Quantitative surface inspection methods for metal castings

by

Michelle M. Voelker

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Industrial Engineering

Program of Study Committee:
Frank E. Peters, Major Professor
Matthew C. Frank
Cameron A. MacKenzie
Stephen B. Vardeman

Iowa State University

Ames, Iowa

2016

To my family and friends

who have supported me on my educational journey

regardless of how many years it has taken.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Frank E. Peters, my committee members, Matthew C. Frank, Stephen B. Vardeman, and Cameron A. MacKenzie for their guidance and support throughout the course of this research. I would also like to thank L. Scott Chumbley for sitting in on my defense last minute. In addition, I would like to thank my graduate and research support group, Deborah McDonough, Niechen Chen, Keon Stallard, and Paul Kemper, for aiding with the data collection process and providing technical (and moral) support throughout the course of my graduate education.

Next, I would like to thank the Steel Founders Society of America and the Army Research Laboratory for their support and funding to complete this research. The feedback of industry partners from my conference presentation at the 69th Annual Steel Founders Society of America Technical and Operating Conference was extremely beneficial in developing a standard that was easy to understand and receiving buy-in from individuals who will actually be using the standard. I would also like to express my gratitude for all of the mangers at my internships (Whirlpool, SpaceX, Procter and Gamble, and Rolls-Royce) for emphasizing the importance of public speaking and confidence, which was extremely beneficial throughout the course of this research.

Finally, I would like to thank my family and friends for motivating me to pursue my Bachelor's and Master's degree of Industrial and Manufacturing Systems Engineering at Iowa State. My success, both academically and professionally, would not have been possible without your support, and of course, plenty of homemade baked goods!

ABSTRACT

Inspecting castings to verify the quality of a part is critical for foundries to maintain a high level of customer confidence. Current methods employ qualitative methods, and the manufacturer must correctly interpret the inspection criteria set by the customer in order to meet the design specifications. The interpretation of what is acceptable often differs from the customer to manufacturer and even from inspector to inspector. In this thesis, the visual inspection of cast metal are explored in depth, and improvements to current methods are proposed.

First, the risk of Type I and II errors from the inspection process were evaluated based off of varying states of environmental and human factors in the inspection process; however, it was discovered high variation among inspectors still exists due to the subjectivity of the standards. This signals a need for a more quantitative standard to evaluate the surface of a casting. In response a digital standard is proposed, which specifies three parameters to allow the customer to communicate their exact needs in regards to surface finish to the manufacturer. These parameters are calculated based off of a part's true geometry post shrinkage in absence of surface roughness and abnormalities, or underlying geometry. Since the underlying geometry differs from the part's intended geometry, methods will be explored to estimate the underlying geometry from a point cloud of the part's surface. The proposed methods will be compared to identify which approach best estimates the ideal underlying geometry. Once an ideal method is identified, it will be used as a standard method to calculate the underlying geometry in order to create consistency among inspectors at both the customer and manufacturer.

The work completed in this thesis will raise awareness of the risk associated with current visual inspection methods for cast metal surfaces. The new, digital standard will reduce the variation in this inspection process allowing greater confidence in the parts leaving the manufacturer. Additionally, the standard will allow the customer to improve communication with the manufacturer in order to achieve the quality of surface required for their specific needs.

CHAPTER 1: INTRODUCTION

Inspection is a critical process in any foundry to verify the quality of a part. Visual inspection methods are currently employed in the evaluation of cast metal surfaces. Surface abnormalities, such as porosity, inclusions, and fusion, must be identified by an inspector who determines if they are acceptable. An example of an abnormality on a casting can be seen in Figure 1.1. The acceptance criteria set by the customer for these methods must be interpreted by the manufacturer correctly in order to meet the design specifications. Current visual inspection standards are qualitative and make it difficult to interpret these standards. Often times, the interpretation of what is acceptable differs from the customer to manufacturer and even from inspector to inspector. In this thesis, the visual inspection of cast metal surfaces is explored in depth, and improvements to current methods are proposed.

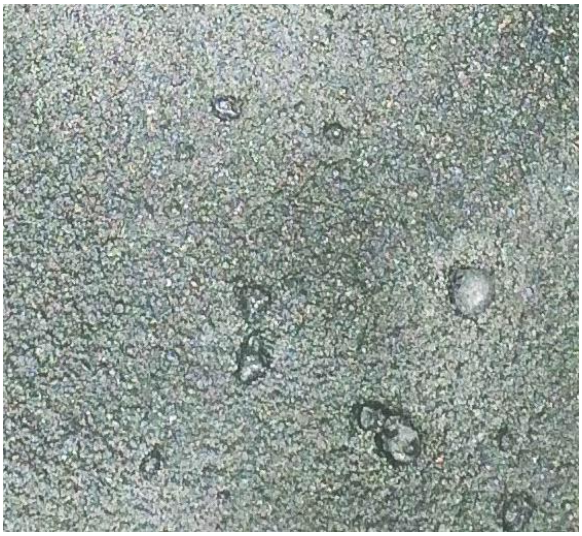


Figure 1.1—Porosity abnormalities on a casting

A risk analysis was conducted on Type I and Type II errors associated with visual inspection processes. This analysis identifies various environmental factors affecting the visual inspection process and uses an influence diagram to identify interactions among them. By using this analysis to target factors with the highest impact, a manufacturer can reduce the probability of error in his inspection process. However, even if the error is significantly reduced, inherent variability between individual inspectors will still exist.

Due to the high risk associated with the visual inspection of castings, a digital standard is proposed. This standard quantifies surfaces based on data obtained from digital scanners. These non-contact methods efficiently collect large amounts of data from the surface of the casting, which allows for increased confidence in the model as opposed to its contact counterparts, to be used in the calculation of quantitative parameters. These parameters include the baseline roughness, abnormality level, and abnormality percentage. The baseline roughness is the roughness average of the cast surface excluding abnormalities, or anomalies of the surface. The abnormality level is the maximum allowed deviation from the actual part geometry, and the abnormality percentage is the percentage of the surface contained in the region bounded by twice the baseline roughness and the abnormality level. This allows the customer to specify exactly what is needed in regards to surface requirements and allows the manufacturer to more accurately determine whether or not a part is acceptable. Additionally, it can be used to train inspectors by verifying surface results to reduce the risk of error and lay the groundwork for an automated tool for use in confirming inspection results. Inspecting castings using the new digital standard can improve communication between the customer and manufacturer in addition to reducing the discrepancies between inspectors' interpretations. This will allow for a greater confidence in the inspection process.

One challenge with calculating surface parameters from digital scans is determining the actual geometry of the casting post shrinkage without roughness and abnormalities resulting from the casting process. This geometry, known as the underlying geometry of the casting, is used to calculate parameters of the digital standard. Slicing and subsampling methods were evaluated to eliminate the roughness and abnormalities from the surface of the casting in order to estimate the underlying geometry without loss of detail in the complexity

of the castings' geometry. This is because "points on the true underlying [geometry] are not directly observable, but are observed only in the presence of error" [1].

Exploring the visual inspection of cast metal surfaces in depth allowed a greater understanding in the causes of error in this process. Since current visual and tactile methods for surface inspection were shown to have high variation among inspectors, improvements are proposed, which use a digital process to specify surface criteria. This digital standard specifies three parameters to allow the customer to communicate their exact needs in regards to surface finish to the manufacturer: the baseline roughness, the abnormality level, and the abnormality percentage. These parameters are calculated based off of a part's true geometry, or underlying geometry. Methods to calculate the estimated underlying geometry were explored in order to identify best practices to eliminate roughness and abnormalities from the data. Once an ideal method is identified, it will be used as a standard method to calculate the underlying geometry in order to create consistency among inspectors at both the customer and manufacturer.

The work completed in this thesis will raise awareness of the risk associated with current visual inspection methods for cast metal surfaces. The new, digital standard will reduce the variation in this inspection process allowing greater confidence in the parts leaving the manufacturer. Additionally, the standard will allow the customer to improve communication with the manufacturer in order to achieve the quality of surface required for their specific needs.

Research Motivation

The motivation of this research stems from the issues arising from the highly variable visual inspection process of cast metal surfaces and ambiguity in communicated surface criteria between the customer and manufacturer. The goals for this research are as follows:

1. To identify the effects of visual inspection on Type I and II errors
2. To develop a digital standard in order to reduce the subjectivity of visual inspection of cast metal surfaces
3. To evaluate and propose methods to determine the underlying geometry of castings for use in the digital standard

Thesis Organization

This thesis contains a general literature review of present research, journal articles relevant to this research, and general conclusions of the research. Three journal papers are constituted for Chapters 2-4 including the following: a risk analysis on visual inspection, an overview of the quantitative standard, and slicing methods to determine the underlying geometry. References are provided at the end of each chapter corresponding to in-chapter citations. Additionally, graphics are labeled first with the chapter they reside followed by the number of the graphic within the chapter for clarity. These chapters are followed by general conclusions and future work. Appendix A provides a draft of the digital standard.

References

[1] Castillo, Enrique Del, Bianca M. Colosimo, and Sam Davanloo Tajbakhsh. “Geodesic Gaussian Processes for the Parametric Reconstruction of a Free-Form Surface.” *Technometrics* 57.1 (2014): 87-99. Web.

CHAPTER 2: LITERATURE REVIEW

This review will discuss prior works relating to current casting inspection techniques and insights into surface evaluation of castings, including options in surface metrology, reverse engineering from point clouds, and surface cleaning algorithms.

Current Inspection Standards

Variations in the surface of castings can be present as a result of the type mold or pattern and cleaning procedures, among other factors. These can cause surface anomalies such as porosity and inclusions. Inspection methods are present in foundries to identify acceptable and unacceptable states of cast surfaces. These include the ACI Surface Indicator, MSS SP-55, ASTM A802, BNIF 359, and GAR C9 Comparator, among others [1-5]. Inspectors use comparators and images in these methods to visually classify the surface roughness and abnormalities of an actual casting. The methods are primarily qualitative and based on a discretized scale, as opposed to a continuous scale of classification. These standards were used as a reference for developing a new digital standard for cast metal surfaces. A comparison of the current inspection standards and challenges associated with each are covered more in depth in the journal articles found in Chapters 3 and 4.

Surface Metrology

Methods to calculate surface characteristics of castings were explored. These include a variety of roughness calculations, including the roughness average, root mean square, and ten point height. These methods aided in determining the optimal method for use in the digital surface standard.

One means of calculating parameters of the surface is by comparing the surface profile to a mean line. From the mean line, local minima and maxima can be located for a given sample length of a two-dimensional surface profile. Roughness can be characterized through multiple methods including amplitude, slope, and spacing of data points in a two-dimensional profile; however, the amplitude parameter is the most commonly used in the engineering field. The roughness average, denoted R_a, is the most commonly used amplitude measure and is the "arithmetic mean of the magnitude of the deviation of the profile from the mean line" [6]. Figure 2.1 shows a visual example of a R_a calculation. For a perfect fit, all data would fall on the mean line, giving a R_a value of 0.0. Deviations from the mean line would result in values greater than 0, with the roughest of surfaces having greater positive values. However, one downfall to specifying surfaces based on a R_a value is this parameter may have difficulty discriminating between surfaces as seen in Figure 2.2 [7].

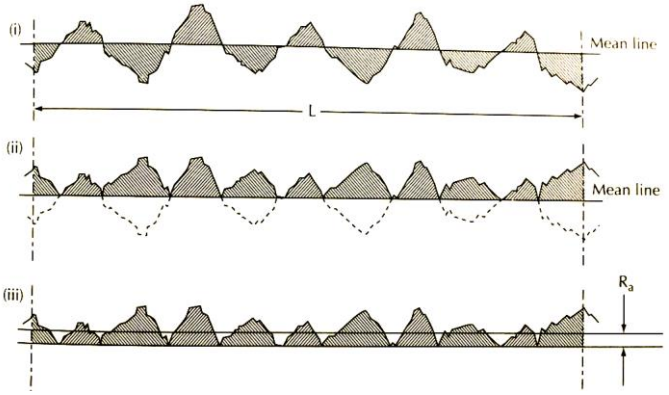


Figure 2.1—Example of roughness average calculation: i) locating mean line, ii) taking the absolute value of all points compared to the mean line, iii) calculating the average [6]

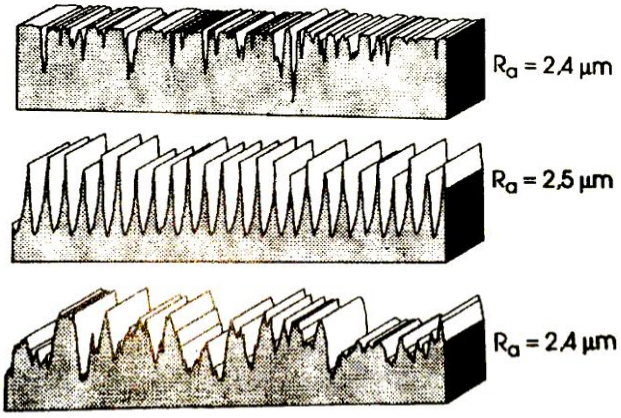


Figure 2.2—Roughness average comparisons of three unique surfaces [7]

Several less common surface parameters using the mean line may also be used in some instances. The root mean square, R_q, is another amplitude method to classify surfaces. This

method is more sensitive to outliers and is calculated by taking the square root of the average squared distances from the mean line. Other methods, such as peak to valley parameters, may also aid in classifying a rough surface with respect to the mean line. In fact, ISO 4287/1 from Japan classifies surfaces based on a ten-point height parameter, where the five most extreme points from both peaks and valleys are averaged, while in some methods, the peak count is used to estimate the number of peaks over a given sample [6]. These are universally used and commonly known methods.

One disadvantage to using amplitude measures involves outliers. If peak and valley cutoffs are not specified, the roughness parameters may not be accurate. By not specifying cutoff values, the mean line could be pulled up toward the peaks, which could skew amplitude calculations, as seen in a segmenting length example in Figure 2.3 [6, 7, 8].

