Chapter 10

Conclusions and Future Work

As Professor Bill Ramsey says when he speaks of J. S. Bach's *Saint Matthew's Passion*, "That was a long song."

The first half of this dissertation is devoted to the tasks of representing wafer map probabilities and developing methods to generate and label synthetic wafer maps. Five geometries are chosen for representation, resulting in nine Shape categories: Segments, Bands, A-Disks and B-Disks, A-Annuli and B-Annuli, and A-Rings, B-Rings, and C-Rings. These shapes are regarded as objects, and geometrical probability concepts are applied to discover expressions for the uniform pdf of each object. These pdf's allow us to create a wafer map as a random sample of one of the shape populations. Equations to calculate feature variables derive from a knowledge of the geometry. The feature variables and pdf's are used to create `wafgen`, our wafer map pattern generator.

The second half of the dissertation is devoted to the construction and testing of a variety of classifiers. Work to create a theoretical model of classification in noise is described, as is a new analysis of noise clustering in industrial wafer maps. We carry out six classification studies and analyze the results. In study one, we prove experimentally that the Bayes error of `wafgen`'s random wafer map population is less than 1.5% for Location classification and smaller still for the other four features. In studies two through six, we carry out a series of classification experiments, the results of which allow us to make the following general statements:

1. N20 is less sensitive than kh to low levels of random noise ($p \leq 0.15$ or so).

2. N20 is more robust against changes to the input priors.

3. kh classifies industrial wafer maps more accurately than N20.

4. XdatJ, made from industrial wafer maps, identifies other industrial wafer maps as accurately as N20 does.

5. Xdata is noticeably less capable than N20 on an input set of synthetic wafer maps.

10.1 Contributions

The work described in this dissertation was initiated by process engineers at Digital Equipment Corporation. They wondered whether it would be possible to create software that could pick out "interesting" failure patterns. After several discussions between the engineers and myself, it became apparent that a large fraction of what is interesting is perceived geometrically. Patterns such as bars, filled circles, and half-wafers were suggested by the engineers. Variations in area, location, and orientation were clearly relevant. A demonstration I produced using 9×9 wafer maps proved the need for a larger format. Other decisions and research contributions are mine alone.

I recognized early that simple geometrical objects made from lines and circles were likely to be ripe for mathematical characterization and hence synthetic imitation by computer. The construction of a statistical and geometrical representation using geometrical probability principles is a major contribution and is described in Chapters 2 through 4.

Chapter 2 lays the foundation for this contribution. Shape categories are defined in terms of boundaries and zones, and constraints are imposed in order to restrict the possibilities. Nine Shapes are obtained, five with two zones—Segments, A-Disks, B-Disks, A-Annuli, and B-Annuli—and four with three zones—Bands, A-Rings, B-Rings, and C-Rings. Two additional one-zone Shapes—Pass and Fail—are added to complete the set. Although the Digital engineers suggested shapes like these, I formalize them and add Annulus patterns in order to close the set of one-zone and two-zone patterns over logical inversion.

Chapter 3 introduces the principle of the invariant pdf, which is one that does not change under scaling, rotation, and translation. The invariant pdf defines the uniform random distribution for a geometrical object, transforming geometry into statistics. I find each invariant pdf along with its domain and derive all of the marginal pdf's. This is the first application of geometrical probability concepts to wafer map characterization.

In Chapter 4, the geometrical variables of Chapter 3 are used to derive equations for features such as area, location, and orientation. In addition, the statistical distribution of these features is characterized for each of nine shapes. Some shapes cannot be identified uniquely using only area and location, so features such as width, orientation, and direction of curvature are added in order to count and classify some objects properly. There are three contributions in this chapter. First is the definition of each feature variable such that its value changes continuously even as one pattern shape morphs into another. Second is the development of equations that define the feature variables in terms of the geometrical variables. Third is the derivation of domain limits for each feature variable.

