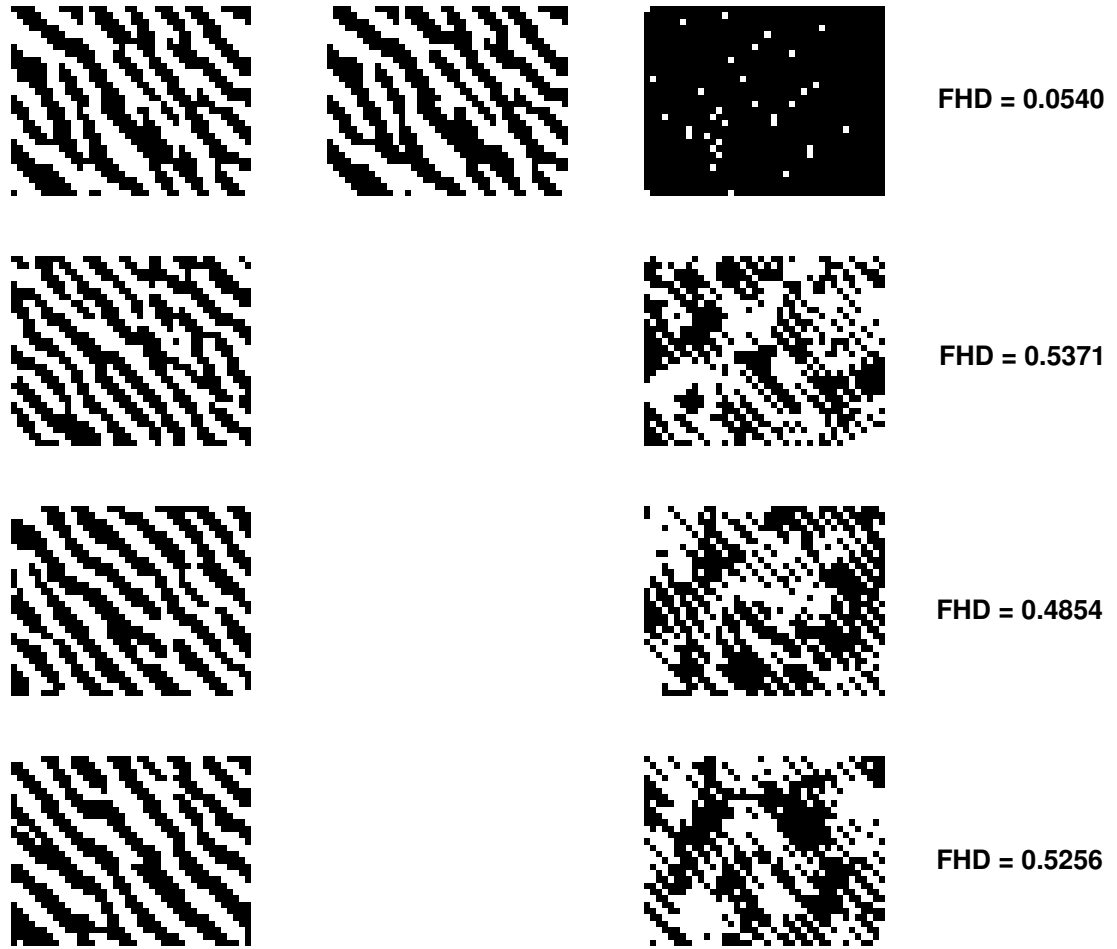


FIGURE 8.3 VISUAL DEMONSTRATION OF EXPERIMENTAL RESULTS FOR PREVIOUSLY ENROLLED TOKEN

across all bit locations, a bit is equally likely to be set. According to Shannon [61], the entropy of a code with n states is maximized if all the states are equally likely. The entropy, in bits, is

$$H = - \sum_{i=1}^n p_i \log_2 p_i \quad 8.2.1$$

where p_i is the probability of the i th state. The probabilities must satisfy

$$\sum_{i=1}^n p_i = 1 \quad 8.2.2$$

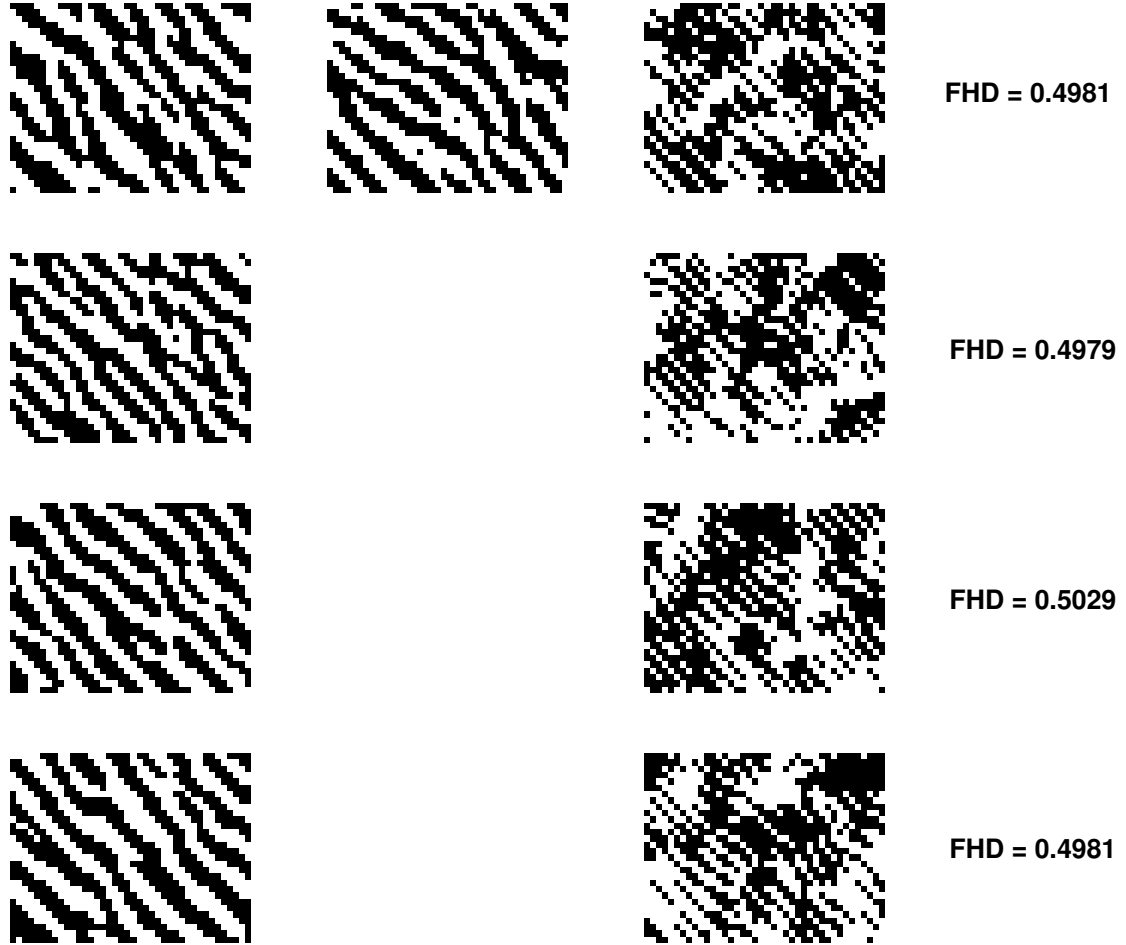


FIGURE 8.4 VISUAL DEMONSTRATION OF EXPERIMENTAL RESULTS FOR PREVIOUSLY UN-ENROLLED TOKEN

The entropy is maximized if $p_i = 1/n$. In our case $n = 2$, which implies that entropy is maximized when $p_i = 0.5$. We therefore conclude that the Gabor hash is a *bitwise maximum-entropy code*. This conclusion is also borne out by the fact that there are no systematic deviations from 0.5 in the graph. This suggests that there is no predisposition in the system for any particular code bit to assume a specific value, which is a reflection of the randomness in the 3D microstructure.

We now turn our attention to the distributions of the Fractional Hamming Distances for “like” and “unlike” speckle patterns. For this we treat the 576 acquired speckle patterns as the database. In order to plot the “like” HD distribution, we acquire 576 new speckle patterns from the same tokens interrogated with the same probe. We then determine the FHDs between the hash strings in the database and the corresponding strings from the newly acquired speckle patterns. The distribution is depicted in figure 8.6. The mean of this distribution is 0.2525, the median is 0.2456, and the variance is 0.0047. Of 576 hash strings, 327 have FHDs less than the mean FHD.

The crucial point to note is that there is no FHD value of zero. In other words,

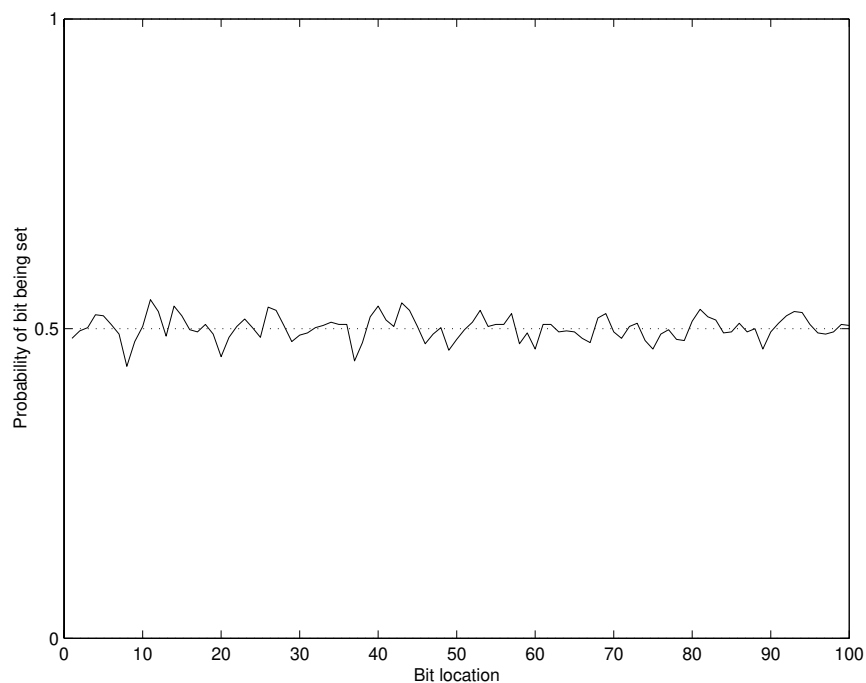
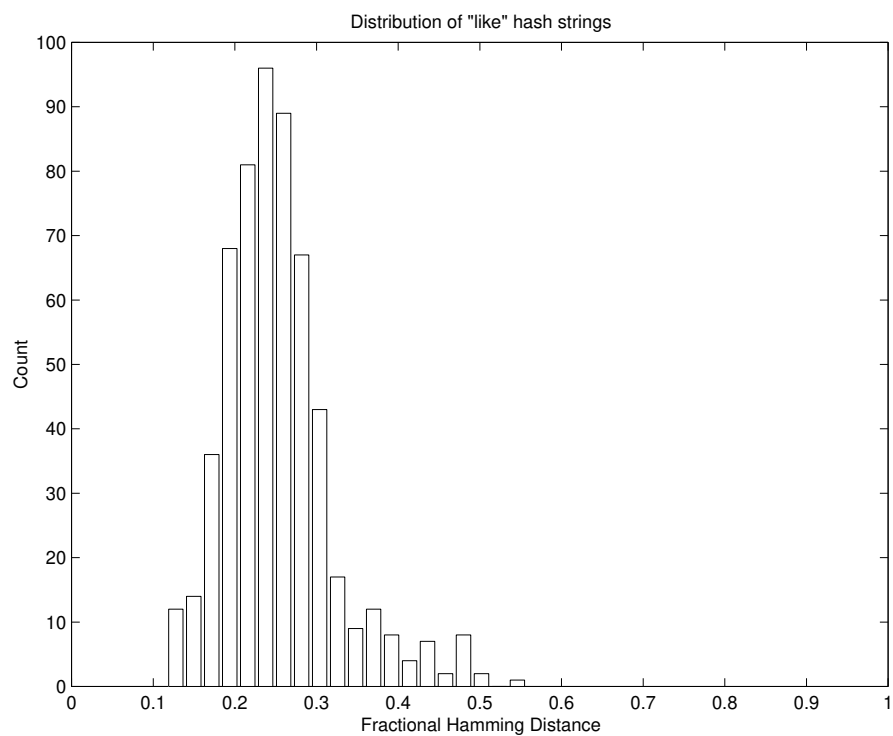


FIGURE 8.5 PROBABILITY OF A BIT BEING SET IN A SPECIFIC LOCATION

FIGURE 8.6 HISTOGRAM OF *LIKE* FRACTIONAL HAMMING DISTANCES

two speckle patterns obtained from the same structure by interrogating it from

the same angle and with the same wavelength never produce identical hash strings. This is due to several factors: sensitivity of the speckle pattern to changes in the environment, photon noise, and noise in the detector.