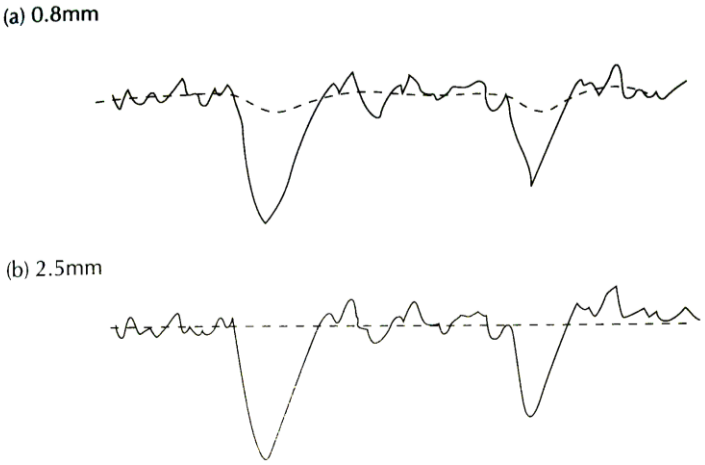


Figure 2.3—Effects of sample lengths of a) 0.8mm and b) 2.5mm on mean line calculations [6]

Other unique methods have also been implemented in industry. In a proposed method of surface evaluation for die-castings, the entire surface is broken into a grid, and a plane is fit using the least-squares method to each unit on the grid. Each unit is approximately the size of a typical surface abnormality, so units containing an abnormality will have a fit plane higher than the surrounding units. This method begins by determining the global flatness of a part from point cloud data. After, the area is broken into a grid. A filtering algorithm reduces the

noise of the sample by averaging sets of nine points in each unit, and each of the individual units is fit to a plane to determine the local flatness. This value is compared to the global flatness. Upper control limits are then set off of the global flatness, and any local flatness height exceeding the control limit is flagged as a potential "nonconformity." This method is designed to only detect surface protrusions, so surfaces with surface depressions would not apply to this method [9].

Additionally, various surface metrology methods were explored in order to collect surface data from the casting. Having a repeatable method in collecting data from surfaces is essential for consistent evaluation of surface parameters. Methods explored include contact methods, such as stylus profilometry, and non-contact methods, such as white light and laser scanning.

In contact methods, such as stylus profilometry, a stylus is pulled across the surface of the part at a constant velocity and the profile deviations in the z-direction are recorded, while other methods using a coordinate measuring machine take individual points on the part. These methods have two main disadvantages. First, contact methods are time consuming and require experienced operators to set up the equipment for proper data collection. Second, large amounts of data of an entire surface are difficult to obtain. Data obtained in these methods are a result of sampling a surface. Since the data collection process is highly manual, operators can selectively place the stylus or probe to yield results they want or test an area not representative of the entire surface. For example, stylus profilometry results in the profile roughness, R_a, from a single line of data as opposed to the entire areal roughness, S_a [10]. Additionally, this data could be limited by the size of the stylus tip. If the stylus tip is too large, the stylus may not be able to trace the profile of narrow valleys on the surface [11]. For

probe based methods using a coordinate measuring machine, the data collection process results in a set of semi-random, sparse points, which would make it difficult to obtain a surface roughness.

In laser scanning, areas with low visibility may be difficult to detect without multiple set-ups. However, articulating arms and turntable devices are used in order to obtain a complete scan of the surface with minimal registration error. The point density and distribution of data points is important for obtaining a sufficient data set for manipulation. However, due to their poor metrological performance through limited control over point acquisition location, line and structured light scanners should be used in correlation with high performing metrological devices, such as coordinate measuring machines, to improve measurement accuracy through multi-sensor data fusion [10]. Alternatively, the noise picked up by these scanners can be cleaned to remove any outlying data points prior to data manipulation [11].

Point Cloud Manipulation

In order to determine the underlying geometry, previous studies related to reverse engineering from point cloud data and surface cleaning algorithms were explored. Each method has a unique means of handling noise in either the scanning device or surface profile of the data. First, methods in manipulating point cloud data to reconstruct an accurate surface will be explored. This is important for replicating the actual surface for data manipulation for and comparison to the underlying geometry. Then, cleaning algorithms to smoothen surface profiles will be discussed for use in determining the underlying geometry.

Triangulation methods are used to convert point cloud data to a mesh. A study by Scheidigger [13] looked at various methods of triangulation to most accurately capture the true part geometry from a point cloud. Some triangulation methods, such as powercrust and cocone reconstruction, use every data point in surface reconstruction, which typically results in a non-smooth, noisy mesh model as seen in Figure 2.4b and 2.4c. Cleaning the cloud using the moving least-squares (MLS) method to reduce the noise improves the accuracy of the mesh curvature; however, directly triangulating the surface post cleaning increases the number of facets compared to the other methods (Figure 2.4d). This method is similar to direct convolution, which uses the mean height of three points in a kernel to reduce surface noise [14]. The proposed optimal curvature algorithm varies the triangle size based on the curvature of the facets, unlike standard triangulation methods (Figure 2.4e). The triangles are strategically chosen based on local curvature of the part. "Osculating circles" are constructed to choose the optimal location of each triangle to control the approximation error and determine the appropriate number of triangles for a specific curved region (Figure 2.4f). Only isosceles triangles are used to construct the mesh. This method works best with highly dense point clouds due to the MLS filter applied prior to running the algorithm to assign curvatures to the points [13].

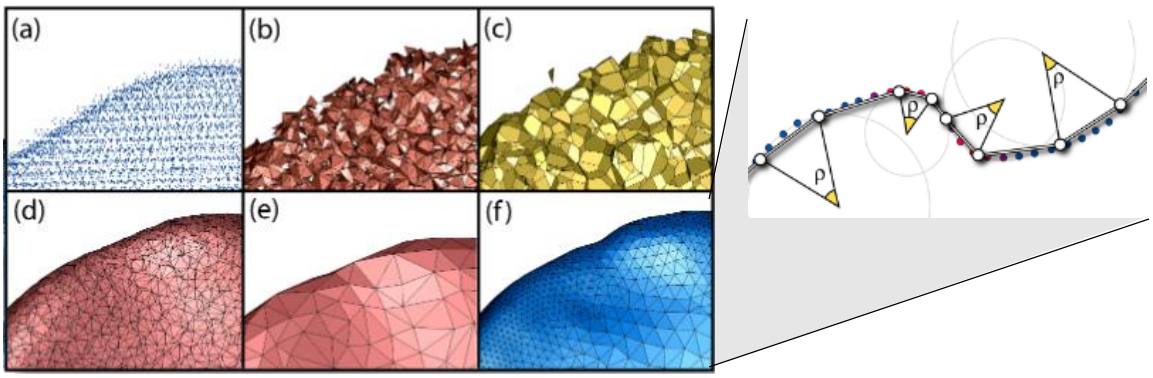


Figure 2.4—Point cloud to mesh triangulation methods a) original point cloud, b) powercrust reconstruction, c) cocone reconstruction, d) moving least-squares cleaned triangulation, e) standard triangulation, f) optimal curvature-adaptive triangulation [13]

Fitting least-square surfaces or patches to point clouds is another method in constructing a surface mesh. Shape recognition can be used in order to classify region of the point cloud with geometric shapes. This method, called random sample consensus (RANSAC), determines a best fit shape by randomly sampling the point cloud to find a minimum set of points to define the shape [15]. Weighing functions, such as a moving average iterative weighing function or Nadaraya-Watson predictor, can also be integrated into this least-squares model to minimize the effects of outliers; however, often these methods have issues at the ends of a data set or other locations where the point density is not consistent [16, 17, 18]. To accommodate for this issue, a line can be fit through the localized data to minimize the end effects [17]. Often non-rational b-spline (NURBS) surface patches are fit to highly dense point cloud data containing little noise due to their ability to accommodate for complex, three-dimensional geometries; however, splines are not as effective at minimizing the effects of dense areas of outliers and, thus, should not be used on noisy data [18, 19, 20, 21].

For surfaces with large abnormalities, constructing a mesh representative of the actual surface is difficult. The aforementioned methods take into consideration the entire surface, including abnormalities, so when the final mesh is created, the abnormalities are still present. Region growing segmentation can be used to identify these abnormalities for potential elimination. This algorithm separates a point cloud into different faces based on curvature values. Each point is observed to its neighbors, and a comparison of the angles between the normal vectors is made. If the difference in the angles falls within some specified threshold, then it is considered part of the same face [22]. This separates the part face from the face of the abnormalities. If this data is eliminated, holes will exist in the point cloud. Many methods for filling incomplete data, or sparse data, are user intensive [23]. In a method used for reverse

engineering, a point cloud is sliced, and the points are condensed into the center of each slice; each point is weighted based on the distance from the center of the slice third-order b-spline curve [24]. The effectiveness of this method would depend greatly on the complexity of the part's geometry and the quantity of remaining data to reconstruct the surface.

Various methods of removing curvature from surface profiles have been explored. A segmented filter is a simple method to eliminate non-uniform waviness from a surface profile. This method, also known as high-pass filtering, segments the data into equal sample lengths [7]. Alternatively, piecewise splines or polynomials can be fit to the profile using a similar process [18]. Each segment is then fit to straight lines as seen in Figure 2.5. Additionally, polynomial filters can be used to eliminate waviness. For short lengths of data, polynomial curves are fit to the data using least-squares method, seen in Figure 2.6 [7].

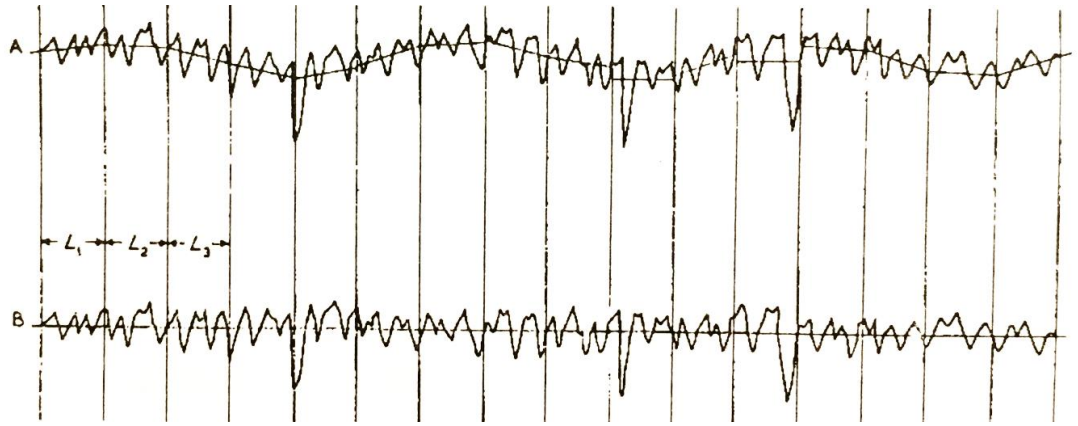


Figure 2.5—Fitting straight lines through segments to remove curvature [7]

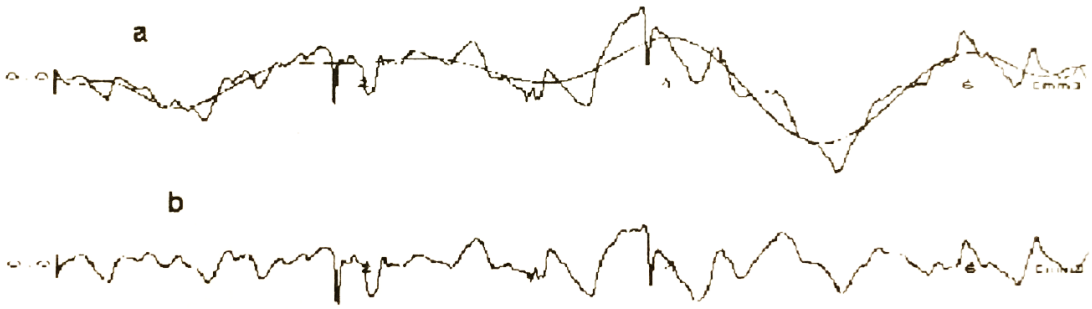


Figure 2.6—Fitting polynomials through a surface to remove curvature [7]

As with any method, there are limits to fitting lines and polynomials to data sets. For instance, the data may not fit as perfectly to the data set as anticipated, as seen in Figure 2.7. If the least-squares method is used to identify the truly best-fit polynomial, the overfitting of data may occur. In order to obtain the best polynomial for the surface, one must know what the profile looks like and fit it to the appropriate order of polynomial, which requires additional time and manual intervention.

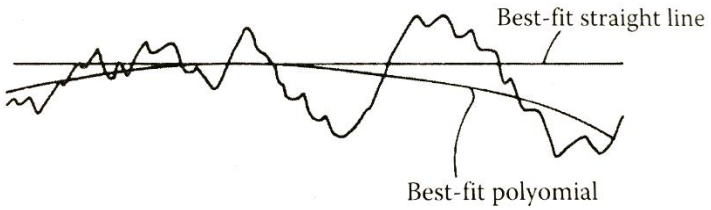


Figure 2.7—Comparison of a best-fit line versus best-fit polynomial for surface profile [25]

Another method to clean data is using a mid-point locus line, as seen in Figure 2.8. In this method, a window of a specified width is moved across the profile and the average height is plotted. Typically, the window should overlap the previous region when it is shifted, however, the shifting distance does not have to be uniform [25].

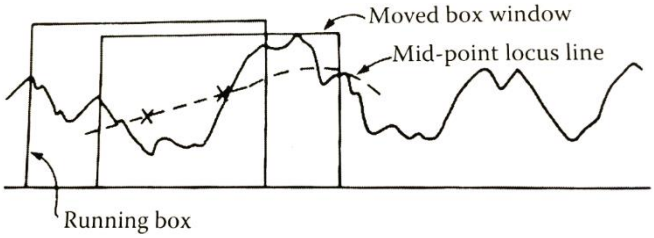


Figure 2.8—Use of the mid-point locus line to remove curvature from surface profile [25]

Gaussian filters are used frequently in surface smoothing applications. They are typically skewed toward large abnormalities; however, if no abnormalities are present, Gaussian filter can give an accurate representation of a data set if the parameters are set appropriately [8]. Much like spline filters, Gaussian filters do not fit data well to the ends of the data sets; therefore, the beginning and end of the data sets are typically fit to straight lines to accommodate for these issues [25]. Some methods replace locally extreme values with the mean line before executing Gaussian filters [19]. This limits the effect outliers and

abnormalities on the final result. However, if a large number of abnormalities exist on one side of the mean line, as they do in honed surfaces, the mean line will be pulled toward the abnormalities and skew the resulting data.

Morphological filters can also be used to decompose surface profiles. A bandpass filter is used to decompose the surface into different bands based on the filter scale compared to the peaks and valleys, as seen in Figure 2.9. By increasing the scale on the filter, smoother profiles are created, which can separate the surface into profiles for curvature, waviness, and roughness [19]. This process linearizes the surface so calculations based on the roughness profile can be executed.