Chapter 5 describes `wafgen`, a software package that applies the equations developed in Chapters 3 and 4 to the computer generation of synthetic wafer maps. These synthetic wafer maps comprise an entirely new, statistically well-defined, extremely large, and automatically labeled dataset. In related work, Section 8.1 offers empirical evidence that the synthetic wafer map population has a very small Bayes error—an unusual characteristic for so large a dataset. Such a new and interesting dataset is itself a contribution, and one I hope will prove useful to pattern classification research. A separate contribution related to `wafgen` is the discovery, described in Appendix B, that an orthogonal regression must be used, and special cases handled appropriately, in order to correctly calculate the Orientation of a discrete wafer map.

Chapter 6 develops procedures whereby any real-world wafer map can be modified to fit a standard format. So-called 'standardization' appears to work very well for large-scale failure patterns. The use of such re-formatting procedures allows access to all classification methods, including those requiring fixed input formats. This re-formatting algorithm is truly automatic yet universally applicable.

Chapters 7 through 9 describe the construction and execution of six studies designed to accomplish the decomposition of errors from the classification of industrial wafer maps. These studies form the first wafer map analyses of their kind and, together with their results, constitute the second major contribution of this dissertation.

Chapter 7 establishes discrete categories for each of the continuous feature variables and explains the rationale behind each set of choices. The literature contains several examples of ways to divide a wafer map into regions of various shapes and sizes, but all of these are ad hoc. The work of Section 7.1 is systematic, and the contribution is in the logical construct.

Chapters 8 and 9 present the studies themselves and their results. Table 9.7 is a significant contribution and would not exist without `wafgen` and its underlying mathematics. Additional contributions are found in knowledge of the causes of classification inaccuracy in the six studies, in the theoretical noise model of Section 8.4, and in the analysis of cluster noise in Section 9.9. Finally, Section 9.10 demonstrates that real wafer maps can be "measured," *i.e.*, values can be assigned to and a marginal pdf estimated for each of the geometrical variables. These results demonstrate that objective measurement is feasible and that geometrical probability concepts can be applied to wafer map characterization.

10.2 Future Work

More than anything else, I would like to install and use the classifiers in a production fab. Doing so would provide the kind of data one really needs to create and qualify a good classification system. Were the fab equipped with an existing knowledge system, a classifier could be added to it as one of the diagnostic tools.

I would like to add a set of rejection criteria in order to expand the classifiers into recognizers. Then either an input pattern would be rejected as 'Unknown' or classified as in this dissertation. The addition of rejection criteria may require the use of another kind of classifier, perhaps one that uses Radial Basis Functions.

It would be valuable to create data management software for the classifiers. Such software could provide an interface to the probe test and process test databases, would allow classification results to be combined with standard statistical methods, and could generate reports for use as part of auto-signoff or as part of yield management activities. With data management and recognition in place, data mining would become feasible as well.

I would like to carry out additional work in the realm of classification itself. As mentioned in Chapter 8, it is unlikely that prototype methods could be extended to three-zone patterns, *i.e.,* Bands and Rings. I would like to explore methods that would work well on two-zone patterns and be extensible to three-zone patterns. Also, I would like to experiment with the use of pre-processing procedures. The 'perfecting' methods of Section 9.2 might be applicable, but only with modifications. Pre-processing must be done with care, but these perfecting methods make some big assumptions, chief among them that the wafer map shape is identified correctly. If a wafer map shape were determined tentatively, perhaps through automatic classification, then these perfecting methods could do sufficient violence so as to *make* the wafer map look like the shape that had been identified. Obviously, less imperious pre-processing methods must be sought.

Finally, more work is needed to further the development of wafer map measurement. Classification and measurement are related, but measurement must be more precise. Consider Fig. 1.4 on page 5, where a circular boundary is imposed upon a wafer map. To my mind, it would be the ability to place a boundary with such accuracy, and then evaluate the quality of its placement, that would make a real measurement possible. It would also provide an entirely new basis upon which to evaluate wafer map data.