The distribution for unlike hash strings is easier to obtain. For all the strings in the database, we determine the FHD between each of them taken two at a time, since by definition, they are all derived from distinct speckle patterns. This gives us

$$\binom{576}{2} = 165600 \quad 8.2.3$$

distinct FHDs. The histogram of these FHDs are plotted in figure 8.7. This

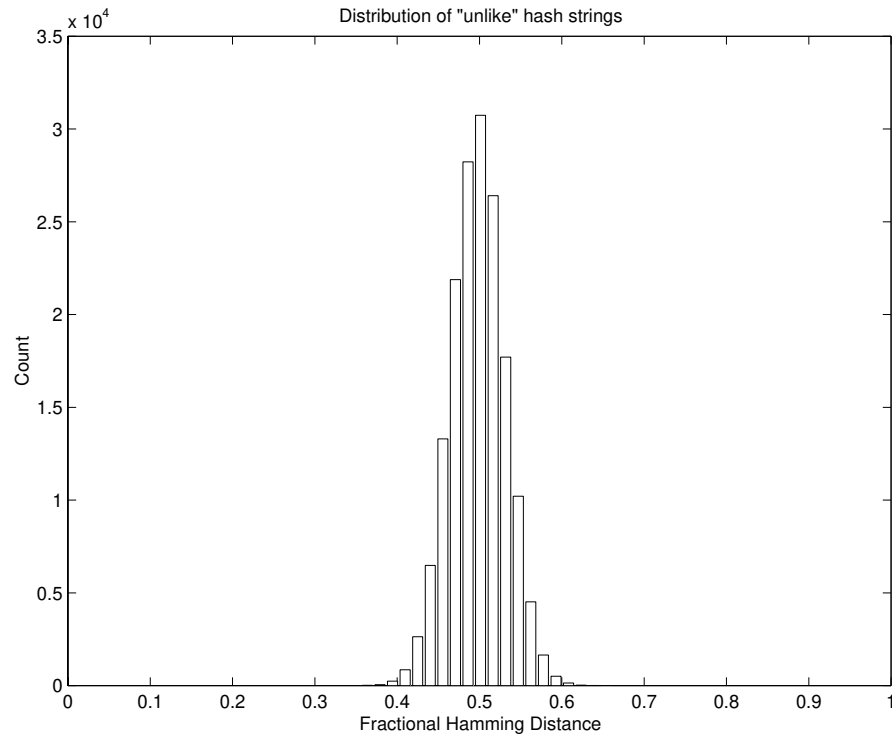


FIGURE 8.7 HISTOGRAM OF *UNLIKE* FRACTIONAL HAMMING DISTANCES

distribution is more sharply peaked. The mean of this distribution is 0.4981, the media is 0.4979, and the variance is 0.0011. This distribution is clearly symmetric about the mean.

If every bit in the 2400 bit string were independent of every other bit, then the expected distribution of the FHD between unlike strings would be a binomial distribution with $p = 0.5$ and $N = 2400$. In other words, the distribution of FHDs would look exactly like that obtained by doing 2400 coin tosses a large number of times and counting the fraction of heads in each round of 2400

tosses. However, as we saw in section 5.3.4, the speckle pattern itself has a correlated structure. Even if the speckle pattern were random, running it through a set of Gabor filters at multiple scales introduces correlations that are approximately equal to the reciprocal of the bandwidth of the filters [62]. The *effective* number of independent bits in the Gabor hash string is, however, determined by looking at the experimental mean and variance.

$$N = \frac{p(1-p)}{\sigma^2} \quad 8.2.4$$

which, for $p = 0.4981$ and $\sigma^2 = 0.0011$, is

$$N \approx 228 \text{ bits} \quad 8.2.5$$

In summary, if all the bits were independent, we would *expect* a ($N = 2400, p = 0.5$) binomial distribution, but in practice we *observe* an ($N = 228, p = 0.5$) binomial distribution. Both these distributions are plotted in figure 8.8. In figure 8.9, we superimpose the theoretical ($N = 228, p = 0.5$) on

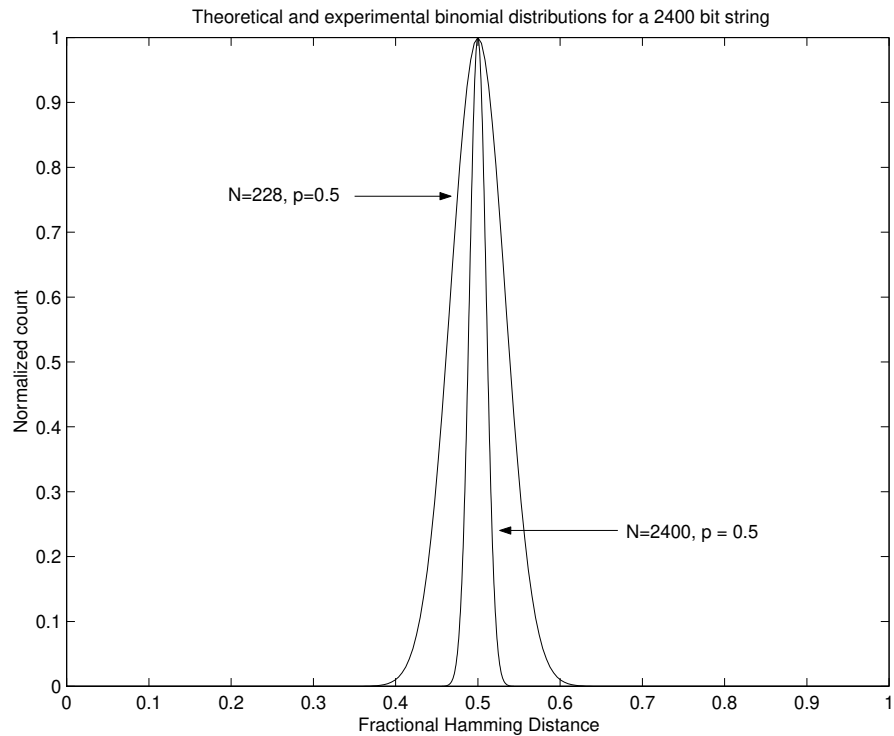


FIGURE 8.8 EXPECTED AND EMPIRICAL BINOMIAL DISTRIBUTIONS FOR A 2400 BIT STRING

the histogram from figure 8.7 to demonstrate that the observed histogram is

indeed capable of being fit by a binomial distribution with $N = 228$ and $p = 0.5$. This suggests that, given our specific implementation of the physical

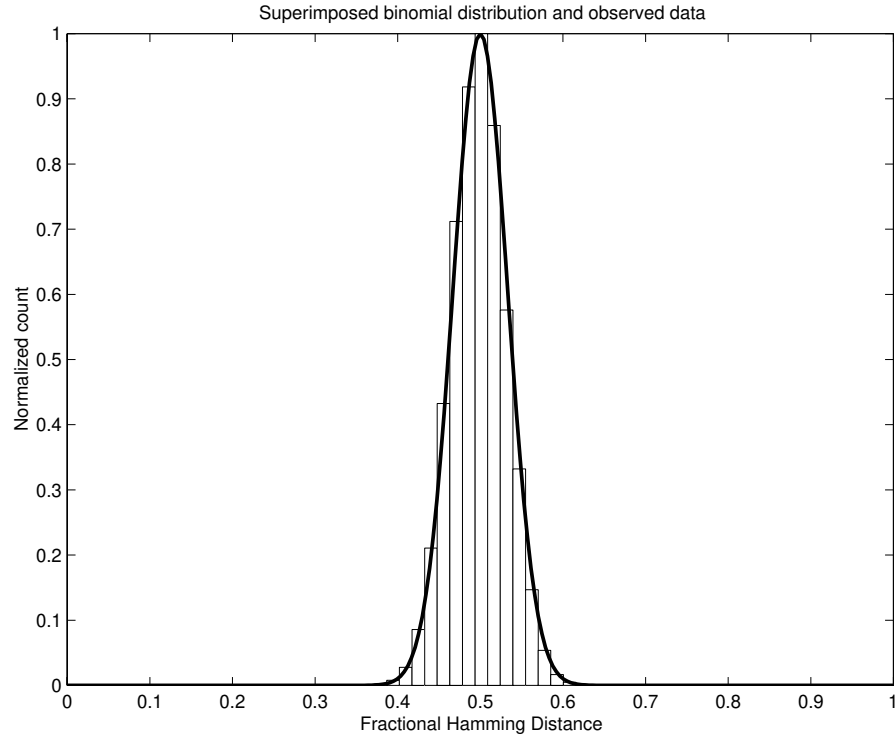


FIGURE 8.9 SUPERPOSITION OF OBSERVED *UNLIKE* DATA AND FITTED BINOMIAL DISTRIBUTION

authentication system, there appear to be 228 independent binary degrees of freedom in the 2400 bit Gabor hash string. Therefore the total number of unique identifiers we can obtain from our physical authentication is on the order of $2^{228} \approx 10^{69}$. This also means that the likelihood of the hash strings obtained from two distinct speckle patterns agreeing by chance is 10^{-69} .

We follow the same procedure outlined above to fit a binomial distribution to the *like* data and determine that the best binomial fit is given by a ($N = 41, p = 0.2525$) binomial distribution. This is shown in figure 8.10 below.

The two distributions are unambiguously separable, a fact that is easily visible when both the like and unlike histograms are appropriately normalized and plotted on the same axes (figure 8.11) along with their respective fitted binomial distributions. Looking at the data from this perspective allows us to formulate a decision criterion based on the FHD. If the FHD between a candidate Gabor hash string and a string is greater than the criterion, we declare that the two strings did not originate from the same speckle pattern. If the FHD is less than the criterion, then we can say that they did originate from the same speckle pattern. In this case, the FHD at which the two binomial distributions cross-over is the criterion and is approximately equal to 0.41. This implies that a candidate Gabor hash string would have to differ from one stored in a database in at least 984 bit positions before we declare that the two