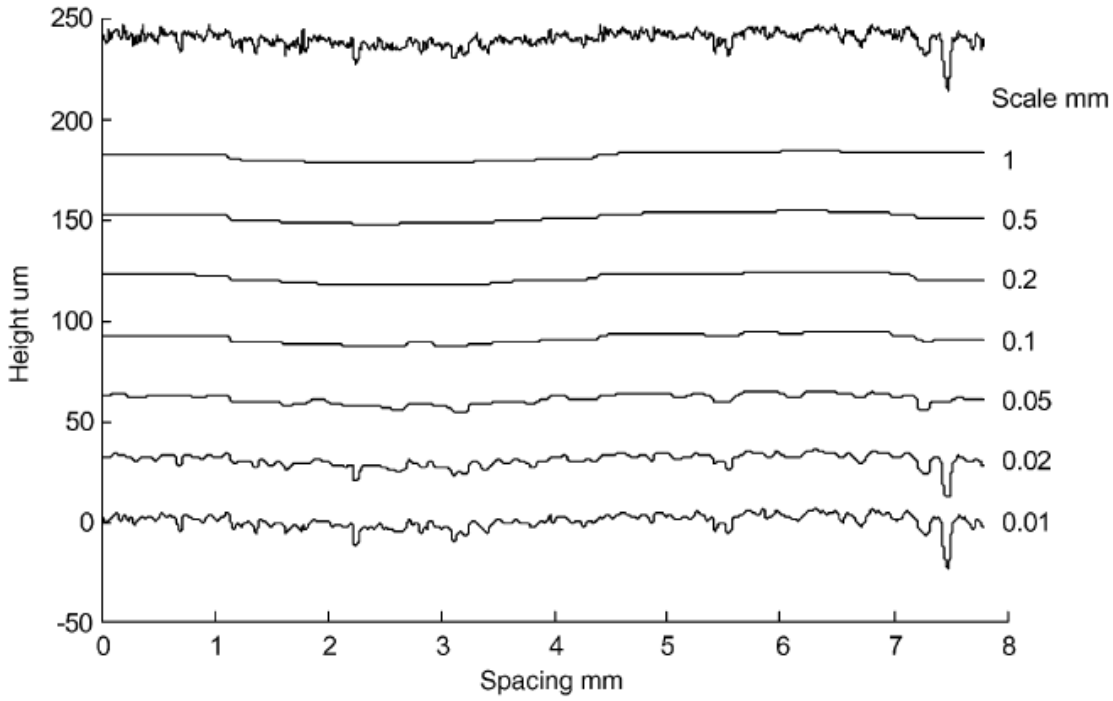


Figure 2.9—Morphological filters decomposing surface profile at various band widths [19]

Robust Gaussian profile filtering use weighted functions in order to smoothen surface profiles. Typically, statistical regression is used in order to weigh the points according to a specific order of polynomial to minimize deviations for smoothing, but with the robust

method, a vertical weighing function is also applied in order to minimize the effects of abnormalities in the profile [8, 26]. This weighting function compares the original mean line to each data point and assigns a weight to each point based on its distance from the mean. A new mean line is constructed using these weighted values, which minimizes the effect of the abnormalities. Unlike traditional Gaussian filters, complex surface profiles can be analyzed since end effects are not as prevalent, and abnormalities have little effect on the resulting profile, as seen in Figure 2.10 [19].

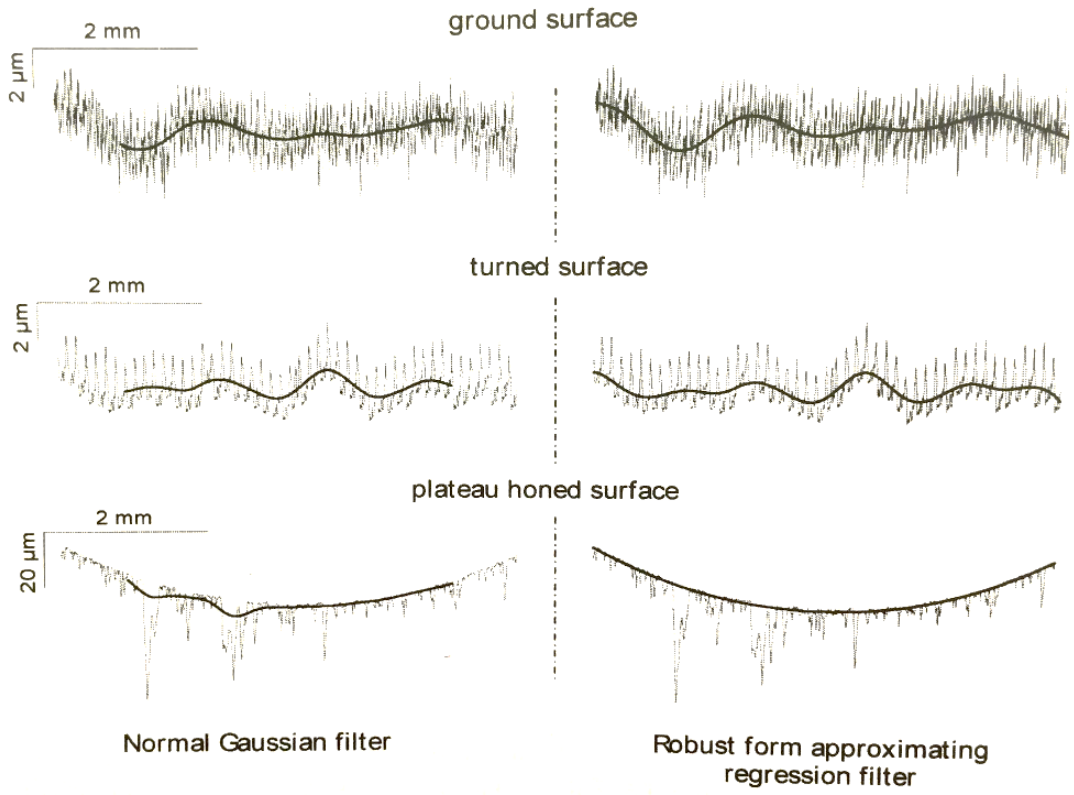


Figure 2.10—Normal Gaussian compared to robust filter on various surface profiles [8]

Manual methods of cleaning and repairing data are often used in point cloud cleaning and reconstruction; however, manual intervention often differs between users. When determining the underlying geometry, the underlying geometry could differ between

individuals interpreting the data, which would essentially nullify the results of any inspection in this manner. For example, if the customer uses one method while the manufacturer uses a different method, the results may differ. This can be avoided by eliminating manual intervention and creating a standard for determining the underlying geometry.

In this section, various methods were described to manipulate point cloud data. In reverse engineering applications, fine details of the scanned surface were desired, and the main goal was to filter out noise caused by the scanning device on mostly smooth surfaces. This proved to be helpful when looking at the original scan data from this casting research, but they failed to remove all roughness from the surface. In the exploration of surface cleaning algorithms, some of the inherent noise from the surface itself could be eliminated; however, these methods were varied on effectiveness based on the data profile. Chapters 5 and 6 will discuss the benefits and drawbacks of some of these methods when determining the underlying geometry.

References

[1] Eubanks, P. E. et al, "Standard for the Visual Inspection of Casting Surfaces," Steel Founders' Society of America Alloy Casting Institute, Des Plaines, IL, 1969.

[2] Manufacturer Standardization Society, "Quality Standard for Steel Castings for Valves, Flanges and Fittings and Other Piping Components (Visual Method)," SP-55, Manufacturers Standardization Society of the Valve and Fittings Industry, New York, NY, 1961.

[3] American Society for Testing and Materials International, "Standard Practice for Steel Castings, Surface Acceptance Standards, Visual Examination," A802/A802M-95, ASTM International, West Conshohocken, PA, 2006.

[4] Editions Techniques des Industries de la Fonderie, "Characterization of Surface Condition of Castings," BNIF 359, National Library of France, Paris, France, 1996.

[5] American Society of Mechanical Engineers, “Surface Texture: Surface Roughness, Waviness and Lay,” ASME B46.1-2009, American Society of Mechanical Engineers, New York, NY, 2009.

[6] Whitehouse, D. J. *Surfaces and Their Measurement*. New York: Taylor & Francis, 2002. 50-67. Print.

[7] Thomas, T. R. *Rough Surfaces*. London: Longman, 1999. 122-74. Print.

[8] Blunt, L., and Xiangqian Jiang. *Advanced Techniques for Assessment Surface Topography: Development of a Basis for 3D Surface Texture Standards “surfstand”* London: Kogan Page Science, 2003. 6-23. Print.

[9] Bračun, Drago, Valter Gruden, and Janez Možina. “A Method for Surface Quality Assessment of Die-castings Based on Laser Triangulation.” *Measurement Science and Technology Meas. Sci. Technol.* 19.4 (2008): Web.

[10] Colosimo, Bianca Maria, Massimo Pacella, and Nicola Senin. “Multisensor Data Fusion via Gaussian Process Models for Dimensional and Geometric Verification.” *Precision Engineering* 40 (2015): 199-213. Web.

[11] Nwaogu, U. C., N. S. Tiedje, and H. N. Hansen. “A Non-contact 3D Method to Characterize the Surface Roughness of Castings.” *Journal of Materials Processing Technology* 213.1 (2013): 59-68. Web.

[12] Poon, Chin Y., and Bharat Bhushan. “Comparison of Surface Roughness Measurements by Stylus Profiler, AFM and Non-contact Optical Profiler.” *Wear* 190.1 (1995): 76-88. Web.

[13] Scheidegger, Carlos E., Shachar Fleishman, and Claudio T. Silva. *Triangulating Point Set Surfaces with Bounded Error*. Proc. of Eurographics Symposium on Geometry Processing, Austrian Academy of Sciences, Vienna. U of Utah, 2005. Print.

[14] Mainsah, E., J. A. Greenwood, and D. G. Chetwynd. *Metrology and Properties of Engineering Surfaces*. Boston: Kluwer Academic, 2001. 124-43. Print.

[15] Schnabel, R., R. Wahl, and R. Klein, “Efficient RANSAC for point-cloud shape detection,” *Comput. Graph. Forum*, Vol. 26 (2007): 214–226.

[16] Bashiri, Mahdi, and Amir Moslemi. “A Robust Moving Average Iterative Weighting Method to Analyze the Effect of Outliers on the Response Surface Design.” *IJIEC International Journal of Industrial Engineering Computations* 2.4 (2011): 851-62. Web.

[17] Vardeman, Stephen, and Analytics Iowa, LLC. “Kernel Smoothing Methods.” Modern Multivariate Statistical Learning. Ames. June 2013. Lecture.

[18] James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: With Applications in R.* New York, NY: Springer, 2013. Print.

[19] Raja, J., B. Muralikrishnan, and Shengyu Fu. "Recent Advances in Separation of Roughness, Waviness and Form." *Precision Engineering* 26.2 (2002): 222-35. Web.

[20] Yoo, Dong-Jin. "Three-dimensional Surface Reconstruction of Human Bone Using a B-spline Based Interpolation Approach." *Computer-Aided Design* 43 (2011): 934-47. *Elsevier*. Web. 1 Jan. 2015.

[21] Zhao, Xiuyang, et al. "IGA-based Point Cloud Fitting Using B-spline Surfaces for Reverse Engineering." *Information Sciences* 245 (2013): 276-89. Web.

[22] "Region Growing Segmentation." *Point Cloud Library (PCL)*. 2014. Web. 18 Jan. 2016.

[23] Yin, Kangxue, et al. "Morfit: Interactive Surface Reconstruction from Incomplete Point Clouds with Curve-driven Topology and Geometry Control." *TOG ACM Trans. Graph. ACM Transactions on Graphics* 33.6 (2014): 1-12. Web.

[24] Liu, Bin, et al. "Triangular Mesh Model Reconstruction from Scan Point Clouds Based on Template." *Tinshhua Sci. Technol. Tsinghua Science and Technology* 14.S1 (2009): 56-61. Web.

[25] Whitehouse, David. "Characterization." *Handbook of Surface and Nanometrology*. 2nd ed. Bristol: Institute of Physics Pub., 2001. 12-21. Print.

[26] Schall, O., A. Belyaev, and H.-P. Seidel. "Robust Filtering of Noisy Scattered Point Data." *Proceedings Eurographics/IEEE VGTC Symposium Point-Based Graphics, 2005* (2005): 71-78. Web.

CHAPTER 3: RISK ASSESSMENT ON VISUAL INSPECTION METHODS FOR CAST METAL SURFACES

A paper submitted to *Quality and Reliability Engineering International*.

Michelle M. Voelker[1], Cameron A. MacKenzie[2], Frank E. Peters[3]

Abstract

Current methods for visual inspection of cast metal surfaces are variable in both terms of repeatability and reproducibility. Because of this variation in the inspection methods, extra grinding is often prescribed; much of this is over processing in attempt to avoid rework or customer rejection. Additionally, defective castings may pass inspection and be delivered to the customer. Surface specifications are often interpreted differently between the customer and manufacturer. A risk assessment employing an influence diagram assesses the probabilities of errors in the inspection process based on different environmental and human factors. The risk assessment determines the probability of Type I and II errors, which can be costly for all parties involved in the production and use of castings. A manufacturer can use this analysis to identify factors in its foundry that could reduce the probability of errors.

I. Introduction

Inspecting parts to meet quality standards is important for meeting customer needs. In metal casting, current standards use qualitative methods to determine acceptability of surface

[1]Primary author; Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa 50011; 0000-0002-3521-694X
[2] Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa 50011; 0000-0002-4851-8509
[3] Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa 50011; 0000-0002-8998-0062

quality. The inspection process involves one or more trained operators to visually examine the surface to determine if the part is acceptable. Variation exists among interpretation of the standard not only in relation to the repeatability and reproducibility of the inspection process, but also in regards to interpretations between the manufacturer and the customer. In fact, the variability in the casting process itself is often less than that of the visual inspection process [1]. This stack-up in variation results in inconsistencies in acceptance criteria and increases the occurrence of Type I and II errors. A Type I error, also known as a false alarm, occurs when a defect is identified on the casting although no defect is present. Type II errors, or misses, occur when a casting passes inspection with a defect present. Although the determination of Type I and II errors is in itself subjective, these errors could be detrimental to the performance of the parts and could lead to problems between the manufacturer and customer if not interpreted as intended.

This paper will combine various sources of uncertainty associated with Type I and II errors, in addition to the consistency of identifying defects, in attempt to model the effectiveness of cast metal surface inspection. This paper develops an influence diagram in order to calculate the probability of a Type I or Type II error. Although influence diagrams have frequently been used to assess risks and identify the optimal alternatives in business and public policy decisions, they have only rarely been applied to manufacturing decisions. This paper is unique because it develops an influence diagram to incorporate and predict the impact of several factors that contribute to Type I and II errors. Management at a manufacturing company can use this type of model to identify factors that would decrease the number of Type I and II errors the most.