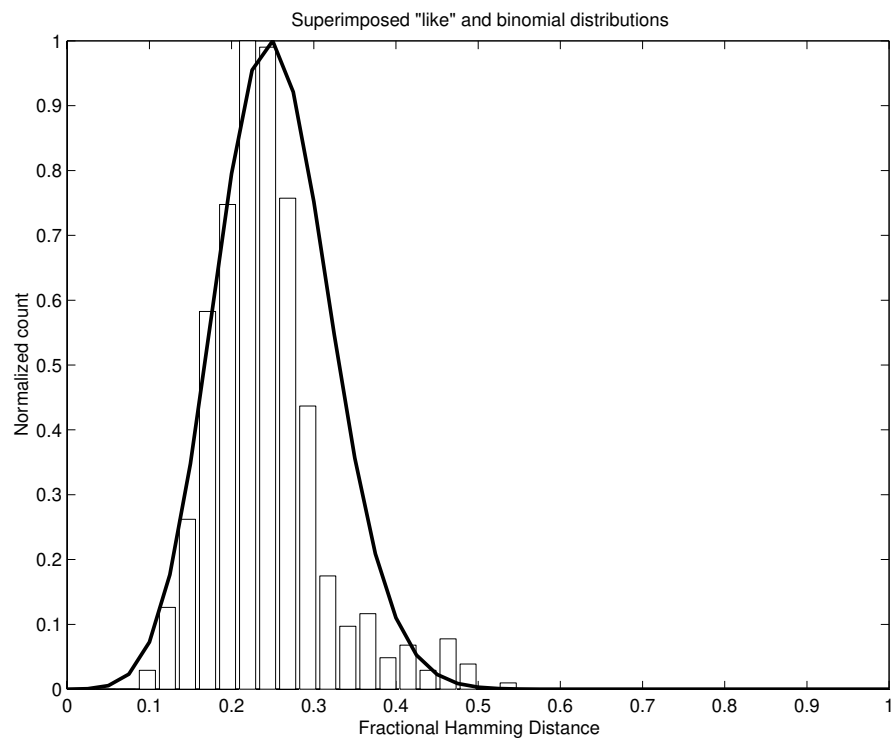


FIGURE 8.10 SUPERPOSITION OF OBSERVED *LIKE* DATA AND FITTED BINOMIAL DISTRIBUTION

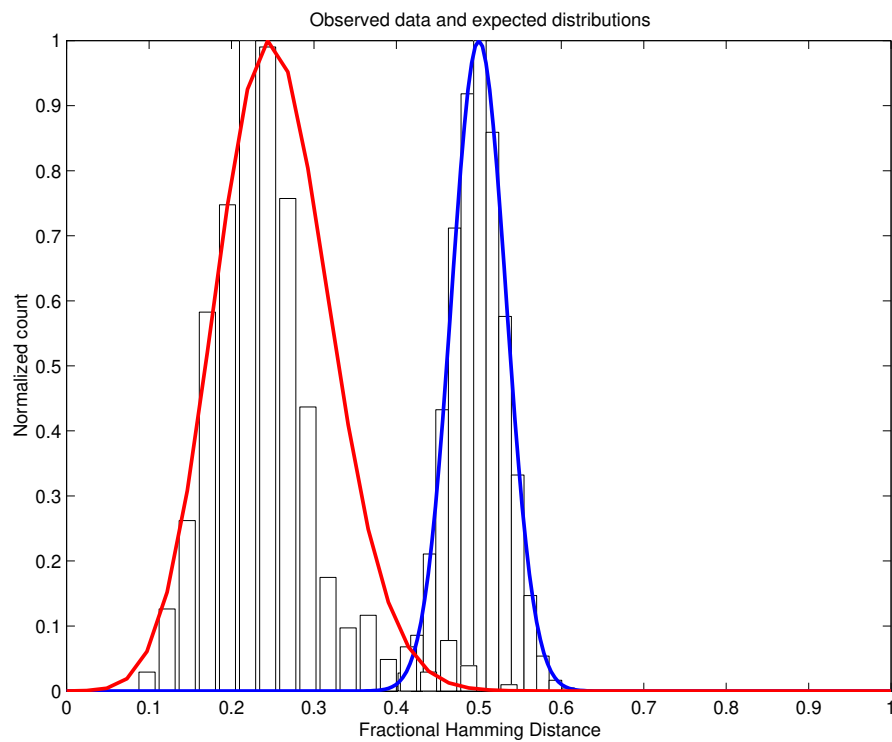


FIGURE 8.11 BOTH *LIKE* AND *UNLIKE* DISTRIBUTIONS ON THE SAME GRAPH

strings do not originate from the same speckle pattern.

8.3 Demonstration of tamper resistance

8.3.1 The setup

In this section we briefly take a look at the tamper-resistance properties of the token. We have claimed that the speckle pattern is extremely sensitive to changes in the structural configuration of the token. We have also likened this phenomenon to the avalanche effect exhibited by computational one-way hash functions (see section 6.4.4). Here we present experimental evidence to support this claim.

The experiment is quite simple. We acquired a speckle pattern and computed its Gabor Hash string. We then used a 1mm diameter drill to make a very shallow indentation in the token. The diameter of the indentation was smaller than the diameter of the laser beam used to interrogate the structure. The token was then re-interrogated and a new Gabor Hash string was obtained.

8.3.2 Results

In figure 8.12, the top two images are the constituents of the unaltered gabor hash string and the lower two images are those of the Gabor Hash string obtained from the “tampered” token.

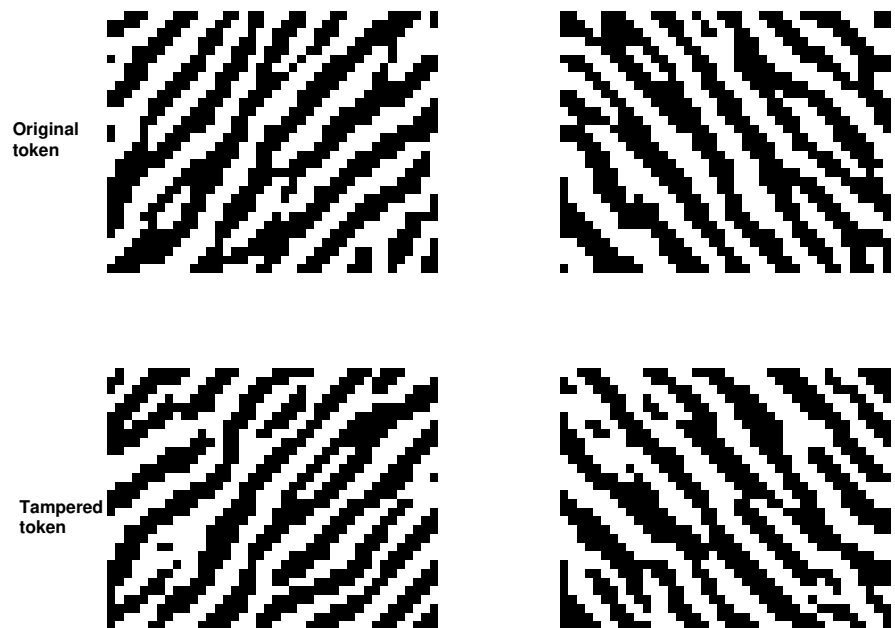


FIGURE 8.12 CONSTITUENTS OF THE GABOR HASH STRINGS OF UNALTERED (UPPER TWO IMAGES) AND TAMPERED TOKENS (LOWER ONES)

In the next figure (8.13), we show the pixel-wise XOR of corresponding images from figure 8.12. As before, a white pixel indicates that the corresponding pixels from the two images are distinct, while a black one reveals similarity. The salient point is that there are a very large number of white pixels, and a quick calculation shows that the FHD or 0.464 or,

equivalently, the number of differing bits is 1113.



FIGURE 8.13 XOR OF CORRESPONDING PIXELS FROM THE PREVIOUS FIGURE

By way of comparison, we now take a look at equivalent performance of an oft-used computational one-way hash function called *Message Digest 5* or *MD5* [63]. As input to the function, we provided three text fragments which were different by a single character. They were

One-Way
One Way
OneWay

We then determined the FHD between the outputs of MD5 when the above three fragments were inputs. We obtained an FHD of 0.5390 between the first and second fragment, 0.5078 between two and three, and 0.5938 between the first and the third fragment. This is an example of avalanche - a small change in the input causes approximately half the bits in the output to flip. Clearly, physical one-way functions behave in the same way.

8.4 Summary

This chapter was devoted to experimental procedures and results. The first experiment was a proof of concept one. The goal of the experiment was to determine that the Gabor Hash is indeed a statistically significant determinant of token identity, which in turn is intimately dependent upon the structural configuration of the 3D microstructure. This was clearly demonstrated in section 8.1.2, where we showed that: physical one-way functions can determine identity and, more importantly, non-identity clearly.

The second experiment was geared towards determining the statistics of a large number of speckle patterns. We showed first that our 2400 bit Gabor Hash string is a bitwise maximum-entropy code. Then we looked at the statistics of the FHD for like and unlike Gabor Hash strings. For unlike Gabor Hash strings, the average FHD was 0.5, and the distribution of FHDs was symmetric about the mean. We then determined the best fitting binomial distribution that explained the data. This turned out to be an ($N = 228, p = 0.5$) binomial distribution. This led us to conclude that the number of independent binary degrees of freedom in our 2400 Gabor Hash is actually 228, primarily because each bit in the Gabor Hash is not statistically independent of its

neighbors.

We also fit an ($N = 41, p = 0.2525$) binomial distribution to the histogram of the like FHDs. The fit in this case was not very good partly because of the small amount of data. Better performance can certainly be expected when more data is collected. We then showed that the two fitted binomial distributions may be used to determine a threshold FHD, which is the decision boundary between accepting a Gabor Hash string as authentic or not.

Finally, we demonstrated the analog of the avalanche effect in a physical authentication system. We did this by acquiring two speckle patterns from the same token. The only difference between the two instances was that we intentionally made a small change to the structural configuration in the second case. We then showed that the FHD between the two resulting Gabor Hash strings was close to 0.5 as expected.

9 Protocols

The last few chapters concentrated on formulating the principles of physical one-way hash functions and demonstrating their properties. Here we consider how a physical authentication system might be used in practice.

In the world of algorithmic cryptography, one-way functions and one-way hash functions are extremely useful primitives from which other, more complex, cryptographic functions are built up. For example, John Rompel [15] showed that one-way functions are necessary and sufficient for secure signatures. Halevi and Micali [64] demonstrated a practical bit commitment scheme based solely on collision-free hash functions. Public-key encryption of the Diffie-Hellman flavor [9] relies on *trapdoor* one-way functions, where inversion is efficient given the trapdoor, but intractable otherwise. Other applications of one-way (hash) functions are in coin-flipping, digital signatures, and authentication. All these cryptographic functions use one-way (hash) functions as *primitives* and are accomplished by *protocols*.

A protocol is a *series of steps, involving two or more parties, designed to accomplish a specific task* [17]. Generally, everyone involved in the protocol must know the protocol and all the steps to be followed in advance. Everyone involved in the protocol must agree to follow it. Finally, the protocol must be unambiguous and complete - all the steps must be well defined and there must be a specified action for every possible situation. In this chapter, we will present two simple protocols employing physical one-way functions as primitives.

First, we will present a protocol where a small number of all possible speckle patterns obtainable from a given 3D microstructure are used to authenticate a transaction, and *never reused*. This is reminiscent of the one-time pads used extensively before the invention of public-key cryptography. We then present an elementary bit-commitment protocol. The objective of this chapter is not to provide protocols which have practical utility. Rather, given the existence of physical one-way hash functions and their performance, we show how such protocols might be constructed. Our goal is to demonstrate that we can think about physical one-way hash functions in a cryptographic framework.

9.1 A bit of history

Before we press ahead with our protocols, we take a small excursion into cryptographic history. In 1917, Gilbert Vernam was given the task of inventing an encryption system that could not be broken. His efforts yielded the one-time pad cryptosystem.

The security of the secret-key one-time pad cryptosystem rests on three principles. First, adding a random number to a known one produces a random number. Second, a secret key is never reused. Third, the message and the key must be the same length.

For ease of explanation, we use binary notation in the ensuing discussion. The one-time pad cryptosystem is implemented as follows. Two identical copies

of a large set of random numbers are created. They are then distributed to the sender and the receiver. When the sender wants to send, say, a 100 bit message, he XORs the first 100 bits of the pad with the message to produce a random 100 bit string. This randomized message is then sent to the receiver, who XORs the received message with the identical 100 bits of *his* copy of the pad. The result is the original message. A fascinating account of an innovative method of making and using one-time pads provided in [65].

A distinguishing feature of this cryptosystem is that exactly two copies of the pad exist, one with the sender and one with the receiver. An adversary without the pad would have to do a dictionary search of all 2^{100} possible keys in order to decrypt an intercepted 100 bit message. Each additional bit in the message doubles the number of keys to be searched.

9.2 One-time pad protocol

9.2.1 Motivation

We motivate our one-time pad (OTP) protocol with a simple application scenario. The application involves using a regular credit card to purchase goods or services. The basic protocol for the transaction is shown in figure 9.1. The cardholder presents her card to the terminal, which reads the data on the magnetic stripe, transmits it via a network to the server in a different location which either authorizes or declines the transaction and transmits the binary decision back to the terminal. The problem with this system is that a magnetic stripe card can easily be cloned. Cloned cards can then be used to buy goods and services which are charged to the original owner of the card. A common method of cloning is dubbed *skimming a card* whereby a palm-sized device is used to gather all the data encoded in the magnetic strip without the knowledge of the legitimate user of the card.

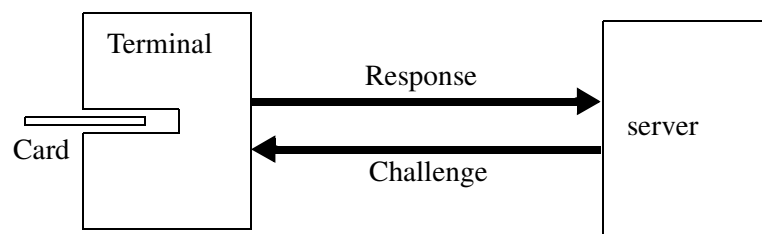


FIGURE 9.1 DATA PIPELINE FOR MODEL TRANSACTION SYSTEM

We look at two different scenarios. In the first case, the terminal, the network, and the server are completely trusted. This might correspond to the case where the terminal is an Automated Teller Machine (ATM), the network is a private network owned by the card-issuing bank, and the server is in a secure location. In the second case, the terminal might be situated at an insecure location, such as a department store, and hence is regarded as an untrusted terminal. The network between the terminal and the server is also treated as insecure. The server, however, is still in the same secure location and is trusted. We assert that *in either of the two situations, a cloned card (CC) is*

indistinguishable from an authentic card (AC).

9.2.2 Augmenting the card and terminal

We are interested in determining if the addition of a physical one-way function system to the card will make it much more difficult to clone the cards. Consider augmenting the credit card by adding a 3D inhomogeneous microstructure to it. This would cause the card to look very similar to the tokens we have been using for our experiments (see, for example, figure 7.1).

We also stipulate that the terminal, in addition to its magnetic stripe reader, is augmented with a modified physical one-way function reader, as described in section 7.7. With this modified reader in place, we are able to illuminate the 3D microstructure with a coherent wavefront which has been modulated by a bitmap. Everything else about the terminal is kept constant.

We also assume that the microstructure is at least as large as the image of the DMD and the DMD has $M \times N$ mirrors, where M and N are in the region of 1000. The first assumption guarantees that each pixel of the DMD is imaged to a separate spot on the microstructure. The utility of this property will become clear very soon.

9.2.3 Nonlinearity in the microstructure

We also assume that the 3D inhomogeneous microstructure is weakly nonlinear, as described in section 5.4. The reason for this will become clear in due course.

9.2.4 Assertions

We now assert, based on results from previous chapters, that:

- It is impossible to clone a microstructure.
- It is impractical or infeasible to simulate the passage of coherent light through the microstructure, i.e., given the complete information about the scatterers in the structure, there is no practical or efficient algorithm, which, when provided with the state of the probe, will output the speckle pattern.
- A small change in the configuration of the 3D microstructure produces a very large change in the Gabor Hash string.
- The probability that two 3D microstructures produce an identical Gabor Hash is very small and decreases exponentially with the size of the microstructure.
- Each distinct state of the probe produces a speckle pattern whose Gabor Hash string has a Fractional Hamming Distance of approximately 0.5 from that of any other bitmap. Each probe state may be regarded as a **challenge**, and the resulting speckle pattern may be regarded as the **response**.

9.2.5 Notation

For the remainder of this chapter, we use the following notation, developed in section 6.1.

$\Lambda = \{\lambda\}$ is the set of all possible wavelengths of coherent radiation used to probe the structure.

$\Theta = \{\theta\}$ is the set of 3D angles of the radiation incident on the structure.

$B = \{I(p, q)\}$ is the set of complex bitmaps used to spatially modulate the wavefront before it impinges on the structure. Each I is a bitmap image of size $M \times N$. There are $2^{M \times N}$ possible bitmaps.

Each challenge is an element of the set $P = \{(I, \lambda, \theta) | I \in B, \lambda \in \Lambda, \theta \in \Theta\}$.

The set of responses (speckle patterns) is denoted by $S = \{s_i\}$ where $i = 1, 2, 3, \dots, |P|$. In this case, $|P| = |B| \times |\Lambda| \times |\Theta|$.

For the remainder of this discussion, we will assume a single wavelength and angle of illumination and focus solely on the multiplicity of bitmaps which modulate the coherent radiation. That is, we assume $|\Lambda| = |\Theta| = 1$. Therefore, $|P| = |B| = 2^{M \times N}$.

We recall the assumption of weak nonlinearity of the 3D microstructure here. If the microstructure were linear, then the responses to each of the $2^{M \times N}$ possible challenges may be predicted by knowing the response to each of the $M \times N$ probes obtained by illuminating the structure one DMD pixel at a time. These basis responses can be combined coherently to predict responses to illumination with combinations of pixels. A weak distributed nonlinearity in the 3D structure makes this prediction problem much harder by requiring knowledge of *all* $2^{M \times N}$ possible responses, which is an exponential increase in the number of stored responses.

9.2.6 The protocol for trusted terminals

The set of challenges (and, therefore, possible speckle patterns) is P .

When the bank issues the card to a certain user, in addition to the magnetic stripe data, it also stores the challenges and resulting Gabor Hash strings for a subset of all possible challenges denoted by $P_s \subset P$. The challenges as well as their number are chosen at random for each card. Therefore, each bitmap is uniformly chosen from $2^{M \times N}$ possible bitmaps, and the number of bitmaps is picked from the range $[1, 2^{M \times N}]$. In practice, the number of bitmaps is kept small.

Here's the protocol:

- The user presents her card to the terminal.
- The terminal verifies the identity of the card from the magnetic stripe and transmits it to the server.

- The server requests the terminal to generate P_q Gabor Hashes for a random subset of the stored challenges P_s for the card.
- Upon receiving the hashes, the server computes the FHD between them and the corresponding stored hashes, and makes a binary decision about the authenticity of the card.

With each “*accept*” decision, the server is more confident that the card presented to the terminal is authentic, because it is impossible to clone the physical microstructure, and it is infeasible to simulate the response of a microstructure to a specific pattern of radiation. Because the terminal is trusted, the cardholder is unable to discern what the challenges were or what their sequence was. Therefore, the challenges may be reused for subsequent transactions.

9.2.7 The one-time pad protocol for untrusted terminals

We now assume that the transactions in the system are primarily carried out at untrusted terminals. In this case, an adversary could observe the challenges and responses and create a look-up table that maps a challenge to a response. When the server issues a challenge, the adversary simply plays back a stored response. This is an example of a *replay attack* [17]. Replay attacks are possible because old responses still have value. We now present a protocol where replay attacks are not possible.

As before, when the card is issued, the bank acquires the Gabor Hash strings resulting from $P_s \subset P$ challenges to the structure. Here $|P_s|$ is a large number, substantially larger than in the previous section.

The protocol proceeds as follows:

- The cardholder presents the card to a potentially untrusted terminal.
- The terminal verifies the identity of the card from the magnetic stripe and transmits it to the server.
- The server challenges the card with some randomly chosen subset of stored challenges.
- Upon receiving the hashes, the server computes the FHD between them and the corresponding stored hashes, and makes a binary decision about the authenticity of the card.

Up to this point, the protocol is identical to the one described in the previous section. However, here is where we deviate from the old protocol.

- When the next transaction is initiated, the server queries the card with a **disjoint** subset of stored queries. No previous challenges are reused.

- When the number of stored challenges diminishes, the server requires that the cardholder visit a trusted terminal where it **re-acquires** the speckle patterns corresponding to a new set of challenges.

If the protocol is carried out as specified, each challenge, and the corresponding response, is used *exactly once* during the lifetime of the card. An adversary with a cloned card does not know which set of challenges will be issued during a particular transaction, and given the assertions above, will not be able to either simulate the responses or replay old responses, since each response is used only once. Clearly, there is a tradeoff between the mean time between “refills” and the amount of data required to be stored per card in the bank’s database. The more data stored, the less frequent “refills” have to be.

Classical one-time pads work because there exist exactly two copies of the pad, one with the sender and one with the receiver. In our case, the 3D microstructure is analogous to the pad, and interrogating it with a specific probe is tantamount to using one sheet from the pad. However, the difference between classical one-time pads and our protocol is that we can, by design, have only a single pad. In essence, we simulate the existence of the other pad with memory. In our protocol, the memory is in a secure location, and the unclonable “pad” resides with the cardholder.

9.3 Bit commitment

9.3.1 Background

We now turn our attention to another primitive which plays a fundamental role in many other cryptographic protocols: *bit commitment*. This is a primitive which can implicitly be traced back to very early public-key papers. It is a basic component of *coin-tossing protocols* where two parties, historically Alice and Bob, who do not trust each other want to toss a coin over a telephone line. In *zero-knowledge proofs*, Alice wants to prove the validity of an assertion to Bob without revealing anything else to him other than the fact that the assertion is true.

The basic idea of bit commitment is simple. Alice wants to follow a procedure by which she can commit a bit b to Bob. By this we mean that she has a bit in mind that she wants to commit in such a way that Bob has no information about what bit it is. However, once she commits the bit, she must not have the ability to change her mind. There are several protocols designed to accomplish this objective, including some quantum bit commitment protocols. Here we look at one such protocol which involves the use of algorithmic one-way functions.

We note that a bit commitment scheme has two important characteristics: *concealing* and *binding*. The it is concealed from the receiver, and it is binding on the sender.

Assume that Alice has a secret s that she wants to commit. She simply sends Bob $c = f(s)$, where c is the output of a one-way function f . To de-commit, she sends Bob the original secret s and he verifies that Alice was not lying during the commit phase. This simple protocol has some problems, which we

analyze here. First, if there are a limited number of secrets, then Bob could simply apply f to all secrets and compare the output to c . One of the outputs matches c at which point Bob has divined the secret without going through the de-commit phase. Obviously, if the number of secrets is extremely large, Bob will spend a long time trying all of them. The other problem is that if the one-way function is not one-to-one, then Alice has room to cheat. She sends $c = f(s_1)$ but claims her secret is s_2 because $f(s_1) = f(s_2)$. What is needed here is a collision-resistant hash function.

9.3.2 The bit-commitment protocol

We now discuss a rudimentary bit-commitment protocol which involves physical one-way hash functions.

As before, Alice wants to commit a set of bits to Bob in such a way that she cannot change her mind later and that Bob cannot divine the set of bits. We assume that Alice and Bob are in the same location, and each possesses an identical modified reader as described in section 7.7. We also assume that there is a large supply of tokens which each encapsulate inhomogeneous 3D microstructures available and that both Alice and Bob are honest, i.e. they follow the protocol as stated and make no attempt to cheat.

The protocol proceeds as follows:

Commit phase:

- Bob draws a token at random from the large supply of available tokens.
- Alice presents this token to her reader.
- Alice then interrogates the token with a pattern of illumination that is a function of her set of bits. In other words, her bits are mapped into spatial coordinates of the DMD pixels. This mapping is part of the protocol and is available to both participants.
- She obtains a Gabor Hash string the resulting speckle pattern.
- She hands over the token as well as the resulting Gabor Hash to Bob.

De-Commit phase:

- Alice declares her bits to Bob.
- Bob maps the bits into spatial coordinates as prescribed in the protocol and interrogates the token in his reader.
- Bob derives the Gabor Hash string for his speckle pattern.
- Bob computes the FHD between his Gabor Hash and the one Alice gave him
- If the FHD is below the threshold prescribed by the protocol, he knows Alice was telling him the truth.

Does this protocol fulfill the two requirements of concealing and binding? If we indeed have collision-resistant hash functions as primitives, then the probability that Bob will invert the committed Gabor Hash to derive the bits Alice had in mind is very low. In respect of binding - if each (of $2^{M \times N}$) of the bitmaps produces a distinct pattern, the probability that Alice will be able to find two illumination bitmaps that produce an identical Gabor Hash is very small and falls exponentially with the size of the 3D microstructure.

9.4 Summary

We conclude that physical one-way hash functions have the potential to be primitives in rudimentary cryptographic protocols. We point out once again that these simple protocols are not intended to have practical utility. Rather, we hope these protocols point the way for designing protocols which do indeed have such utility.