Section II presents the current industry standards in visual inspection standards. The influence diagram to determine the likelihood of Type I and II errors and the effects of different interacting factors on them will be outlined in Section III. Finally, Sections IV and V will explore discussions and conclusions based on the constructed model.

II. Current Visual Inspection Standards

Many qualitative standards exist for the surface inspection of cast metal including company and industry specific standards. The Manufacturer Standardization Society (MSS) SP-55 Visual Method, American Society for Testing and Materials (ASTM) Steel Castings Research and Trade Association (SCRATA) A802, Alloy Casting Institute (ACI) Surface Indicator Scale, and GAR Electroforming Cast Comparator C9 are the most commonly used metal casting standards in industry.

MSS SP-55 Visual Method

Images are used for comparison to cast surfaces in the MSS SP-55 method. Twelve abnormality types, ranging from porosity to weld repair areas, are identified and images of acceptable and not acceptable surfaces are provided for each [2]. Figure 3.1 shows an example of the images provided for reference.

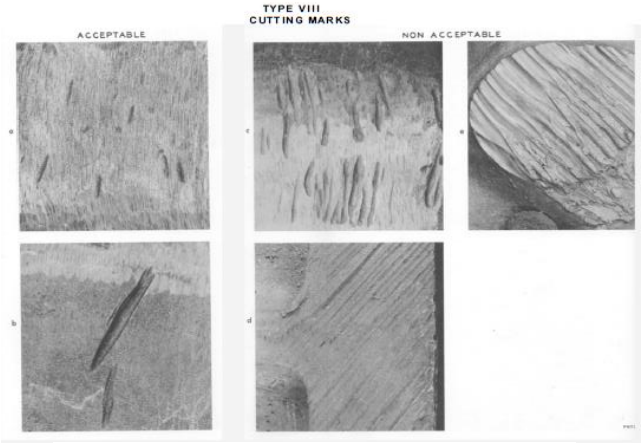


Figure 3.1—MSS Method Example [2]

ASTM A802

Plastic replications of actual metal castings are used for comparison in ASTM A802, more commonly referred to as the SCRATA method. Lettered plates representing one of nine abnormalities are used, each with various severity levels as seen in Table 3.1. The abnormalities represented are similar to the MSS method. This standard is the most widely used standard in the U.S. steel casting industry.

Table 3.1—Visual Inspection Acceptance Criteria [3]

Surface Feature	Level I	Level II	Level III	Level IV
Surface texture	A1	A2	A3	A4
Nonmetallic inclusions	B1	B2	B4	B5
Gas porosity	C2	C1	C3	C4
Fusion discontinuities	…[A]	D1	D2	D5
Expansion discontinuities	…[A]	…[A]	E3	E5
Inserts	…[A]	…[A]	F1	F3
Metal removal marks:				
Thermal	G1	G2	G3	G5
Mechanical	H1	H3	H4	H5
Welds	J1	J2	J3	J5
[A] No reference comparator plate is available for this surface feature and level.				

ACI Surface Indicator Scale

The ACI Surface Indicator, as seen in Figure 3.2, evaluates "general smoothness, height and depth of irregularities extending beyond the range of general variations, and frequency and distribution of such irregularities" [4]. Designations SIS-1 through SIS-4 correspond to the root mean square (RMS) average deviation in micro-inches. The standard also specifies criteria for the height and frequency of surface abnormalities.

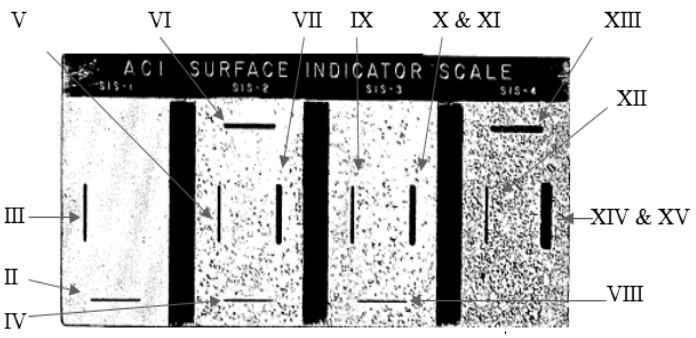


Figure 3.2—ACI Surface Indicator Scale

GAR Microfinish Comparator C9

Less widely used than the other methods, the GAR C9 Comparator is seen in Figure 3.3. Comparator swatches quantify the surface roughness based on root mean square (RMS) values in micro-inches. No abnormalities are defined in this standard. In addition to a visual examination, inspectors are instructed to "draw the tip of the fingernail across each surface at right angles" to match the texture of the inspected part [5].

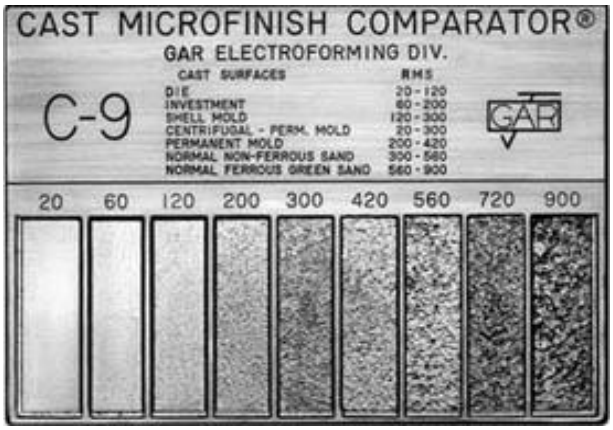


Figure 3.3—C9 Cast Microfinish Comparator [5]

While machine vision is readily applied for some casting surface inspection tasks, it is limited to a range of defects in certain areas. Automation methods are not applicable for the several in process inspection steps of a wide variety of castings within the production facility.

III. Construction of Influence Diagrams

The methods discussed in Section II are used to help determine whether or not a part is defective; however, errors are frequent with these methods. This section builds a model in order to assess the likelihood of Type I and II errors in the visual inspection of cast metal surfaces and the effects of different interacting factors on them.

An influence diagram—also called a Bayesian belief net or a decision diagram—models factors that contribute to a final outcome or uncertainty [6, 7]. Figure 3.4 depicts an influence diagram to calculate the probability of a Type I error and a Type II error. The

diagram is constructed in Netica in order to analyze various scenarios causing errors efficiently. A Type I error (false alarm) occurs when a defect is identified on the casting although no defect is present. A Type II error (miss) occurs when a casting passes inspection with a defect present.

As depicted in Figure 3.4, a manufacturer makes two decisions that influence these probabilities: the training for the inspector and the judgement type used in the inspection process. The manufacturer can determine the judgement to use in the inspection process and training type. Additionally, three uncertain factors directly influence the likelihood of errors: defect density, environmental impact, and human capabilities. The environmental impact depends on the noise, lighting, and work atmosphere. Human capabilities depend on the health and fatigue of the inspectors. After probabilities are assessed for all of the uncertainties, the influence diagram can be solved to calculate the probability of a Type I and II error for each alternative in the training and judgment type decision.

Influence diagrams have been popular modeling tools for analyzing the risks of engineered systems [8, 9], decision making in business and public policy [10, 11, 12], and diagnosing disease [13].

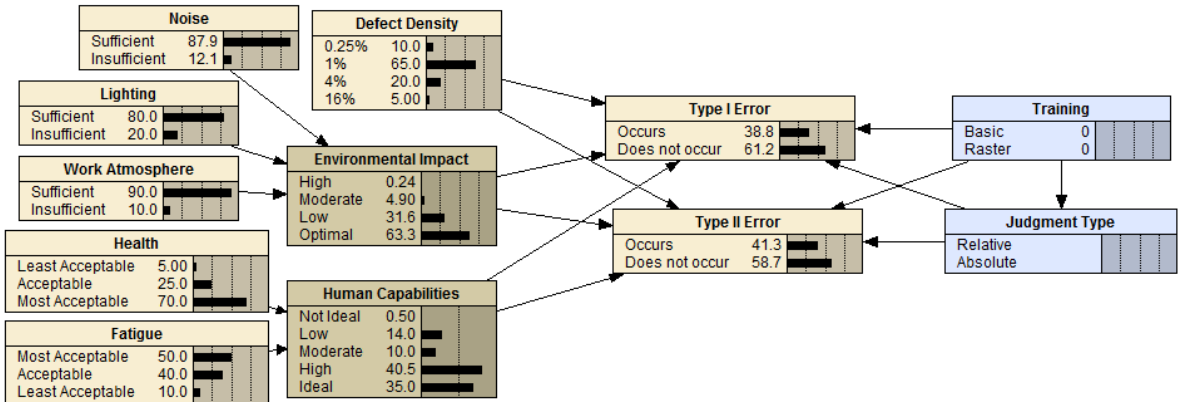


Figure 3.4—Netica model for Type I and II Errors for cast metal surface inspection

Their role in assessing manufacturing problems and uncertainties has been much more limited, however. Some exceptions include diagnosing faults in manufacturing systems [14, 15, 16], optimizing a maintenance policy [17], modeling manufacturing processes with several control variables [18, 19], and determining the optimal site for a manufacturing facility [20, 21].

The remainder of this section analyzes each factor (training and judgement type, environmental factors, human capabilities, and defect density), assesses probabilities for the uncertainties, and explains each factor's impact on Type I and II errors. The probabilities are based on previously conducted experiments and research and the authors' own expertise and knowledge about manufacturing conditions.

Training and Judgement Type

Methodologies used to calibrate inspectors affect the likelihood of Type I and II errors and consistency of identifying defects. This can be attributed to the enforcement of inspection procedures and effectiveness of training.

Enforcing methodologies for inspection is a major factor in the consistency of identifying defects. This consistency helps analyze the reliability of the estimates for our Type I and II errors since the judgment of these errors are, in fact, as subjective as the inspection process. The type of judgment as well as the inspection sampling method impacts how defects are identified.

A manufacturer can choose to enforce a relative or absolute judgement in visual inspection. Relative judgement occurs when the inspector has a comparator or image of the inspection criteria in hand for direct comparison to the cast part, while absolute judgement

occurs when the inspector recalls the criteria from memory. A study to determine the differences in relative versus absolute judgement was conducted in relation to eye-witness accounts [22]. In the relative judgement experiment, participants were asked to compare two individuals and pick which was previously shown in an image. For the absolute judgment experiment, the same participants were shown a single individual and asked if he or she had appeared in the previous image. Accuracy of absolute judgement in the study was found to be 69%, whereas for relative judgment it was found to be 80% as seen in Table 3.2. Although this study did not directly relate to the casting inspection process, these values can be used as insight into the impact of judgement type on Type I and II error. An incorrect identification leads to a Type I or Type II error.

Table 3.2—Judgement type's effects on identification of defects

Judgement Type	Correct ID	Incorrect ID
Absolute	0.69	0.31
Relative	0.80	0.20

In a study evaluating inspection of castings using comparators, data was collected in relation to Type I and II errors. Participants in the study were asked to categorize 25 castings as acceptable or not based on their evaluation of the surface. For some surfaces, participants were given the comparator to use for references (relative), while others were to recall the criteria from memory (absolute). Table 3.3 shows the results of this study [24].

Table 3.3—Judgment type's effects on Type I and II errors

Judgement Type	Type I Error	Type II Error
Absolute	0.33	0.26
Relative	0.22	0.30

Training techniques also can impact error in visual inspection. In one case study, basic training and raster training were evaluated in casting inspection using absolute judgement [24]. Raster training involves teaching inspectors to systematically scan the part in a zig-zag pattern. This study also used eye tracking software to determine the percentage of the casting

viewed under these conditions. Overall, the specific technique used to locate defects not only allowed the individual to view a greater percentage of the part, but it decreased the effects of Type I and II errors in the inspection process. The results of this study are found in Table 3.4; however, it is noted Type II error in raster training was about 16% more variable than for basic training. The inspectors in this study had no prior experience with inspecting castings, which allowed for an unbiased result in the analyzing the overall effectiveness in training [22, 24].

Table 3.4—Training effects on Type I and II errors and percent of part viewed [32]

Training	Type I Error	Type II Error	% Part Viewed
Basic	0.408	0.454	68
Raster	0.264	0.549	75

Environmental Factors

Inspectors can be influenced by various environmental factors including the physical environment and work atmosphere. These aspects can reduce the inspector's effectiveness in the visual inspection process.

The physical work environment includes auditory noise, light level, temperature, and humidity [1]. These can all distract the inspector and even reduce his capability to locate defects. For example, the just noticeable difference between the defect and surrounding area will reduce significantly if the lighting is poor, making the defect more difficult to locate. In general, both Type I and II errors increase in suboptimal conditions [23]. Additionally, the temperature and humidity can affect the inspector's cognitive ability. In fact, the ideal humidity of 65% and temperature of 70 degrees Fahrenheit in the presence of a fan can actually stimulate brain activity and increase alertness of the inspector [24].

The work atmosphere can also affect the inspector's likelihood to locate defects. In some workplaces, workers are rewarded for doing their job well while others are disciplined

if quality is subpar. In some corporations, inspectors are required to re-inspect parts, either from his previous inspections or from another inspector. These are referred to as motivational losses. If an inspector receives a part that has already passed inspection once or knows a part will be inspected later, he may not look as closely for defects because he feels like it is a poor use of time. Both of these instances will increase the likelihood of Type II errors [23].

As depicted in Figure 3.4, the factors of noise, lighting, and work atmosphere were assigned binary states of sufficient or insufficient in the influence diagram. These three factors were chosen based on the estimated impact of each on the inspector. The environmental impact can either be high, moderate, low, or optimal based on the noise, lighting, and work atmosphere. The deterministic states are conservative estimates and their impacts is listed in Table 3.5. For example, if lighting and work environment are considered sufficient but the noise level is insufficient, then the environmental impact is low, and the probability of Type I and II errors will increase by 0.05. The previous studies discussed in Subsection III assumed optimal conditions.

Table 3.5—Deterministic values of environmental impact on Type I and II errors

Deterministic Value	Insufficient States	Impact on Error
High	3	+0.20
Moderate	2	+0.10
Low	1	+0.05
Optimal	0	0.00

The current states of all factors associated with the environmental impact are subjectively estimated based on previous reports and the authors' expertise. Each of the main factors (noise, lighting, and atmosphere) are examined to determine the likelihood that each is in an acceptable or unacceptable state.