10 Scaling, attacks, and fabrication complexity

We address three different issues in this chapter. The first question we tackle is: how does a physical authentication system scale? Scaling might occur in two ways: the number of tokens could increase or the size of the physical system used in each of the tokens might increase. In either case, the scaling performance is determined by the specific choice of the physical system and is very system dependent. In the first section of this chapter, we discuss scaling issues with respect our implementation of a physical authentication system.

Sections 10.2 and 10.3 present variety of possible attacks against a physical authentication system. The types of attacks we discuss are the *brute-force* attack, the *birthday attack*, and the *replay attack*. As before, we will only discuss these attacks in the context of our specific implementation of a physical authentication system. Specifically, we assume the model for a transaction system to be the same as in the previous chapter, and retain all the assumptions about the physical authentication system that we made in section 9.2.

We then take a brief detour to discuss currently available methods of microfabrication with a view to gaining some insight into the cost, complexity, and limitations of these methods as applied to the problem of constructing inhomogeneous 3D microstructures.

In section 10.5, we broaden our view to present the notion of *fabrication complexity* of any physical system. In this section we take a closer look at the question: how hard is it to clone a physical system? Fabrication complexity is a metric of the resources required it is to clone a physical system to some specified accuracy.

Finally, we discuss how a 3D microstructure might be cloned using a parallel fabrication attack, and relate the parameters of a fabrication attack to the fabrication complexity.

10.1 Scaling issues

10.1.1 Scaling the size of the physical structure

Here we discuss the implications of scaling the physical structure. Consider that we have in our possession an inhomogeneous 3D microstructure of dimensions $L \times L \times L$ and contains N spherical scatterers of diameter d uniformly distributed throughout its volume $V = L^3$. We assume that the spheres are large enough so as to be in the geometrical optics regime, i.e., the wavelength of the probe $\lambda \ll d$.

We now assume that the linear dimensions of the structure are doubled, i.e., they are now $2L \times 2L \times 2L$, where $2L < L_{abs}$, the absorption length. Let us see how this increase in scale affects the system parameters and performance.

- The new volume is $8V$ and in order to keep the density constant we have to add $7N$ more spheres to the volume.

- The mean free path remains unchanged.
- Each photon now undergoes, on average, four times as many scattering events as it travels through the structure, i.e., $(2L/l)^2 = 4(L/l)^2$
- The fraction of scatterers in a given path through the structure is now half its previous value, i.e., $l/(2L) = 0.5(l/L)$.
- The angle at which the C_1 correlation is effectively zero is now half its previous value, i.e., $\lambda/(4\pi L) = 0.5(\lambda/2\pi L)$. The engineering implication of this scaling is that the performance of the angular positioning system must go up by the same factor.
- The average transmitted intensity decreases by a factor of 2.
- Speckle sensitivity to the motion of a single scatterer increases by a factor of 2
- The total number of structures distinguishable by a probe increases exponentially with the size of the structure.

Clearly, then, increasing the size of the structure has both advantages and disadvantages. The advantages are increased sensitivity to tampering and increased effort to simulate the output. The disadvantage is that the mechanical performance (token registration and laser positioning) must be more accurate, which translates to increased cost. However, the cost of the token itself remains almost constant with increase in size.

10.1.2 Scaling the number of tokens

Now we are concerned with gaining some insight into the performance of the system as the number of tokens is increased. Specifically, we would like to know if there is any performance *degradation* as the number of tokens increases.

Consider the histograms plotted in figure 10.1 (which are recalled from figure 8.11). In an ideal world, the mean value of the *like* distribution (on the left in the figure) would be 0 and the that of the *unlike* distribution would be exactly 0.5. Both the distributions would be δ -functions. In practice, however, the mean value of the like distribution is closer to 0.25 and that of the unlike one is 0.5. Both these distributions have finite variances. There are two primary reasons why this is so. We list them in decreasing order of influence.

- The first reason for the large value of the mean of the like distribution is noise in the system. This noise has many origins, but the biggest source of noise is mechanical misregistration. We recall that speckle pattern is very sensitive to angle of incident radiation. Despite our best efforts at precision engineering, there is still some misregistration between successive instances of presenting the token to the system, as we have previously seen in figure 7.12. Further, the phase of the Gabor Transform scrolls as the token undergoes small rotations in the horizontal plane.

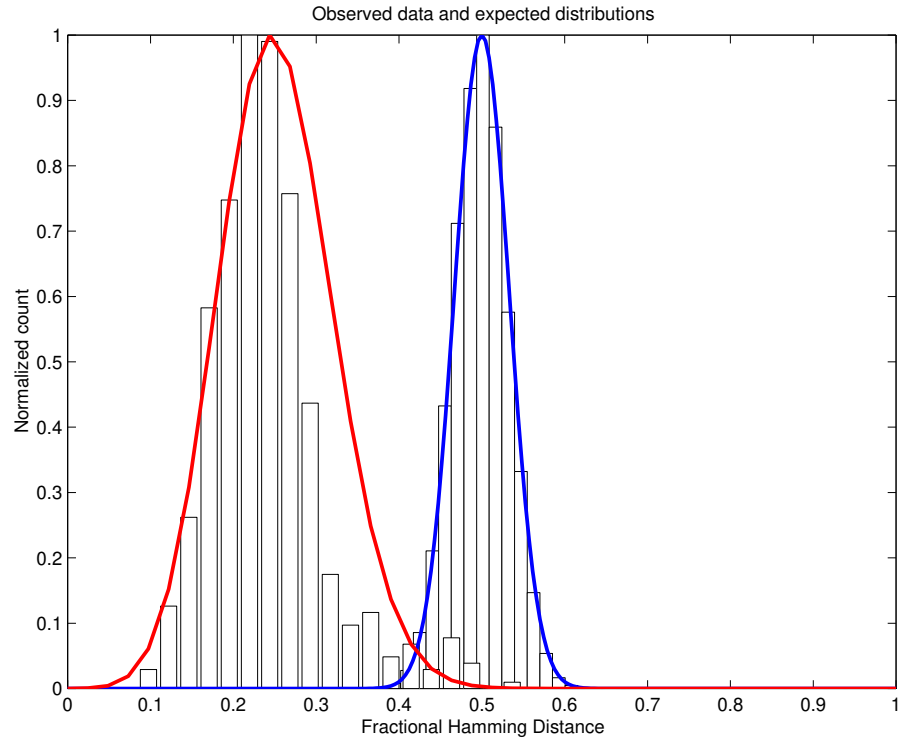


FIGURE 10.1 HISTOGRAMS FROM LIKE AND UNLIKE TOKENS PLOTTED ON THE SAME GRAPH

Because we threshold about zero, several pixels whose phase is close to zero to begin with flip in value because of small rotations. Thus, misregistration is a major reason why bits in the Gabor hash string tend to flip. There is no fundamental reason why the tokens have to be misregistered. Better engineering could easily eliminate (or at least diminish substantially) this source of noise.

- The second reason why bits tend to flip is that the speckle pattern is very sensitive to small changes in the structure. Minute scratches can cause substantial changes in the pattern. We address this issue by selecting features that are on the same scale as the lobes of speckle and ignoring changes at other scales. This is one of the big advantages of a multiscale thresholding scheme. The problem could be further diminished by encapsulating the microstructure with clear, scratch- and scuff-resistant film.
- Finally, photon noise is also a reason why bits flip. Although, we have used a complex-valued transform to negate the effects of changes in ambient levels of illumination, multiple scattering routinely affect a few pixels in the speckle pattern. This issue could be dealt with by enclosing the complete optical train in a light-tight enclosure. In our implementation, for ease of experimentation, we enclosed only the detector in a light tight enclosure. An optical chopper could also take care of this problem.

If the token-reader were re-engineered with the suggestions above, a substantial reduction in the mean value of the like distribution can be expected. However, because of noise, it is unlikely that the mean will be zero.

Even without these improvements, we have seen (section 8.2) that an average of 984 bits (from a total length of 2400 bits) have to be flipped before a false reject occurs. Also, the probability that two unrelated speckle patterns are called related (false accept) is in the region of 10^{-69} . Let us also recall that the number of possible tokens for this example is in the region of 10^{60} . The number of tokens is almost the same as the total number of Gabor hash strings, and the average Fractional Hamming Distance between unrelated Gabor hashes is 0.5. This suggests that robustness of the system will not diminish significantly as the number of tokens is increased, as long as the total number of tokens is smaller than the number of possible tokens.

There is also a more principled, engineering-independent way to "move" the two distributions in figure 10.1 apart by using a *privacy-amplification* protocol. This protocol essentially, over several rounds, converts the two distributions into δ -functions located at 0 and 0.5 respectively. The interested reader is referred to [84] and [85] for more details.

10.2 Brute-force and birthday attacks

We now consider two common attacks on physical authentication systems. These two attacks, for algorithmic one-way functions, are described in section 2.2.3. The brute force attack is completely independent of the physical system used. Assume we have a POWF system whose output is n bits long and that a specific token is required in order to authenticate a particular transaction. In the absence of the token, an adversary could compromise the system by simply trying all 2^n possible strings. Following the analysis in section 2.2.3, the number of attempts k to break the system with unity probability is given by $k = 10^{n/3.3}$. In the specific case of our system, n is effectively 228 (see equation 8.2.5). This gives us

$$k = 10^{228/3.3} \approx 10^{69} \quad 10.2.1$$

This is the number of strings an adversary would have to try to compromise a POWF-based authentication system as implemented in chapter 7.

If the adversary were to try the birthday attack instead, and try to find two 3D structures which hash down to the same value with probability at least 0.5, then she would have to try on the order of 10^{35} tokens. This number increases to $k \approx 10^{45}$ in order to increase the probability of success to 0.99.

These two attacks are not unique to POWF-based systems. They apply to all authentication systems which is why they are not of too much interest to us here.

10.3 Replay attacks

The replay attack is perpetrated when an adversary stores up old Gabor Hashes and plays them back as needed. There are many ways in which a replay attack could be made in a POWF authentication system. We discuss two of them here.

- *Store up all possible challenges and responses:* Here the adversary has access to a valid token and a token reader and wants to build a look-up table mapping each possible challenge to a response. Let us consider what this entails. Since we are using a 1000×1000 DMD, there are 10^6 possible bitmaps with a single pixel turned on. If the structure is linear, then the response of the structure to *any* bitmap illumination could be calculated by *coherently* superposing speckle patterns from these 10^6 basis bitmaps. This implies that an adversary would have to record the complex amplitudes of the speckle patterns resulting from each of these basis illuminations. While this number may not appear very large, it is still significant when we consider that recording complex amplitudes essentially entails making holograms. We note that recording the complex amplitude of any wavefront is a non-trivial problem. One possible approach would be to build an automated hologram recording machine which could produce and store all possible speckle patterns corresponding to a valid token digitally. Building such a machine entails a significant research and financial commitment.

We note in passing that a weakly nonlinear structure would dramatically increase the number of stored holograms from 10^6 to on the order of 2^{10^6} . This is clear from the discussion in section 5.4. At the conclusion of that section we saw that the nonlinear speckle pattern is exponentially sensitive to the structural configuration of the token and the state of the probe. Inducing a weak nonlinearity in the structure would render the replay attack much harder by requiring the adversary to store an impractical number of holograms.

- *Simulate the responses computationally:* One approach which avoids recording a large number of holograms would be to determine (at least in the linear case) the scattering matrix of the 3D structure. Once the scattering matrix is at hand all responses to challenges may be computationally simulated.

Determining the scattering matrix is a non-trivial task because of the very large number of scattering matrix elements (on the order of $(1/\lambda^2)^2 \sim 10^{20}$) involved. We recall from section 5.3.7 that the squared magnitude of each complex scattering matrix element s_{ab} determines the probability that radiation incident on the 3D microstructure at angle a exits at angle b . Because we are using coherent radiation, it is essential to determine the complex value of each scattering matrix element. It is possible, in principle, to determine each $T_{ab} = |s_{ab}|^2$ (but *not* s_{ab}) by using a *polar nephelometer*, which sends light in to the slab at a specific angle and detects the amount of light scattered in all directions. The reader is referred to [59] and [60] for more details of the construction and use of nephelometers.

Even if all we wanted to discover were the T_{ab} , and not the s_{ab} , we would have to make $O(N^2) = 10^{20}$ queries to the 3D structure with the nephelometer. Let us assume for a moment that we are able to determine $O(10^9)$ scattering matrix elements per second. Then it would take on the order of 10^{11} seconds which is about 10^5 years. We also note that storing all the elements of the scattering matrix would require disk space well beyond the capabilities of modern storage devices.

We conclude that it is impractical to attempt to discover the scattering matrix for such the inhomogeneous microstructure we are using. If, however, the scattering matrix were made available, determining a single response would require $O(N) \sim 10^{10}$ operations, which *is* practical, given current computing speeds.

In summary, an adversary could either store up all the responses to all potential challenges, which requires determining the complex exiting wavefront for each challenge. Alternatively, she could try and determine the scattering matrix which completely determines the response of the structure to any challenge. We have just demonstrated that this is an impractical task.

10.4 Fabrication methods

The security of a POWF-based authentication system depends on the difficulty of fabrication of arbitrary three-dimensional microstructures with some predetermined accuracy. The demand on accuracy comes from the probe, whose wavelength determines the smallest length scales. In this section, we look at available methods of microfabrication in terms of their complexity and cost. Our goal will be gain some intuition into the resources required to clone an inhomogeneous 3D microstructure. Fabrication techniques may be divided into two classes: top-down and bottom-up. Top-down approaches start by defining the required structure and proceed by engineering a method to fabricate it. Bottom-up approaches rely on chemical and statistical forces to create the structure. We will discuss only the former class of techniques here.

10.4.1 Photolithography

The standard fabrication method used to make essentially all microelectronic devices is photolithography where 2D patterns are defined using masks and then transferred to a substrate. Thus, there are two distinct steps: pattern definition and pattern transfer. 3D structures are created by stacking 2D layers.

The substrate is first coated with a polymeric photoresist whose chains break down on exposure to ultra-violet light. Then, the photoresist is exposed to UV light through a mask and some reduction optics. The exposed photoresist is then washed away using a developer to leave a pattern on the substrate. The entire substrate is then coated with the material to be deposited, usually an insulating or metallic layer, and the resist is then lifted by dissolving in a solvent. This leaves the deposited material in the regions where holes existed in the photoresist layer. In general, photolithography can be done with metals

or insulators. There are stringent requirements on alignment because the feature size is on the order of 0.1 micron.

The primary barrier, however, to the use of microlithography in fabricating arbitrary 3D structures is cost. In the table below we present the current cost [81][82] of each of the key steps in a microlithography fabrication process. It is clear that the most expensive steps in the process are exposing the substrate and photoresist processing. We also note that these costs are provided for approximately 30 layers. In order to fabricate a 1 mm thick 3D structure at a longitudinal resolution of 0.05 microns, we would need to expose 20000 layers. Finally, we note that this process is geared to the production of a large number of identical structures, as opposed to a distinct one each time.

Process	Cost (millions of dollars)
Exposure tools	816
Automated photoresist processing	288
Etching	280
Cleaning and stripping	30
Automation	30
Infrastructure	992

Clearly, the cost and complexity of microlithography puts it out of the reach of casual attackers. This is a principal advantage of using physical one-way hash functions for authentication: there is a significant asymmetry between the cost of making a single token and cloning it.

10.4.2 Electron beam lithography

Photolithography can produce feature sizes on the 0.1 micron scale. Smaller features are possible by using electron beam lithography, which brings us into the realm of nanofabrication. In this case, a computer-controlled electron beam alters the chemistry of a resist, usually poly-methyl methacrylate (PMMA). Although this technique can produce feature sizes on the order of 20 – 30 nanometers, it is a serial process and very slow. It is not a practical process if rapidity is a criterion.

10.4.3 Scanned probe lithography

This technique uses variations on a scanning probe microscope to either “plow” a groove through or cause local changes in the electrochemistry of a substrate. This technique can produce 20 – 30 nanometer features. Limitations of this process include: low temperature requirements, ultra high vacuum, and, usually, conducting substrates.

10.4.4 Summary

It is clear from the foregoing discussion that the existing techniques of micro- and nanofabrication are expensive and are largely geared towards the production of regular structures. Ultimately, it might be possible to build a

machine to copy any arbitrarily random microstructures but this is likely to be a complicated and expensive process. We suggest that the asymmetry in cost, between producing an arbitrarily complicated 3D structure and cloning it, is a security resource and can be exploited to build authentication, and perhaps, cryptosystems.

10.5 Fabrication complexity

We are now ready to discuss fabrication complexity which we define as *the minimal computational and physical resources required to clone a physical system*.

Let us recall why we are interested in determine fabrication complexity. When we defined physical one-way functions in section 6.2. one part of the definition stipulated that materially constructing a distinct physical system containing the same secret should be “hard”. We did not define what we meant by “hard”. Ideally, we would like cloning to be impossible in the same way that it is theoretically impossible to clone single quantum states. However, in general, it is not impossible to clone arbitrary classical physical systems. Here we ask what the difficulty is.

10.5.1 Notation

We use the same notation as before, recalled here for convenience.

Let Σ be a physical system containing a secret $X \in \{0, 1\}^l$. X is some property or microstate of the physical system and l is a polynomial function of some physical resource such as volume, energy, space, matter *et cetera*.

Let $z \in \{0, 1\}^k$ be a specific state of a physical probe P such that k is a polynomial function of some physical resource. Henceforth, a probe P in state z will be denoted by P_z .

Let $y = f(X, P_z) \in \{0, 1\}^n$ be the output of the interaction between system Σ containing secret X and probe P_z .

10.5.2 Problem definition

We approach the problem of determining fabrication complexity as follows: let us assume that we already know the secret X of a given physical system Σ to some predetermined accuracy. We would like to transmit a compact description this secret to a machine that produces a distinct physical system Σ' containing an identical secret $X' = X$. We will be satisfied with the reproduction if, for every possible probe state, the outputs of Σ' are indistinguishable from those of Σ . That is:

$$|f(X, P_r) - f(X', P_r)| < \epsilon \quad 10.5.1$$

for all possible P_r , where P_r is a randomly chosen probe state and ϵ can be made as small as we like.

10.5.3 A Universal Fabrication Machine (UFM)

Fabrication complexity may be thought of as having two separate parts: a compact description of X and a machine which transforms this description into a physical system. We will employ Kolmogorov Complexity (see section 2.5.2) to develop an algorithmically minimal (i.e., compact) description of X and a Universal Fabrication Machine (UFM) to construct the physical system.

A UFM is simply a Universal Turing Machine augmented with a fabrication head. A conceptual diagram is shown below.

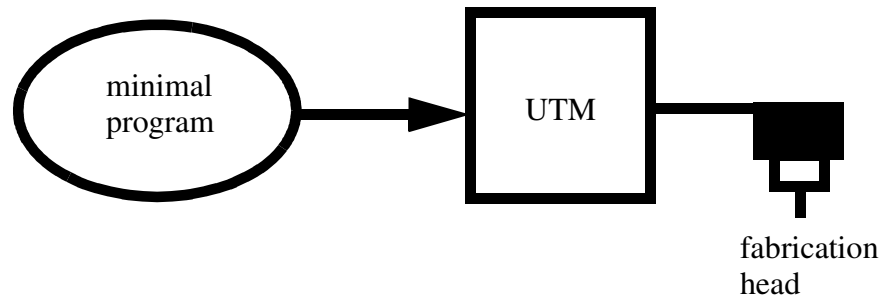


FIGURE 10.2 A CONCEPTUAL DIAGRAM OF A UNIVERSAL FABRICATION MACHINE

In principle, the operation of a UFM is very simple. It receives as input an algorithmically minimal description of the secret and decodes this program to output a spatial description of the physical system. The fabrication head, which has access to a palette of raw materials, then transforms the spatial description into a physical object. In doing so, the fabrication head utilizes physical resources which have a finite cost associated with them. This cost is also a component of the fabrication complexity. Therefore:

$$\text{Fabrication complexity} = (\text{Computational Complexity} + \text{Physical Resources})$$

We note that the algorithmically minimal description of X is independent of any specific method of fabrication. However, the program which decodes the description must know about the method of fabrication in order to produce output that is compatible with the fabrication method.

10.5.4 Kolmogorov complexity of disordered structures

Consider the case where we have a disordered structure of volume $V = L^3$ and uniformly distributed scatterers located at a mean free path l from each other. For now, we assume that all the scatterers are spheres of the same radius. What is the Kolmogorov Complexity of this structure?

Let us divide up the volume into cubical cells of volume l^3 . Because we assumed that the scatterers are uniformly distributed in the volume, there is a high probability that each cell contains exactly one scatterer. Each of these scatterers requires 3 integers to describe its location in space. The number of bits required to describe the location of one scatterer is then

$$3\log_2\left(\frac{L}{l}\right) \text{ bits} \quad 10.5.2$$

We know that there are approximately $N = (L/l)^3$ scatterers in the volume. Therefore the total number of bits K required to describe the locations of all the scatterers is

$$K = 3N\log_2\left(\frac{L}{l}\right) = 3\left(\frac{L}{l}\right)^3 \log_2\left(\frac{L}{l}\right) \text{ bits} \quad 10.5.3$$

Additionally, we need a few more bits to describe the radius and to make the program self-delimiting, i.e., the program contains all the information required to produce the spatial description and halt. This correction is usually a constant and we will implicitly add an $O(1)$ term to our Kolmogorov Complexity.

Equation 10.5.3 is an estimate of the number of bits required to describe the structure as derived from first principles. We know that the minimal program contains approximately the same number of bits. This is the program that is provided as input to the UFM.

10.5.5 Physical resources used in fabrication

The fabrication head of the UFM receives as input a spatial description which it transforms into a physical object. The resources used in order to accomplish this will vary with the specific fabrication method used. Assuming there is a minimum cost per operation, however, we can reasonably expect that the cost increases polynomially as the size of the system to be fabricated increases. For example, we can estimate the physical resource cost of fabricating the structure in figure 1.1 by determining the energy it takes to place one atom on the substrate and multiplying by the number of atoms. In 3D we expect the number of atoms in a given volume to increase as the cube of the linear dimension. Hence, we conclude that the physical resource cost is a polynomial function in the size of the structure.

We observe that a polynomial increase in the size of the structure results in an much larger exponential size in the number of possible structures as viewed by a probe.

10.6 Parallel fabrication attack

Having discussed the various fabrication methods and the concept of fabrication complexity, we now briefly look at another kind of attack that might be perpetrated. We refer to this kind of attack as a *parallel fabrication attack*.

Our version of a UFM is a serial machine. It is not difficult to imagine a

machine which copies 3D microstructures in parallel. Conceptually, such a machine could operate as follows. A tomographic technique, operating in a resolution regime below that of the physical probe, could image slices of the 3D microstructure and use those images to create layers of the 3D microstructure. One way to do this might be to selectively cure epoxy using the tomographic images as masks.

Let us consider the requirements on such a parallel fabrication machine. First, its lateral resolution must be greater than the wavelength λ_p of the probe used to derive the unique identifier. This could be achieved by using illumination whose wavelength $\lambda_T \ll \lambda_p$. This would enable the construction of features which are continuous with respect to probe radiation. Second, the longitudinal resolution would also have to be substantially smaller than λ_p . Finally, the time take to construct each layer would have to be small because a large number of layers need to be built. For our example, with $L = 10$ mm and $\lambda_p = 0.5$ microns, we would require a lateral and longitudinal resolution of greater than 0.05 microns at which wavelength about 200,000 layers must be fabricated.

A machine with these capabilities does not exist today. However, it is not inconceivable that it could be made in the near future. We believe that the greatest threat to the security of POWF-based authentication systems is posed by such parallel fabrication machines.

10.7 Summary

In this chapter, we discussed a few issues that could not logically fit into any other chapter. First, we took a look at scaling issues of physical authentication systems. Scaling can occur in two ways: either the size of the 3D structure could increase, or the number of tokens could go up. The scaling performance with respect to increase in the size of the structure is well understood. Essentially, the number of structures distinguishable by a probe is exponential in the ratio (L/l) , where L is a linear dimension of the structure and l is the mean free path between scatterers.