The noise element is a major environmental factor in steel foundries. Based on data collected in foundries, the noise level of the processes can range from 70 decibels in areas

further from equipment to well above 85 decibels with some as high as 110 decibels; this not only affects the environment in which they currently work, but it can also affect long term health of the individual [25]. The probability of the noise level in a foundry was modeled using a triangle probability distribution with the minimum, mode, and maximum at 70, 85, and 110 decibels, respectively. Most foundries require their employees to wear at minimum noise reduction rated (NRR) 25dB hearing protection; therefore, the distribution was shifted to the left nine units to account for this practice. To determine the state, current industry standards were used based on industrial safety requirements were used. According to the Occupational Safety and Health Administration, exposure to sound levels above 90 decibels for an eight-hour work day can cause hearing damage, so values above this level were classified at an unacceptable state [26]. Therefore, the probability the noise level is considered insufficient is 12.1% for this model.

Additional lighting at inspection stations is typically installed to increase visibility of the inspector; however, if the light levels become too bright, individuals may experience glare on the surface of the part reducing his ability to effectively inspect the surface. Placement of the casting in the lighting can also play a major role in successfully detecting defects due to shadows that may appear on the surface [27]. Based on a study on casting inspection, the range of lighting seen in inspection stations was from 150 to 15,000 lux with a mean of approximately 675 lux [28]. This was modeled using a beta probability distribution. Ideally, the acceptable range to avoid glare-out and excessive shadows on the part is from 500 to 900 lux. Therefore, light levels outside of this range is considered insufficient, which is 20% for this model.

Larger foundries typically have more than one inspector for each casting process whether it be on the same or different shifts. These foundries are likely to be more at risk for providing rewards to high performing inspectors or creating unintentional competition among the inspectors increasing the likelihood for error. According to a study in the United States, 20% of foundries were considered large businesses, which consisted of 100 or more employees [29]. Since this behavior has not been studied in great detail, we make a conservative assumption that 50% of the large businesses create an insufficient work environment.

Human Capabilities

The capabilities of the individual performing the inspection also play a role in his or her ability to detect defects. These capabilities can be either physical, such as vision ability, or perceptual, such as memory ability.

As shown in Figure 3.4, two uncertainties impact an inspector's capabilities: health and fatigue. Visual, mental, and physical fatigue in inspectors can affect the judgement of whether or not a defect is present. When inspectors are tired, they can lose focus in the task at hand and become easily distracted [23]. Although fatigued inspectors may take additional time to view each part, errors generally increase [24].

The age and health of the inspector can also be a limiting physical capability. This includes vision impairment, such as near or far sightedness, which could reduce the individual's ability to identify defects. Additionally, one's haptic capabilities may also be used to feel whether the surface requirements match the criteria, such as in the GAR C9 comparator. The presence of calluses on the fingertips or loss of feeling in the fingers that

may come with age or prolonged interactions with handling castings may reduce one's ability to detect differences between the casting and the comparator.

The factors of fatigue and health were assigned states in the influence diagram in Figure 4. The states of fatigue and health are least acceptable, acceptable, and most acceptable. These factors were chosen based on the estimated impact of each on the inspector. The impact of the human capabilities are a deterministic function based on the states of each factor: least acceptable (LA), acceptable (A), and most acceptable (MA). The deterministic states and their impact on Type I and II errors are listed in Table 6. Previous studies of Type I and II errors assumed ideal conditions and human capabilities.

Table 3.6—Deterministic values of human capabilities on Type I and II errors

Deterministic Value	States	Impact on Error
Not Ideal	2 LA	+0.20
Low	LA + MA/A	+0.15
Moderate	2 A	+0.05
High	A + MA	+0.02
Ideal	2 MA	0.00

An individual's health and fatigue levels can be impacted by several factors in a foundry environment. These include air quality, heat exposure, and overtime [29, 30]. Individual health related to hereditary, such as vision, were also considered to assess the probability of health and fatigue. Due to a combination of all of these factors, conservative estimates were placed on human capability factors in order to aid in the modeling process.

Defect Density

An inspector's perception of a task can greatly influence the likelihood of finding Type I and II errors. This includes developing a memory of past inspections and expectations over time.

Inspectors who inspect the same part over and over again develop a memory for where defects are most common. This may cause them to overlook other areas of the part to be inspected where defects are less common. In general, Type I errors become less common, and Type II errors increase [23].

Defect density, or the overall number of defects on a part, can affect Type I and II errors. Generally, as the defect density decreases, Type I and II errors increase. For example, if an inspector recalls from previous experience that the number of overall unacceptable parts was approximately one every five parts, he may begin to second guess himself if he finds ten or more in a row without any errors causing a Type II error. Similarly, if he experiences many parts with a lower number of defects, he may look past parts without as many defects causing a Type I error. A study using test samples with 0.25, 1, 4, and 16% defect densities was administered to 80 inspectors with no prior inspection experience. These inspectors were asked to identify all defects on each sample without being told how many defects to expect. If the inspector could not decide whether a specific feature was considered a defect, the test monitor acted as an inspection supervisor and advised them on how to classify the area in question [31]. Results from this study can be found in the Table 3.7.

Table 3.7—Defect density's effect on Type I and II errors

Total Defects	Type I Error	Type II Errors
0.25%	0.85	0.42
1%	0.41	0.29
4%	0.15	0.25
16%	0.05	0.18

Interpretations

Since various standards can be used to inspect cast metal surfaces and there is no easy way to calibrate inspectors, the results from visual inspection are subjective [1]. This can

contribute to Type I and II errors from inconsistencies based on the inspection standard used and interpretation of specifications.

There are many inspection standards for cast metal surfaces. As discussed in Section II, they consist of methods using images while others use physical comparators. Additionally, some of these standards identify specific types of abnormalities to look for when inspecting. This causes several issues. First, if an abnormality is not defined by the standard, there is no way for the customer to specify what is desired. On the other hand, if the customer only specifies criteria for porosity and the part has inclusions, the inspector has to make the decision whether to only inspect for the porosity (what was specified), or if the other abnormalities should also be considered when inspecting the part. In fact, an inspector may not even be able to determine a cause via visual inspection. This causes confusion for both parties.

Additionally, the interpretation of the standard contributes to uncertainty. This includes interpretations between the customer and the manufacturer, among a single inspector, and between multiple inspectors. The consistency of identifying defects has an effect on the overall error; however, the effects on Type I and II errors are not known. Although this variation does not specifically play a role on Type I and II errors, it reveals how consistent defects are identified in the visual inspection process.

Many discrepancies exist between the customer and the manufacturer. At times, the manufacturer can complete inspection, but this inspection may not meet the customer's standards. For both parties, time constraints can play a big role in the effectiveness of inspection. For example, if the manufacturer is behind schedule and needs to deliver parts, he may be more likely to ship out bad parts in order to be on time. The customer may reject all

of the bad parts; however, if the customer is also being pushed to deliver its product, it may be more lenient as to what surface quality is acceptable in order to meet its deadlines. These situations have happened before in industry, but studies are not available to quantify these values due to the variability and confidentiality associated with this factor.

Issues with repeatability (variation for a single inspector) and reproducibility (variation between inspectors) may also arise within a company's inspecting team, which affects the consistency of identifying defects. Visual inspection methods show large variation in measurement error for both repeatability and reproducibility due to inconsistencies for a single inspector between parts and between inspectors on the same part [33]. From this study, the average repeatability across six operators from three foundries was 66.83%, while the average reproducibility for operators at the same facilities was 63.33% as seen in Table 3.8 [34]. Figure 3.5 shows an example of the variation of defect detection on the same part for two operators. Although improving the inspectors' ability to interpret the standard consistently would inherently reduce the overall error, its effects on the individual types of errors are unknown. Since these factors do not directly contribute to Type I and II errors, they were not included in the influence diagram.

Table 3.8—Repeatability and reproducibility

	Same Results	
	Yes	**No**
Repeatability	0.668	0.332
Reproducibility	0.633	0.367

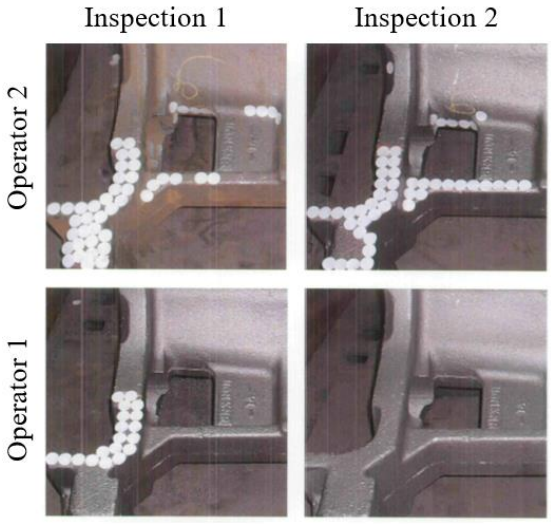


Figure 3.5—Repeatability and reproducibility example where operators marked defects with white stickers [34]

Since the consistency of an inspection requires that the inspection is both repeatable and reproducible, consistency can be calculated as the product of the probability of repeatability and reproducibility. Judgement type and the inspection method will also impact the consistency of evaluation.

IV. Discussion

Populating the influence diagram in Figure 3.4 requires combining data from different sources in order to form a better overall risk assessment. Since each dataset that relates judgement type (Tables 3.2 and 3.3), training (Table 3.4), or defect density (Table 3.7) to Type I and II errors does not consider the other two elements, an average of the three probabilities are used to determine the probability of an error conditioned on the judgement, training, and defect density. It is also assumed that all studies were conducted under optimal conditions for environmental conditions and the ideal state for human capabilities. Thus, these factors are additive to the overall probabilities of the other factors. The factors in the influence diagram that influence Type I and II errors are not exhaustive; however, they do play a major role on casting inspection. Megaw [27] provides an extensive list of sources that affect the accuracy of visual inspection.

Although prior studies provide estimates of the probabilities for each node in the influence diagram, actual values will vary among individual foundries. These values, in addition to the experimental values from previous studies, may be substituted in the model in order to get an accurate assessment for an individual foundry.

These qualitative standards for cast metal surfaces rely on an individual's capability to judge whether or not a part is acceptable. Individuals must differentiate between the types of abnormality present. It can be unclear if a part is acceptable when an unexpected

abnormality appears on the final part if that abnormality was not taken into consideration by the customer when specifying the surface. Additionally, the interpretation of the standard or specification varies greatly among inspectors and between the customer and manufacturer. This factors increase the risk of Type I and II errors resulting from the inspection process.

Given these assumptions and data, the influence diagram in Figure 3.4 is solved using Netica software to evaluate the impact of each training and judgement decision on Type I and II errors.

Figure 3.6 depicts the probability of Type I and II errors given each alternative for judgement and training type. These probabilities are calculated from the influence diagram in Figure 3.4 and based on the model assumption for the potential state of any given foundry. As seen in the figure, relative judgement and raster training tend to decrease the probability of Type I errors; however, the opposite is true for Type II errors. This is an interesting observation since it appears more robust training and judgement types (raster and relative) decrease the probability of false alarms and increase the probability of misses. This is most likely a result from the high variability in Type II error from raster training used in the assessment of probabilities [32].

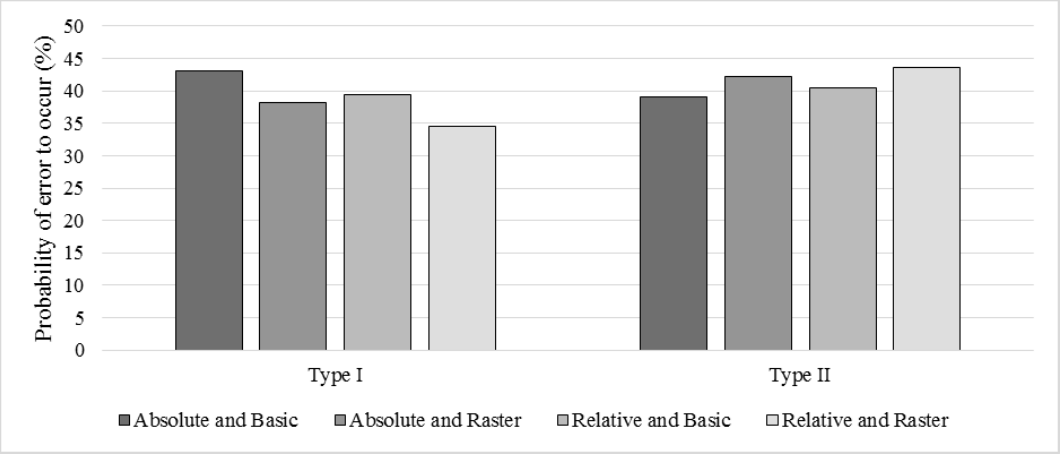


Figure 3.6—Base values of error comparing judgment and training type decision without certainty of other factors

Since many of the probabilities in the model are based on assumptions of how the different factors interact with each other, sensitivity analyses can help determine to what extent the probabilities for Type I and II errors depend on these assumptions. Figure 3.7 shows the sensitivity analysis where the percent change of the Type I and II errors is displayed based on a change from the original state to the worse state of each factor. This change in probability would be added to the base values from the decisions made for judgement and training type seen in Figure 3.6. For example, if noise is at an insufficient state, the probability of a Type I and II error increases by 5%.