We then considered the effects of scaling the number of tokens in circulation. We examined the reasons why the average Hamming Distance between like speckle patterns was as great as it is (0.2525) and suggested ways in which this could be reduced. Based on the facts that (a) the number of possible structures distinguishable by a probe is exponential in the size of the structure (b) the effective number of possible Gabor Hashes was almost the same as the number of structures and (c) the average Hamming Distance between two unrelated Gabor Hashes is 0.5, we concluded that increasing the number of tokens would not cause any significant degradation in robustness, as long as the number of tokens in circulation is smaller than the number of possible tokens.

We then considered three different kinds of attacks on a physical authentication system. The first two, brute-force attack and the birthday attack, are attacks on the Gabor Hash strings, and are completely independent of the physical system. We calculated that with a brute force attack, an adversary would have to try on the order of 10^{69} strings in order to

compromise the system. If the adversary were to try the birthday attack instead, and try to find two 3D structures which hash down to the same value with probability at least 0.5, then she would have to try on the order of 10^{35} tokens.

Replay attacks were then considered. There are two cases. The first is where an adversary stores up all possible responses and the second is when the response is computationally simulated. For each of these cases, we considered plausible ways in which the attack could be perpetrated, and concluded that these attacks, while feasible, were impractical.

Then we briefly considered current-day microfabrication methods in order to gain some insight into their cost and complexity. It is clear that these methods are extremely expensive and are geared towards the production of large numbers of identical and regular structures, and are not, in general, applicable to the construction of arbitrary inhomogeneous microstructures.

This led to a discussion of fabrication complexity, which we define as the minimal computational and physical resources required to clone a physical system. We discussed fabrication complexity in the context of an idealized Universal Fabrication Machine — a Universal Turing Machine augmented with a fabrication head. The purpose of this machine is to transform an algorithmically minimal (in the Kolmogorov sense) description of a physical structure into the structure itself. We calculated the size of an algorithmically minimal description of a disordered 3D structure and showed how the number of structures allowed by this description is on the same order as that derived from the physics of coherent multiple scattering.

Finally, we discussed a parallel fabrication attack, where the 3D structure is cloned a whole layer at a time, and came up with a potential structure for a 3D photocopying machine and some basic limits on its performance.

11 Contributions and future work

In this, the final chapter of the dissertation, I (as opposed to the scientific "we") summarize the original contributions and discuss work that should be undertaken in the future.

11.1 Summary and original contributions

The early goal of this research project was simply to derive unique and tamper-resistant identifiers from *three-dimensional* structures at a very low cost-per-bit. This goal was motivated by several factors. Primary among them were the emergence of value-bearing tokens in large quantities (e.g. downloadable postage stamps, smart cards with monetary value) and the amazing increase in 2D imaging and fabrication capabilities.

An extreme example of 2D fabrication is seen in figure 1.1. 2D scanning is not far behind - in fact, the HP Capshare freehand scanner assembles rectilinear images from different swaths by correlating paper texture features at swath boundaries. Another example is the Microsoft Intellimouse which captures images of the work surface at over 1500 fps in order to accurately track mouse movement. The fact that such advanced 2D imaging is available in consumer-grade electronics costing on the order of a hundred dollars is a threat to physical authentication systems which rely solely on 2D features.

Thus, I decided to focus solely on inhomogeneous 3D structures to derive identity information. Then, the first choice that had to be made was to decide on the 3D imaging technique. The contenders for subsurface imaging were optical coherence tomography (OCT), confocal microscopy (CM), and magnetic resonance imaging (MRI). OCT seemed to be the best choice, given that it could image whole slices of the structure at a time, and was based on very simple optical principles, which meant the token reader had the potential to be inexpensive. However, MRI was attractive as well because the quantum computing group were already engaged in building a table-top NMR spectrometer. However, all these techniques had one fatal flaw - modifying a small region of the structure had a correspondingly small effect on the images. Tamper detection in the face of noisy imaging seemed to be only remotely possible.

Inspired by a patent application by Nabil Amer, David DiVincenzo, and Neil Gershenfeld, I noticed that coherent multiple scattering from inhomogeneous 3D microstructures possessed many of the same properties and, more importantly, asymmetries, as algorithmic one-way functions. This observation led me to the concept of physical-one way functions and to pose the original problem in the framework of these functions.

In the next few paragraphs, I will outline the contributions of this dissertation.

- I believe that framing the present problem of generating unique tamper-resistant identifiers (and any future work related to physical authentication and cryptography) in the language of algorithmic cryptography is in itself a useful contribution. Modern algorithmic

cryptography is divided into two distinct activities: the definitional activity, which is the rigorous definition of cryptographic tasks that capture natural security concerns, and the constructional activity, which is the design and analysis of cryptographic schemes satisfying the definitions[11]. I believe that adopting the same approach in the physical domain is both fruitful and essential in order to examine the security properties of physically-based cryptosystems in a rigorous manner. Previous efforts in the field of physical authentication have, to the best of my knowledge, not made an explicit connection to algorithmic cryptography.

- In this vein, the definition of physical-one way (hash) functions along the lines described above is a contribution. The definitions I proposed are preliminary, and will no doubt be subject to change. However, they are a good starting point for two reasons. First, they will allow us to clearly examine the security properties of physical authentication based on coherent multiple scattering and second, they will enable us to search for other (classical) physical systems which might be usable as authentication and cryptographic systems.
- Identifying a candidate physical system and showing, both theoretically and experimentally, that it satisfies the definition is a contribution. Indeed, the definition mentioned above emerged from the properties of coherent multiple scattering, not the other way round. Specifically, showing that POWFs are collision resistant and that they exhibit the avalanche effect are important pieces of the picture.
- The experimental verification of the fact that coherent multiple scattering allows us to derive unique tamper-resistant identifiers at a very low cost-per-bit consumed a whole year. Simply looking at the final system used to gather data does not betray the vast amounts of time spent in building several (approximately 10) candidate registration systems. Ultimately, however, the final system performed exceedingly well. I believe that demonstrating that our physical authentication system is both practical and useful is an essential contribution to the notion of physical one-way functions.
- Another piece of the puzzle are the two protocols I presented. The one-time pad protocol is needed because one of the inputs to the physical one-way function is a 3D structure and is in possession of a, possibly malicious, user of the authentication system. Pre-acquiring speckle patterns and storing them up at a trusted site simulates the second copy of the pad required in a one-time pad protocol. Of course, there may be hidden flaws which are immediately obvious to a cryptographer, but my goal was to simply show that it is possible to construct protocols based on physical one-way hash functions. The bit-commitment protocol is also very rudimentary but it does demonstrate that, at least for honest users, it is possible to develop a scheme which fulfils all the requirements.

- The discussion of possible attacks on a physical authentication system, both digital and physical, is a useful contribution.
- Finally, I believe that the concept of fabrication complexity, which bounds the information, energy, and time required to fabricate a physical system in a specific internal state is generally a very useful concept. As far as I can tell, previous authors have not addressed the question of the total computational and physical resources required to construct a *specific instance* of a physical system. I suspect that this is partly due to the heavy reliance on statistical entropy as a measure of information which forces thought in the direction of *ensembles* of physical systems. Kolmogorov complexity allows the quantification of information in a specific state and is thus a very useful candidate of bounding the computational component of fabrication complexity. The second reason why fabrication complexity has not been discussed before is because there has not been any real need to do so. In our case, all security vanishes if a physical system is cloned. It is therefore important to know the minimum effort required to clone a given physical system. This effort is captured by the notion of fabrication complexity

11.2 Future work

I view this dissertation as a starting point for the principled and rigorous study of physical cryptosystems. In this section, I want to offer suggestions for future efforts based on the general idea of physical one-way functions.

- On the theoretical front, some work needs to be done to sharpen the definitions of physical one-way functions and fabrication complexity to make them as general and physical-system-independent as possible. The influence of coherent multiple scattering is writ large on the existing definitions.
- In the context of the specific implementation of a physical authentication system, several improvements can be made. A version of the token-reader that is much more portable and compact needs to be built. This could incorporate a high-performance registration system and an isolated optical chain to diminish the amount of noise in the system. In the same vein, tokens whose 3D structure is also isolated from the environment need to be produced. Finally, a lossless compression scheme should be applied to the Gabor hash strings in order to reduce storage requirements and produce identifiers in which each bit is statistically independent of every other bit.
- Some more work needs to be done to precisely classify the complexity class of simulating the passage of light through both linear and nonlinear disordered structures. I have a feeling that simulating the passage of light through nonlinear media is **NP**-complete and would like to verify it.
- I have already mentioned that the computational and physical complexity of the physical system increases as we approach the optical localization threshold, where the wavelength of incident radiation is on the same order

as the mean free path. A precursor to localization, optical weak localization, results in enhanced backscattering of light. One simple way to make the linear physical authentication system harder to simulate and spoof would be to derive *two* speckle patterns - one on the same side of the token and in the same direction as the incident radiation, and one on the opposite side, as we have always done. This would certainly ramp up the space/time computational complexity of simulation. Unfortunately, it might very well increase the cost and complexity of the token reader. As a scaling problem, however, it certainly merits investigation.

- An application that I would like to see implemented is an authenticated camera. Schneier et al. [75] describe develop protocols for an authenticated camera that allows people to verify that a given digital image was taken by a specific camera at a specific time and specific place. These protocols require interaction between the camera and base station both before and after a series of images are taken. It should be possible to implement an authenticated camera with a POWF-based system. One could imagine a small blob of epoxy with several scatterers in it permanently bonded to a portion of the CCD detector of a network-enabled digital camera. The camera could then be augmented with a semiconductor laser to probe the structure. Once the camera is augmented in this way, a one-time pad protocol for untrusted readers can be executed each time the user takes a picture. I believe this general idea could be extended to many other network-enabled objects.
- I have suggested that the general framework of physical one-way functions could be used to search for other physical systems which could be used in authentication. One such system immediately comes to mind: the electronic analog of coherent multiple scattering. This field is usually referred to as electronic transport in mesoscopic systems [76][77]. A lot of the pioneering work in this field was done by Rolf Landauer. There are many similarities between the propagation of light through disordered microstructures and that of electrons through disordered wires. It has been shown, at very low temperatures, that the conductance of disordered wired fluctuates in much the same way as speckle, and the mathematical treatment of conductance fluctuations is identical to that of speckle [78][79][80]. Given this, I can imagine that it would be possible to use POWFs to generate unique identifiers using silicon microstructures. Specifically, I can imagine every silicon chip having its own unique tamper-resistant identifier. Admittedly, this is impractical now, but it might not be in the future.
- Finally, there is one connection to algorithmic cryptography that I have not addressed in this dissertation: *trapdoor functions*. It would be interesting to discover physical systems which can be fashioned into physical one-way trapdoor functions, in order to enable physically inverting the functions in constant time. Of course, we would still require that all simulation of the physical system be impractical or infeasible. It is not immediately clear to me whether physical one-way trapdoor functions will have any practical utility in implementing cryptosystems, given their

dependence on a specific instance of a physical system. However, discovering these physical systems is essential in order to fill out the space of cryptographically motivated physical functions.

12 References

- [1] Wiesner, S., Conjugate coding, SIGACT News, vol. 15, no. 1, 1983
- [2] Bennett, C., Brassard, G., Briedbart, S., and Wiesner, S., Quantum Cryptography, or unforgeable subway tokens, Advances in Cryptology, CRYPTO82, Springer Verlag, 1998.
- [3] Bennett, C., and Brassard, G., Quantum Cryptography: Public key distribution and coin tossing, Proc. IEEE International Conference on Computers, Systems, and Signal Processing, pp. 175-179, 1984.
- [4] Wootters, W. K., and Zurek, W., A single quantum cannot be cloned, Nature, Vol. 299, pp. 982-983, 28 October 1982.
- [5] Papadimitriou, C. H., Computational Complexity, Addison-Wesley, Reading, MA, 1994.
- [6] Greenlaw, R., and Hoover, H. J., Fundamentals of the Theory of Computation, Morgan Kaufman Publishers, San Francisco, CA, 1998.
- [7] Corman, T. H., Leiserson, C. E., and Rivest, R. L., Introduction to Algorithms, The MIT Press, Cambridge, MA, 1993.
- [8] Turing, A. M., On computable numbers and the Entscheidungsproblem, Proc. London Math. Soc. Ser. 2, 42, pp. 230-265, 1937.
- [9] Diffie, W., and Hellman, M., New Directions in Cryptography, IEEE Transactions on Information Theory, vol. IT-22, no. 6, pp. 644-654, 1976.
- [10] Needham, R., in Wilkes, M.V., Time-sharing Computer Systems, Elsevier, New York, pp. 91, 1972.
- [11] Goldreich, O., Modern Cryptography, Probabilistic Proofs, and Pseudorandomness, Springer, 1999.
- [12] Goldreich, O., Foundations of Cryptography (Fragments of a Book), available only online at <http://www.toc.lcs.mit.edu/~oded/frag.html>
- [13] Bellare, M., A note on negligible functions, Technical Report CS97-529, Department of Computer Science and Engineering, University of California at San Diego, March 1997.
- [14] Simmons, G., Contemporary Cryptology: the science of information integrity, IEEE Press, 1992.
- [15] Rompel, J., One-way functions are necessary and sufficient for secure signatures, Proceedings of the Twenty Second Annual ACM Symposium on Theory of Computing, pp. 387-394, 1990.