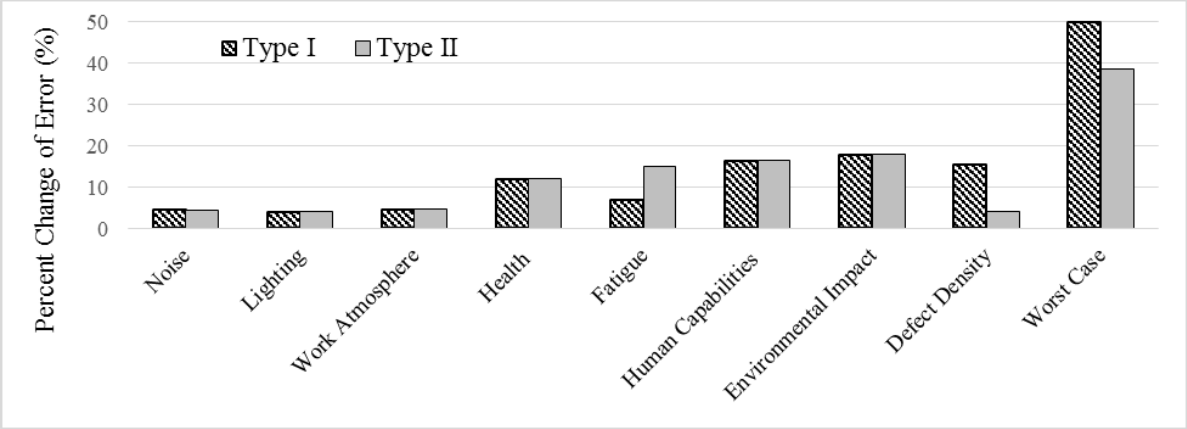


Figure 3.7—Sensitivity analysis of various factors' worst case on Type I and II errors

The worst case, as seen in Figure 3.7, includes a high environmental impact, not ideal human capabilities, and a 0.25% defect density. Such a situation increases the likelihood of a Type I error by 49.9% and the likelihood of a Type II error by 38.7%. Thus, if the basic training and absolute judgement are used with this worst case scenario, the probability of a Type I error is 92.9%. If raster training and relative judgement are used, the probability of a Type II error is 82.3%. If conditions are optimal, the defect density is 16%, the environmental impact is optimal, and human capabilities is ideal. Under this best-case scenario, the

probability of Type I error is 17.8% with relative judgement and raster training, and the probability of Type II error is 29.8% with absolute judgement and basic training. Although these are extreme differences, the probabilities of Type I and II errors are still high for even the best case scenario. In the most likely scenario, with an average defect density of 1%, a low environmental impact, and high human capabilities, Type I errors occur 40.5% of the time and Type II errors occur 43.6% with raster training and absolute judgment.

From the sensitivity analysis in Figure 3.7, the environmental impact and human capabilities have the largest impact on the Type I and II errors; however, each factor contributing to the environmental impact (noise, lighting, and work atmosphere) has little individual effect on the overall outcome. Additionally, defect density appears to impact Type I errors more than Type II errors, and fatigue has the opposite effect. In the case of defect density, when fewer defects are present, inspectors have fewer defects to identify, which increases their tendency to over inspect parts and cause false alarms. When inspectors are fatigued, their attention is less focused, resulting in a tendency to miss defects. Targeting areas like fatigue and defect density would be ideal if a manufacturer wants to reduce one type of effect; this could include requiring visual exercises to reduce eye strain or increasing awareness of defect density among inspectors.

V. Conclusions

Surface standards for metal cast surfaces help to determine the acceptability of surface quality; however, with current standards and capabilities, a large amount of variability exists in the visual inspection process. This paper represents the first use of an influence diagram to model the inspection process of surface capabilities. The influence diagram models and demonstrates how the different factors interact to impact Type I and II errors. The

probabilities in the influence diagram are derived from previous studies and the authors' expertise. According to the model, Type I errors appear slightly less frequently than Type II errors. However, each type of error must be examined independently of one another to understand the impact. In the case of a Type I error, parts are leaving the manufacturer and arriving at the customer in an unacceptable condition. If the customer does not do an in-house inspection before using the parts, they could be assembled into final products and could damage the customer's reputation to the consumer. In the case of a Type II error, acceptable parts are being held at the manufacturer unnecessarily causing an increase in work-in-process inventory and adding additional labor for rework and re-inspection. If multiple inspectors arrive at this same conclusion, the parts may even be scrapped.

The influence diagram developed in this assessment provides additional insight into the visual inspection process. The model of individual factors and their interactions with one another present a broader picture of the problem. Using Netica allowed for a simple means of comparing scenarios when uncertainty nodes changed state or decision nodes were declared. This provides a better understanding of how a variety of factors plays a role in affecting Type I and II errors.

The consistency of identifying defects, however, is extremely variable, which means the estimates for Type I and II errors contain a significant amount of variability. The judgment of these errors are as subjective as the inspection process. Clearer communication of expectations of cast surface specifications is needed between the manufacturer and customer. In order to improve communication in visual inspection, the manufacturer and customer should convene to discuss their expectations of surface quality in regards to the comparator methods available. Additionally, training procedures should be developed so that inspectors

are calibrated with one another. A yearly refresher course, at minimum, would be ideal in order to verify the inspectors remain calibrated throughout the duration of their inspection duties.

To reduce the subjectivity and variability of visual inspection, quantitative criteria should be implemented. A digital surface standard can be developed in order to provide a quantitative method of inspecting cast metal surfaces. This standard should reduce the variation and improve the accuracy in the surface inspection process. The influence diagram could be expanded to assess how the probabilities of errors change with such a standard.

VI. Acknowledgements

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-033. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

VII. References

[1] Monroe, Raymond. “Design “As Is” and “To Be”” Steel Founder's Society of America Meeting. July 2015.

[2] MSS. Quality Standard for Steel Castings for Valves, Flanges and Fittings and Other Piping Components (Visual Method). Tech. no. SP-55. New York: Manufacturers Standardization Society of the Valve and Fittings Industry, 1961.

[3] ASTM International. Standard Practice for Steel Castings, Surface Acceptance Standards, Visual Examination. Tech. no. A802/A802M-95. West Conshohocken: ASTM Int'l, 2006.

[4] Eubanks, P. E. et al. Standard for the Visual Inspection of Casting Surfaces. Manual. Des Plaines: Steel Founders' Society of America Alloy Casting Institute, 1969.

[5] ASME. Surface Texture: Surface Roughness, Waviness and Lay. New York: American Society of Mechanical Engineers, 2009. Print. ASME B46.1-2009.

[6] Howard, R.A. and J.E. Matheson, *Influence Diagrams in The Principles and Applications of Decision Analysis*, R.A. Howard and J.E. Matheson, Editors. 1984, Strategic Decisions Group: Menlo Park, CA.

[7] Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Reasoning*. 1988, Morgan Kaufmann Publishers, Los Altos.

[8] Bedford, T. and R. Cooke, *Probabilistic Risk Analysis: Foundations and Methods*. 2001, Cambridge: Cambridge University Press.

[9] Trucco, P., E. Cagno, F. Ruggeri and O. Grande, "A Bayesian Belief Network Modelling of Organisational Factors in Risk Analysis: A Case Study in Maritime Transportation." *Reliability Engineering and System Safety*, 93 (2008): 823-834.

[10] Varis, O. "Bayesian Decision Analysis for Environmental and Resource Management." *Environmental Modelling & Software* 12.2-3 (1997): 177-185.

[11] Neil, M., Fenton, N., Forey, S., & Harris, R. (2001). "Using Bayesian belief networks to predict the reliability of military vehicles." *Computing & Control Engineering Journal*, 12(1), 11-20.

[12] Paté-Cornell, E. and S. Guikema (2002). "Probabilistic Modeling of Terrorist Threats: A Systems Analysis Approach to Setting Priorities among Countermeasures." *Military Operations Research* 7.4 (2002): 5-23.

[13] Owens, D. K., R. D. Schachter and R. F. Nease. "Representation and Analysis of Medical Decision Problems with Influence Diagrams." *Medical Decision Making* 17.3 (1997): 241-262.

[14] Lewis, R. W. and R. S. Ransing. "A Semantically Constrained Bayesian Network for Manufacturing Diagnosis." *International Journal of Production Research* 35.8 (1997): 2171-2188.

[15] Mehranbod, N., M. Soroush, M. Piovoso and B. A. Ogunnaike. "Probabilistic Model for Sensor Fault Detection and Identification." *AIChE Journal* 49.7 (2003): 1787-1802.

[16] Jeong, I.-J., V. J. Leon and J. R. Villalobos. "Integrated Decision-Support System for Diagnosis, Maintenance Planning, and Scheduling of Manufacturing Systems." *International Journal of Production Research* 45.2 (2007): 267-285.

[17] Vatn, Jørn, Per Hokstad, and Lars Bodsberg. “An Overall Model for Maintenance Optimization.” *Reliability Engineering & System Safety* 51.3 (1996): 241-57.

[18] Agogino, A. M., S. Srinivas and K. M. Schneider. “Multiple Sensor Expert System for Diagnostic Reasoning, Monitoring and Control of Mechanical Systems.” *Mechanical Systems and Signal Processing* 2.2 (1988): 165-185.

[19] Nadi, F., A. M. Agogino and D. A. Hodges. “Use of Influence Diagrams and Neural Networks in Modeling Semiconductor Manufacturing Processes.” *IEEE Transactions on Semiconductor Manufacturing* 4.1 (1991): 52-58.

[20] Canbolat, Y. B., K. Chelst and N. Garg. “Combining Decision Tree and MAUT for Selecting a Country for a Global Manufacturing Facility.” *Omega* 35.3 (2007): 312-325.

[21] Dogan, I. “Analysis of facility location model using Bayesian Networks.” *Expert systems with applications* 39.1 (2012): 1092-1104.

[22] Weber, Nathan, and Neil Brewer. “Confidence-Accuracy Calibration in Absolute and Relative Face Recognition Judgments.” *Journal of Experimental Psychology*: Applied 10.3 (2004): 156-72.

[23] Schorn, T. J. “Management Risks in Multiple 100% Inspections.” American Foundry Society (2013).

[24] Peters, Frank, Richard Stone, et al. “Visual Inspection of Casting Surfaces.” *AFS Transactions* (2013): 45-52.

[25] National Occupational Health and Safety Commission. *Foundry Health Hazards.* Canberra: AGPS, 1989. Safe Work Austrailia. Web. 10 Nov. 2015.

[26] Occupational Safety & Health Administration. “Regulations and Standards.” U.S. Department of Labor. Web. 10 Nov. 2015.

[27] Megaw, E. D. “Factors Affecting Visual Inspection Accuracy.” *Applied Ergonomics* 10.1 (1979): 27-32. Web. 11 Nov. 2015.

[28] Stone, Richard. “Environmental and Social Factors.” IE 577- Human Factors Lecture. Howe Hall, Iowa State University, Ames. 16 Nov. 2015. Lecture.

[29] Spada, Alfred. “Revitalization of American Metalcasting.” American Foundry Society. Minnesota, Twin Cities. 17 Feb. 2015. Lecture.

[30] Occupational Safety & Health Administration, United Food and Commercial Workers International Union. *Hazards of Extreme Temperatures: Handling Heat and Cold*. 2015. Web. 11 Nov. 2015.

[31] Harris, Douglas H. “Effect of Defect Rate on Inspection Accuracy.” *Journal of Applied Psychology* 52.5 (1968): 377-79. Web. 15 Oct. 2015.

[32] Watts, K. *The effect of visual search strategy and overlays on visual inspection of castings.* M.S. Thesis, Iowa State University, 2011.

[33] Daricilar, G. *Measurement Error of Visual Casting Surface Inspection*. M.S. Thesis, Iowa State University, 2005.

[34] Daricilar, G., and F. Peters. “Methodology for Assessing Measurement Error for Casting Surface Inspection.” *International Journal of Metalcasting* (2011): 7-15.

CHAPTER 4: DEVELOPMENT OF A DIGITAL STANDARD TO SPECIFY SURFACE REQUIREMENTS OF CAST METAL SURFACES

A paper accepted for inclusion in the Special Issue on Surface Texturing for *ASTM Materials Performance and Characterization*[4].

Michelle M. Voelker[5] and Frank E. Peters[6]

Abstract

Communication of specifications between a customer and a manufacturer is important for meeting form, fit, and functional requirements of any part. Current standards for the requirements of cast metal surfaces use qualitative methods, including comparator plates and images of surfaces, to specify the surface quality allowing ample room for variation in interpretation of the standard. The length scale of existing contact surface measurements is too small for most casting surfaces. This paper covers a proposed digital standard for specifying cast metal surfaces. The proposed digital standard uses point cloud data of a cast surface, likely attained using a non-contact capture method, in order to identify roughness properties and anomalies caused by the casting process. Unlike current qualitative methods, this standard does not specify the potential causes of surface issues, such as porosity or inclusions. This standard has been developed in order to reduce measurement variation and eliminate confusion between the customer and manufacturer. Assigning quantitative criterion to the surface allows the customer to specify exactly what is needed as opposed to limiting them to a subjective comparator or image to base their requirements. Additionally, this

---

[4] Reprinted with permission

[5] Primary author; Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa, 50011, United States; 0000-0002-3521-694X

[6] Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa, 50011, United States; 0000-0002-8998-0062

quantitative method can be used to verify visual inspection results among the inspectors within a production facility to reduce their measurement error and improve productivity.

I. Introduction

Inspecting parts to meet quality standards is important for meeting customer needs. In metal casting, current standards use qualitative methods to determine acceptability of surface quality. These methods show large variation in measurement error for both repeatability and reproducibility due to in the inconsistencies subjective decision making for a single inspector between parts and between inspectors on the same part [1]. The proposed digital standard that quantifies acceptance criteria is being developed to reduce the amount of error during inspection to verify results from a visual method. For the customer, a quantitative, or digital, standard will allow them to be able to communicate to the manufacturer exactly what they need or want. It does not limit the customer to a specific set of surface finishes like other standards that use a set of comparators or images to specify requirements. The development of the standard for Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces (Appendix A) is discussed in this article.

II. Current Inspection Standards

The Alloy Casting Institute (ACI) Surface Indicator Scale, Manufacturer Standardization Society (MSS) SP-55 Visual Method, American Society for Testing and Materials (ASTM) A802-95 that reference the Steel Castings Research and Trade Association (SCRATA) comparator plates and its French equivalent, BNIF 359, continue to be the leading

standards used to specify metal casting surfaces. In addition, the GAR Electroforming Cast Comparator C9 is used in some surface roughness inspection processes.

ACI Surface Indicator Scale

The ACI Surface Indicator method uses a metal plate with four surface variations, as seen in Figure 4.1. The method evaluates "general smoothness, height and depth of irregularities extending beyond the range of general variations, and frequency and distribution of such irregularities [2]." The comparator swatches are designated SIS-1 through SIS-4 and correspond to the root mean square (RMS) average deviation in micro-inches. Additionally, the standard specifies criteria for the height and frequency of surface abnormalities through a series of grids of a "controlling square inch."

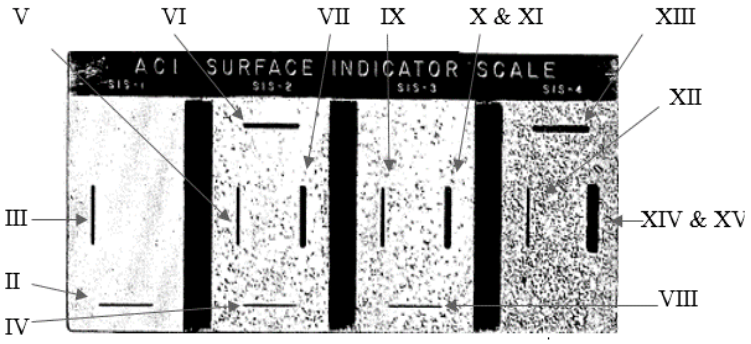


Figure 4.1—ACI Surface Indicator Scale [2]

MSS SP-55 Visual Method

The MSS SP-55 method uses images as a means to specify surfaces. Twelve different types of abnormalities ranging from porosity to weld repair areas are pictured with examples of both acceptable and non-acceptable cast surfaces [3]. An example of the standard is shown in Figure 4.2.

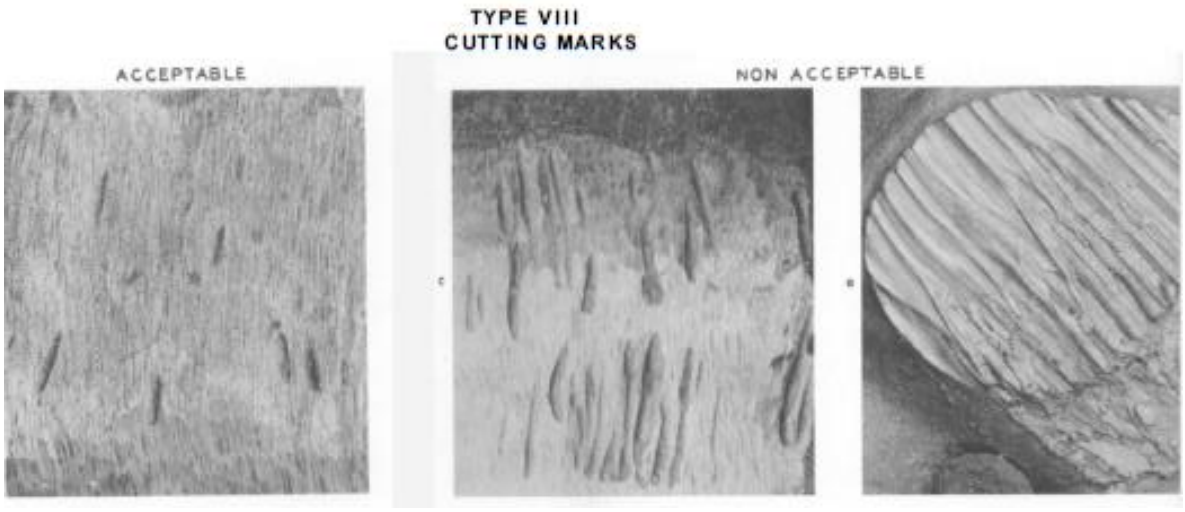


Figure 4.2—MSS method example of acceptable (left) and non-acceptable (right) cutting marks [3]

ASTM A802-95

The SCRATA method uses plastic plates replicated from actual steel casting surfaces for comparison to the finished part. Nine different abnormalities are represented by lettered plates, each with either two or four levels of severity of the abnormality labeled Level I to Level IV as seen in Table 4.1. The roughness nor abnormalities are quantified. These abnormalities are similar to the MSS method with a slight variation in how they are grouped. This method is most commonly used in the U.S. steel casting industry.

Table 4.1—Visual inspection acceptance criteria of ASTM A802 [4]

Surface Feature	Level I	Level II	Level III	Level IV
Surface texture	A1	A2	A3	A4
Nonmetallic inclusions	B1	B2	B4	B5
Gas porosity	C2	C1	C3	C4
Fusion discontinuities	… [A]	D1	D2	D5
Expansion discontinuities	… [A]	… [A]	E3	E5
Inserts	… [A]	… [A]	F1	F3
Metal removal marks:				
Thermal	G1	G2	G3	G5
Mechanical	H1	H3	H4	H5
Welds	J1	J2	J3	J5

[A] No reference comparator plate is available for this surface feature and level.

BNIF 359

The BNIF method is a French standard similar to the SCRATA method in that it uses plastic replicas of cast metal surfaces. A comparison of these comparators can be seen in Figure 4.3. Each comparator is an example of a specific casting process and is classified by the type and amount of finishing required. The three finishing classifications consist of the following: Series n°1: No or limited finishing, Series n°2: Particular finishing, and Series n°3: Special finishing. Suggested values for steel, iron, aluminum and copper are given based on the molding process. A general scale of the roughness average is provided as a general guideline for each suggested process as seen in Figure 4.4 [5].

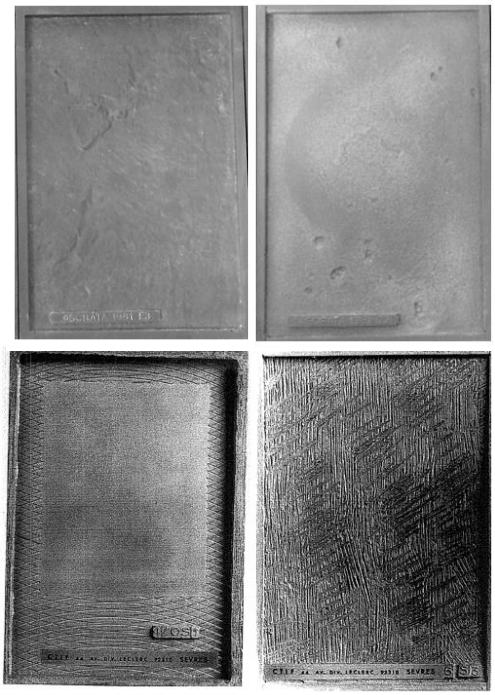


Figure 4.3—Comparison of SCRATA (top-E3, C3) and BNIF comparators (bottom- 4 OS1, S3) [4, 5]

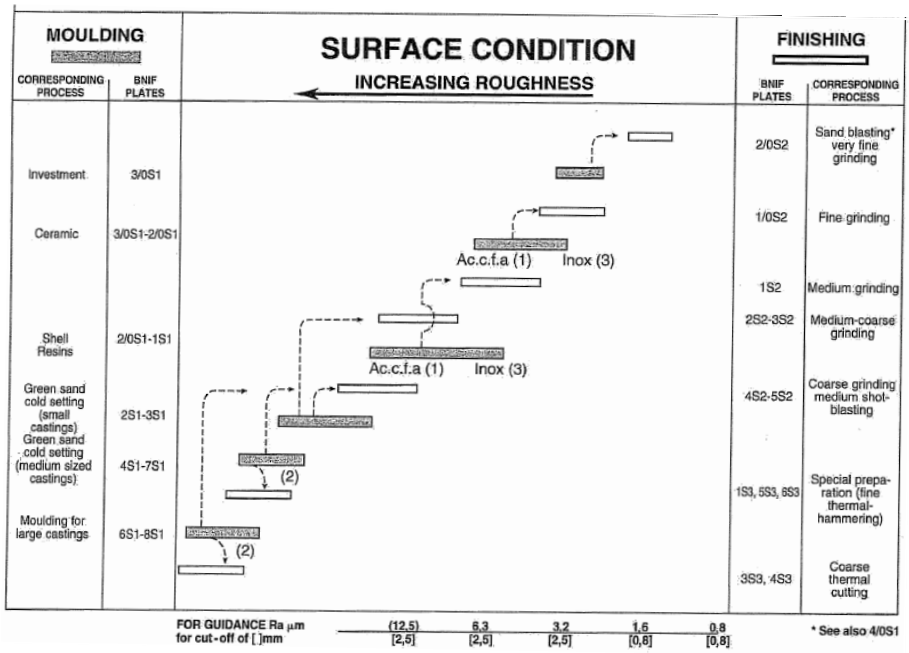


Figure 4.4—BNIF suggestion table for steel castings [5]

GAR Electroforming Cast Microfinish Comparator C9

The GAR C9 Comparator, seen in Figure 4.5, is not as widely used as the aforementioned methods. Each comparator swatch represents the surface texture based on root mean square (RMS) values in micro-inches. This standard provides additional clarity compared to the ACI Surface Indicator Scale, MSS SP-55, and ASTM A805-92 for interpretation of the standard; however, it does not define any abnormalities. In addition, inspectors use this comparator qualitatively with little regard for the measurement assignment. Instructed use of this comparator includes "drawing the tip of the fingernail across each surface at right angles" to match the texture of the inspected part [6].

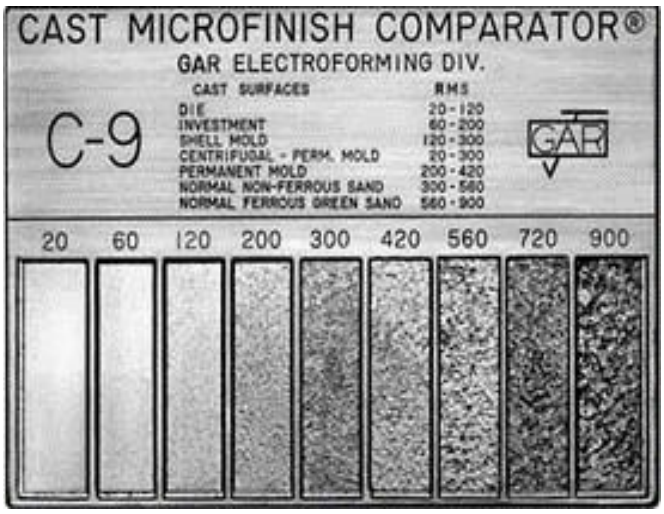


Figure 4.5—C9 Microfinish Comparator [6]

Other

These standards for metal casting specification and inspection have several disadvantages. These disadvantages include the need for subjective interpretation of the standard, expectations of labor, definition of abnormalities, and distribution of abnormalities.

Summary of Current Standards

These standards for metal casting inspection have several disadvantages. These disadvantages include interpretation of the standards, expectations of labor, definition of abnormalities, and distribution of abnormalities.

Standard Interpretation

Variation exists between the manufacturer's and customer's interpretation of the standard due to the complexity of the evaluation criteria and variation in qualitative inspection. A definitive cut off point in which the part can be deemed as acceptable currently does not exist or is unclear in the written standards.

Labor Expectations

Personnel must be trained on the standard and should have the standard documentation in hand in order to make the determination of whether or not the part is acceptable. These methods rely solely on the individual's sensory (visual and possibly tactile) capability as opposed to hard data. Due to the subjectivity of the decision, the cutoff point can move out over time or among people. Research has shown that training must be ongoing to keep personnel 'calibrated' [9].

Undefined Abnormalities

Surface abnormalities not contained within the given standard make it difficult to assign a value to the finished part. Furthermore, many abnormalities cannot be determined via visual inspection and rather require metallurgical analysis. Furthermore, the origin of the abnormality is quite irrelevant to the final casting use in most cases.

Abnormality Distribution

The distribution of abnormalities versus size over the entire part is not clearly specified. For example, if one large crater is acceptable on a part, there is no reasoning behind why multiple craters of smaller size are not acceptable. Or, if the area under question is smaller than a SCRATA comparator plate, the single larger crater could now not be acceptable.

With the decreasing cost of non-contact technologies, such as white light and laser scanning, a quantitative method can be introduced to increase reliability and repeatability of the casting inspection process.

III. Overview of Quantitative Standard

The quantitative standard uses data obtained from three-dimensional scans of a portion of a casting in order to objectively inspect a surface. From this data, the three main parameters specified by the customer are verified, including the baseline roughness, abnormality level, and abnormality percentage.

The baseline roughness, measured in millimeters, is the roughness average, denoted S_a for areal roughness or R_a for a profile, of the cast surface disregarding abnormalities. This parameter is the minimum requirement to be specified by the customer. Default values will be assigned to other parameters if none are specified.

Abnormalities are any surface anomaly present that is not part of random variation due to the actual baseline roughness and are greater than, arbitrarily, twice the specified baseline roughness. Therefore, there is no need for the customer to specify every type of abnormality that could possibly occur, as with the SCRATA standard; all abnormality types are encompassed under the abnormality level parameter. These include, but are not limited to,

porosity, inclusions, and expansion. Abnormalities are considered any point exceeding twice the specified baseline roughness. The abnormality level is specified in millimeters and is represented by the absolute distance of the data point from the underlying geometry. If an abnormality level is not specified, the default level assigned where no abnormalities are acceptable, or twice the specified baseline roughness. (As discussed later, the designer could specify a surface with no allowable abnormalities; however, this could come at a higher acquisition cost.)

The third parameter to describe the surface is the abnormality percentage. This is expressed as the total fraction of the surface area that is considered abnormal, or exceeding twice the specified baseline roughness. The default inspection area is 8 centimeters by 8 centimeters, arbitrarily, unless otherwise agreed upon by the customer. The abnormality area is a percentage of this target area. The target area can be any 8 by 8 centimeter area on the surface, meaning every such area needs to be in specification. This prevents discrepancies between the customer and manufacturer when interpreting the abnormality percentage. If an abnormality percentage is not specified, the default level assigned will be 5%. This standard does not cover dimensional accuracy, unusual visual conditions, such as casting color, nor chaplets. Chaplets are not included in this specification because they represent a likely performance issue, unlike most other abnormalities on the casting surface.

These three parameters should be specified at their maximum acceptable value for use and annotated using the Voelker Surface Ratio (VSR), which is written numerically with dashes as, “VSR [baseline roughness] – [abnormality level] – [abnormality percentage].” An example of this notation is, “VSR 0.30 – 0.60 – 2,” indicating a maximum baseline roughness of 0.30 mm, a maximum abnormality level of 0.60 mm, and the maximum percentage of the

inspected surface considered abnormal of 2%. If the standard only specifies “VSR 0.30,” the defaults for abnormality level and abnormality percentage are assigned as twice the specified baseline roughness, or 0.60, and 5, respectively for any 8 by 8 centimeter area on the casting.

In order to consistently calculate these parameters due to the complexity of cast surfaces, the underlying geometry must be determined. The underlying geometry is the geometry of the surface in absence of the surface roughness and abnormalities. This geometry may differ from the intended part geometry due to contraction, mold movement, and other dimensional changes during the casting process. To illustrate the use of the proposed standard, the process of finding the underlying geometry to calculate surface deviations and identifying abnormalities for a criteria of VSR 1.85 – 12.00 – 35 is found in Figure 4.6. After a surface is scanned and the underlying geometry is determined, the deviations from each point to the underlying geometry are calculated. Based off of the acceptance criteria from the customer and deviations from the underlying geometry, the actual baseline roughness is calculated, and abnormalities are identified and measured.