- [16] Hastad, J., Impagliazzo, R., Levin, L., and Luby, M., A pseudorandom generator from any one-way function, *SIAM Journal on Computing*, vol. 28, no. 4, pp. 1364-1396, 1999.
- [17] Schneier, B., *Applied Cryptography*, John Wiley and Sons, 1996.
- [18] van Renesse, R., *Optical Document Security*, Artech House, 1998.
- [19] Brosow, J., Method and system for verifying authenticity safe against forgery, US patent no. 4218674, 1980.
- [20] Goldman, R., Verification system for document substance and content, US patent no. 4568936, 1986.
- [21] Samyn, J., Method and apparatus for checking the authenticity of documents, US patent no. 4820912, 1989.
- [22] Denenberg, S., System for registration, identification, and verification of items utilizing unique intrinsic features, US patent no. 5521984, 1996.
- [23] van Renesse, R., 3DAS- a 3D structure authentication system, *Proceedings of the European Convention on Security and Detection*, IEE, 1995
- [24] Smith, J. R., and Sutherland, A. V., Microstructure based indicia, *Proceedings of the Automatic Identification Advanced Technologies AutoID'99*, pp. 79-83, 1999.
- [25] Amer, N., DiVincinzo, D. P., and Gershenfeld, N., Tamper detection using bulk multiple scattering, US Patent no. 5790025, 1998.
- [26] Gershenfeld, N., personal communication, 2001.
- [27] Huang, D., et al., Optical coherence tomography, *Science*, v. 254, issue 5035, pp. 1178-1181, 1991.
- [28] Huang, D., *Optical Coherence Tomography*, PhD Thesis (HST), MIT, 1993.
- [29] Navarro, R., et al., Image representation with Gabor wavelets and its applications, *Advances in Imaging and Electron Physics*, 97, P. W. Hawkes, ed., Academic Press, San Diego, CA, 1996.
- [30] Daugman, J. G., Complete discrete 2D Gabor transform by neural networks for image analysis and compression, *IEEE Transactions on ASSP*, 36, pp. 1169-1179, 1988.
- [31] Cristobal, G., and Navarro, R., Space and frequency variant image enhancement based on a Gabor representation, *Pattern Recognition Letters*, 15, pp. 273-277, 1994

- [32] Heeger, D. J., Model for the extraction of image flow, *Journal of the OSA A*, 4, pp. 1455-1471, 1987.
- [33] Daugman, J. G., Spatial visual channels in the Fourier plane, *Vision Research*, 24, pp. 891-910, 1984.
- [34] Daugman, J. G., Two-dimensional spectral analysis of cortical field receptive field profiles, *Vision Research*, 20, pp. 847-856, 1980.
- [35] Daugman, J. G., Uncertainty relation for space, spatial frequency, and orientation optimized by two dimensional visual cortical filters, *Journal of the OSA A*, 2, pp. 1160-1169, 1985.
- [36] Gabor, D., Theory of communication, *Journal of the Institute of Electrical Engineers*, 93, pp. 429-457, 1946.
- [37] Mallat, S. G., A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, pp. 674-693, 1989.
- [38] Adelson, E. H., and Burt, P. J., Image Data Compression with the Laplacian Pyramid, *Proceeding of the Conference on Pattern Recognition and Image Processing*, pp. 218-223, 1981.
- [39] Nestares, O., et al., Efficient spatial domain implementation of a multiscale image representation based on Gabor functions, *Journal of Electronic Imaging*, 7, pp. 166-173, 1998.
- [40] L.J. Hornbeck, "Deformable-Mirror Spatial Light Modulators," *Spatial Light Modulators and Applications III, SPIE Critical Reviews*, Vol. 1150, pp. 86-102, 1989.
- [41] L.J. Hornbeck and W.E. Nelson, "Bistable Deformable Mirror Device," *OSA Technical Digest Series Vol. 8, Spatial Light Modulators and Applications*, p. 107, 1988.
- [42] R.J. Gove, "DMD Display Systems: The Impact of an All-Digital Display," *Society for Information Display International Symposium*, 1994.
- [43] Landau, L. D., and Lifshitz, E. M., *Quantum Mechanics*, vol. 3 of a course in theoretical physics, Pergamon Press, Oxford, 1977.
- [44] Anderson, P. W., Absence of diffusion in certain random lattices, *Physical Review*, 109, pp. 1492-1505, 1958.
- [45] Ioffe, A. F., and Regel, A. R., *Progress in Semiconductors*, 4, pp. 237, 1960.

- [46] Van Albada, M. P., and Lagendijk, A. D., Observation of weak localization of light in a random medium, *Physical Review Letters*, 55, pp. 2692-2695, 1985.
- [47] Wolf, P. E., and Maret, G., Weak localization and coherent backscattering of photons in disordered media, *Physical Review Letters*, 55, pp. 2696-2699, 1985.
- [48] Anderson, P. W., The question of classical localization: a theory of white paint? *Philosophical Magazine B*, 52, pp. 505-509, 1985.
- [49] Goodman, J. W., Statistical properties of laser speckle patterns, in *Laser Speckle and Related Phenomena*, in J.W. Dainty ed., Springer-Verlag, 1975.
- [50] Goodman, J. W., *Statistical Optics*, Wiley Interscience, 1985.,
- [51] Freund, I., Rosenbluh, M., and Feng, S., Memory effects in propagation of optical waves through disordered media, *Physical Review Letters*, 61, pp. 2328-2331, 1988.
- [52] Feng, S., Kane, C., Lee, P. A., and Stone, A. D., Correlations and fluctuations of coherent wave transmission through disordered media, *Physical Review Letters*, 61, pp. 834-837, 1988.
- [53] Doniach, S., and Sondheimer, E.H., *Greens Functions for Solid State Physicists*, Benjamin, London, 1974.
- [54] Spivak, B., and Zyuzin, A., Mesoscopic sensitivity of speckles in disordered nonlinear media to changes of the scattering potential, *Physical Review Letters*, 84, pp. 1970-1973, 2000.
- [55] Skipetrov, S.E., and Maynard, R., Instabilities of waves in nonlinear disordered media, *Physical Review Letters*, 85, pp. 736-739, 2000.
- [56] Akkermans, E., Wolf, P.E., Maynard, R., and Maret, G., Theoretical study of coherent backscattering of light by disordered media, *Journal of Physics, France*, 49, pp. 79-98, 1988.
- [57] Berkovits, R., Sensitivity of the multiple scattering speckle pattern to the motion of a single scatterer, *Physical Review B*, vol. 43, no. 10, pp. 8638-8640, 1991.
- [58] Kogan, E., and Kaveh., M., Random-matrix-theory approach to the intensity distributions of waves propagating in a random medium, *Physical Review B*, vol. 52, no. 6, pp. R3813-R3815, 1995.
- [59] Bohren, C., and Huffman. D., *Absorption and scattering of light by small particles*, Wiley Interscience, 1983.

- [60] van de Hulst, H. C., *Light scattering by small particles*, Dover, 1981.
- [61] Shannon, C., and Weaver, W., *A mathematical theory of communication*, University of Illinois Press, 1949.
- [62] Wiener, N., *Time Series*, MIT Press, 1949.
- [63] Rivest, R., The MD5 Message Digest Algorithm, Internet RFC 1321, 1992. (Available online at <http://theory.lcs.mit.edu/~rivest/publications.html>)
- [64] Halevi, S., and Micali, S., Practical and provably-secure commitment schemes from collision-free hashing, *Advances in Cryptography - CRYPTO '96*, LNCS vol. 1109, pp. 201-215, Springer-Verlag. 1996.
- [65] Marks, L., *Between Silk and Cyanide - a Codemaker's War 1941-1945*, Free Press, 1999.
- [66] Bennett, C., in Zurek, W. H., ed., *How to define complexity in physics, and why*, in *Complexity, Entropy, and The Physics of Information*, Santa Fe Institute Studies in the Sciences of Complexity, volume 8, pp. 137-148, 1990.
- [67] Solomonoff, R., A formal theory of inductive inference, part I, *Information and Control, Part I*, vol 7, no. 1, pp. 1-22, 1964
- [68] Solomonoff, R., A formal theory of inductive inference, part II, *Information and Control, Part II*, vol. 7, no. 2, pp. 224-254, 1964.
- [69] Kolmogorov, A.N., Three approaches to the quantitative definition of information, *Problems in Information Transmission*, vol. 1, no. 1, pp. 1-7, 1965.
- [70] Chaitin, G. J., On the length of programs for computing finite binary sequences, *Journal of the ACM*, vol. 16, pp. 145-159, 1969.
- [71] Bennett, C.H., The thermodynamics of computation---a review, *International Journal of Theoretical Physics*, vol. 21, pp. 905-940, 1982.
- [72] Schneier, B., and Shostack, A., Breaking up is hard to do: modeling security threats for smart cards, *First USENIX Symposium on Smart Cards*, 1999.
- [73] Kirkpatrick, S., and Selman, B., Critical behavior in satisfiability of random Boolean expressions, *Science*, 264, pp. 1297-1301, 1994.
- [74] Hogg, T., Huberman, B., Williams, C., *Frontiers in Problem Solving: phase transitions and complexity*, A special issue of *Artificial Intelligence*, vol. 81, nos. 1-2, 1996.

- [75] Schneier, B., Kelsey, J., Wagner, D., and Hall, C., An authenticated camera, 12th Annual Computer Security Applications Conference, ACM Press, pp. 24-30, 1996.
- [76] Datta, S., Electronic transport in mesoscopic systems, Cambridge University Press, 1995.
- [77] Landauer, R., Conductance viewed as transmission, Reviews of Modern Physics, vol. 71, no. 2, pp. S306-S312, 1999.
- [78] Howard, R. E., Jackel, L. D., Mankiewich, P. M., and Skocpol, W. J., Electrons in silicon microstructures, Science, vol. 231, no. 4376, pp. 346-349, 1986.
- [79] Henderson, G. N., Gaylord, T. K., and Glytys, E. N., Ballistic electron transport in semiconductor heterostructures and its analogies in electromagnetic propagation in general dielectrics, Proceedings of the IEEE, vol. 79, no. 11, pp. 1643-1659, 1991.
- [80] Chan, I. H., Clarke, R. M., Marcus, C. M., Campman, K., and Gossard, A. C., Ballistic conductance fluctuations in shape space, Physical Review Letters, vol. 74, no. 19, pp. 3876-3879, 1995.
- [81] Kailath, T., et al., Multivariable control, simulation, optimization and signal processing for the microlithographic process, Multidisciplinary University Research Initiative Report, Stanford University, (<http://www-isl.stanford.edu/groups/MURI/answers/answers.html>), 1997.
- [82] Bruggeman, B., et al., Microlithography cost analysis, Interface '99 Symposium, (<http://www.mentor.com/dsm>), 1999.
- [83] Datasheet available at <http://www.dalsemi.com/datasheets/pdfs/2401.pdf>
- [84] Bennett, C., Brassard, G., Crepeau, C., and Maurer, U., Generalized privacy amplification, IEEE Transactions on Information Theory, vol. 41, no. 6, 1995.
- [85] Online demo of privacy amplification available from www.inf.ethz.ch/departement/TI/um/research/keydemo/Overview.html
- [86] <http://www.unicate.nl>