A single surface can be specified in different ways. The sample profile in the previous example shows a surface with an abnormality located in the center. For the purpose of simplifying conceptualization, the total number of abnormal points in the two-dimensional profile divided by the total number of points in the profile will be used to illustrate the abnormality percentage. Given this assumption, the profile could be classified as the following variations: VSR 1.85 – 12.00 – 35, VSR 2.32 – 12.00 – 17, and VSR 6.00 – 12.00 – 0. The bounds of each variation where the data points falling outside of the bounds are considered abnormal are shown in Figure 4.7.

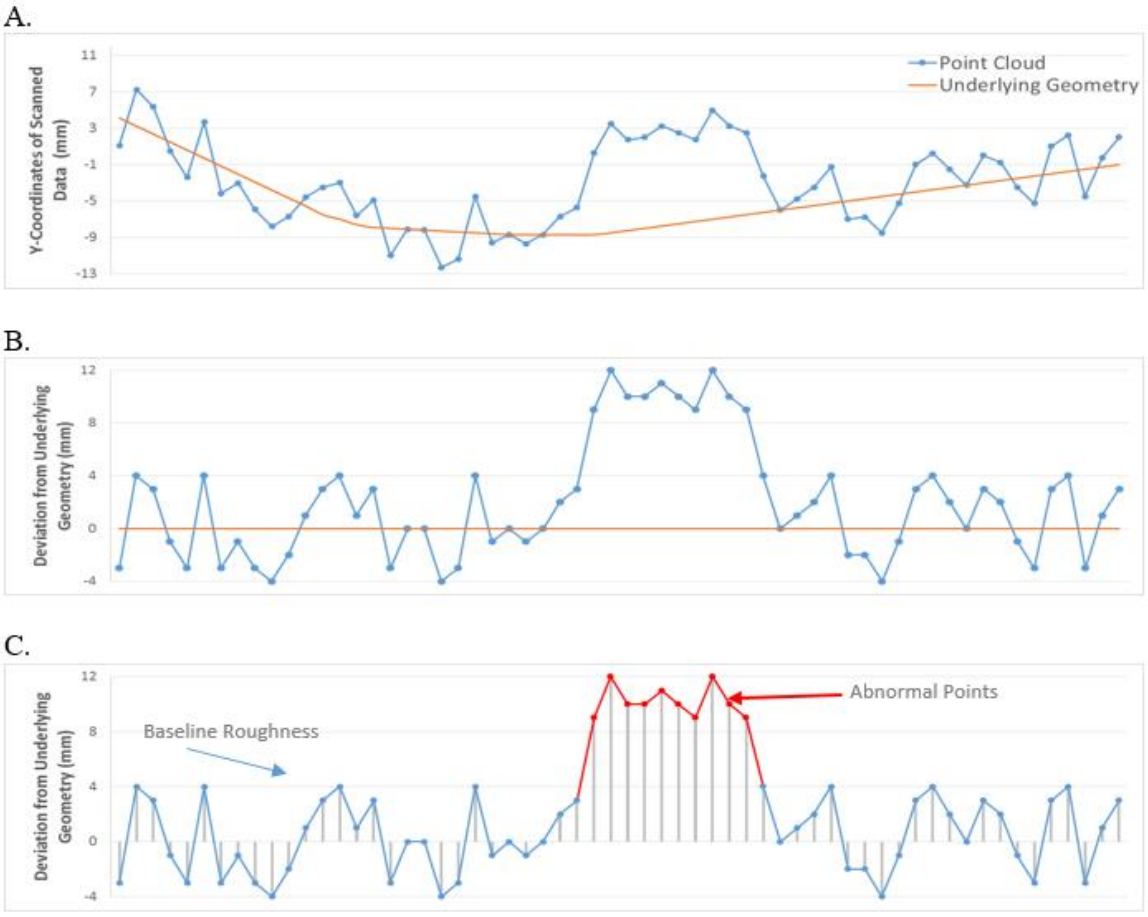


Figure 4.6—Parameter calculation process A) determine underlying geometry, B) calculate deviations from the underlying geometry, C) identify and measure abnormal points based off of the deviations from the underlying geometry and assigned acceptance criteria

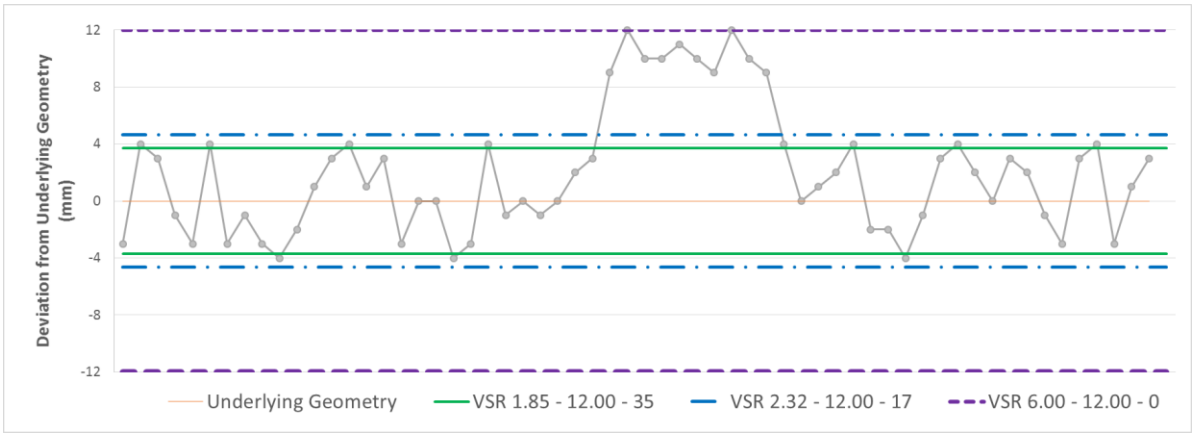


Figure 4.7—Comparison of control limits where data points are considered abnormal based on the specified baseline roughness of each example specification

VSR 1.85 – 12.00 – 35

This specification criteria considers the 21 points with a deviation from the underlying geometry greater than 3.7 mm (twice the specified baseline roughness represented by thick, green solid line in Figure 4.7) as abnormal. These points were omitted from the actual baseline roughness parameter calculation; however, they were captured in the abnormality percentage parameter given. The 21 points over the entire inspected area of 60 points, or 35%, were considered abnormal. This is right at the threshold as presented by the third parameter (twice the specified baseline roughness). The abnormality level sets the maximum deviation from the underlying geometry of the data points to 12. This would mean the part would be rejected if points greater than 12 mm from the underlying geometry were present.

VSR 2.32 – 12.00 – 17

The 10 points with a deviation from the underlying geometry greater than 4.64 mm (represented by alternating dot and dashed blue line in Figure 4.7) are considered abnormal for this specification criteria. The same process was used as part A to determine the parameters of the criteria.

VSR 6.00 – 12.00 – 0

In this scenario, all points within ± 12 mm (represented by purple dashed line in Figure 4.7) of the underlying geometry are not considered abnormal since the abnormality level is exactly twice the specified baseline roughness. All 60 data points are used in calculation of the actual baseline roughness for this criteria. This particular specification does not allow any

point to be abnormal, but it opens up the deviation from the underlying geometry to be considered abnormal.

Other Variations

This surface profile would also be considered acceptable where any of the three parameters are greater than those currently stated, such as VSR 4.12 – 15.00 – 40. This is because the specification notes the maximum acceptable value for use of all parameters. However, one must consider resulting surface variations if specifying values for the baseline roughness and abnormality level greater than their sample surfaces, since a lower quality surface than the sample could be considered acceptable under these increased parameters.

Customers need to be conscientious when specifying cast surfaces as there can be an infinite number of surfaces that would be acceptable for each VSR surface specification. Variations of a surface profile for each criteria assigned in the previous example are seen in Figure 4.8: VSR 1.85 – 12.00 – 35, VSR 2.32 – 12.00 – 17, and VSR 6.00 – 12.00 – 0. Sample A of Figure 8 is identical to the profile found in Figure 4.6. Based on the number of points exceeding the bounds of twice the specified baseline roughness, as previously demonstrated in Figure 4.7, Samples A-B of Figure 8 would be considered acceptable with all three standards previously mentioned. Samples C-F of Figure 8 only correspond to VSR 6.00 – 12.00 – 0 since a greater number of points exceed twice the specified baseline roughness of the other examples. As a general rule, the specified baseline roughness and abnormality percentage are inversely related when assigning different specifications to the same surface. To simplify specification assignment and interpretation, it is suggested the abnormality percentage for an 8 centimeter by 8 centimeter surface area does not exceed 10%.

Examples of acceptable surfaces based on VSR specification samples

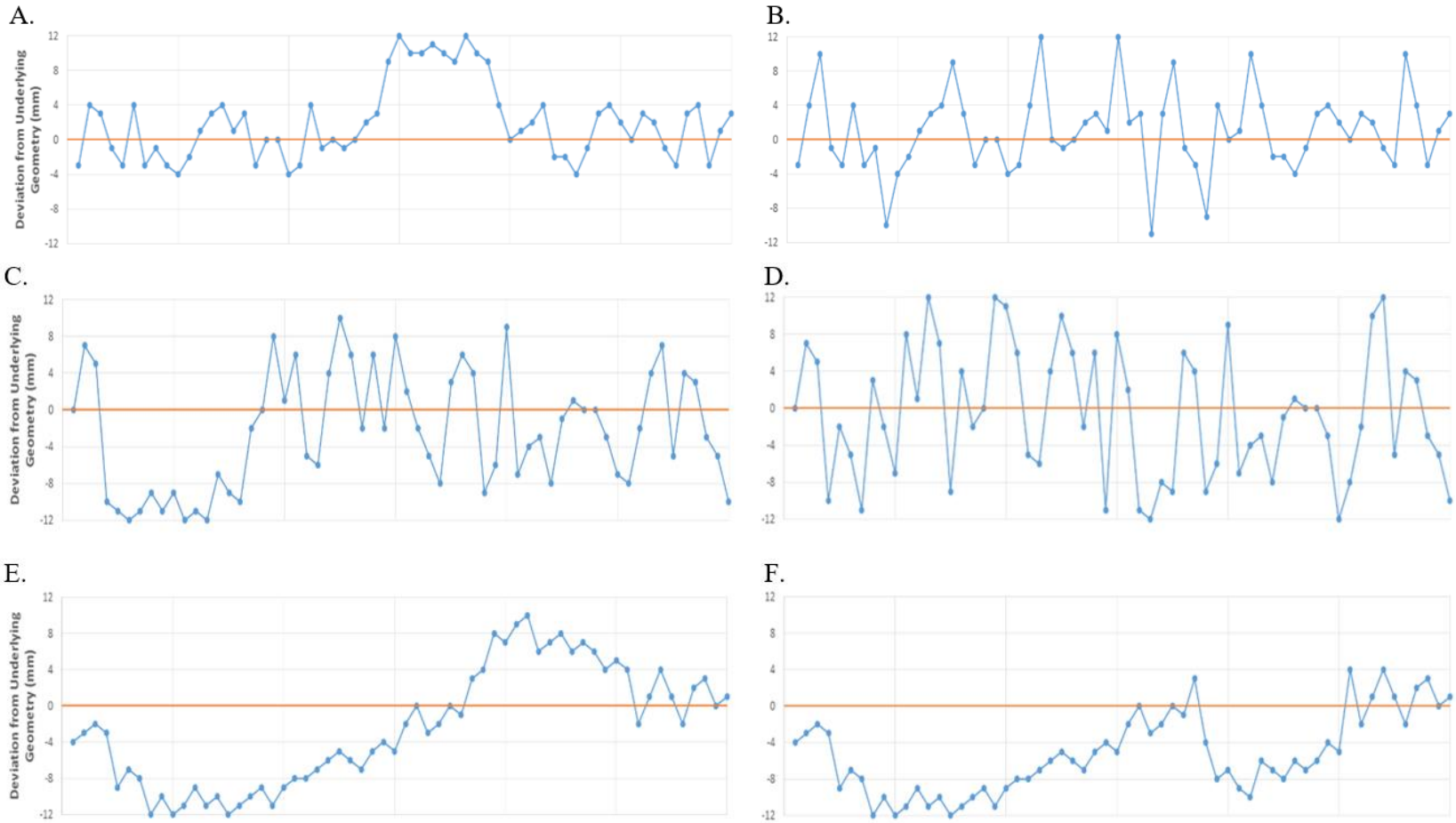


Figure 4.8—Surface profiles representing an 8 centimeter by 8 centimeter constant cross-section of acceptable surfaces specified as VSR 1.85 – 12.00 – 35 or VSR 2.32 – 12.00 – 17 (A-B) and VSR 6.00 – 12.00 – 0 (A-F) based on the distance of each data point from the underlying geometry

Designers must determine the type of surface which is acceptable for their component, and then write the appropriate VSR specification, keeping in mind that more restrictive specifications will increase the procurement cost. A major advantage of the VSR standard is that the designer can quantify the surface that is acceptable, and not rely on comparative methods which may not result in the surface they were expecting.

All parts deemed acceptable through VSR 1.85 – 12.00 – 35 and VSR 2.32 – 12.00 – 17 will also be considered acceptable under the VSR 6.00 – 12.00 – 0 criteria; however unlike the other two requirement examples, VSR 6.00 – 12.00 – 0 also can be specified, which increases the number of allowable points located further from the underlying geometry while maintaining a roughness less than or equal to 6.00 mm. Since an abnormality is defined as greater than twice the specified baseline roughness, any data falling within ±12 mm from the underlying geometry would not be considered abnormal. Therefore, since the sample surfaces do not have any data points falling outside of this range, the abnormality percentage is 0%. This method sets a range on the maximum permissible deviation from the underlying geometry as opposed to calling out any abnormalities and is ideal when specifying no abnormalities can be present on the surface.

In order to begin assigning criteria to their castings, customers can use current castings as a baseline for specifying a standard. To do this, customers can select a part with what they consider the least acceptable surface roughness and abnormality level, or a part that is not of the highest quality but still meets their current surface expectations. After using a non-contact method to collect data points from the surface, the customer can select a criteria for that surface by comparing the data to the underlying geometry. A single acceptance criteria may be specified over the entire cast surface, or multiple criteria may be specified for various areas

of the casting in order to reduce the variation of interpretation using the methods discussed in this section.

IV. Discussion

The quantitative standard eliminates the discrepancies between the manufacturer's and customer's interpretation of inspection criteria, as seen in the qualitative standards. The reduced complexity of the evaluation criteria and variation from qualitative inspection allows for a clearer understanding of expectations.

The quantitative standard uses hard data to evaluate whether or not the surface is or is not acceptable and does not rely on an individual's sensory capability. This hard data does not differentiate between the types of abnormality present, which is beneficial if an unexpected abnormality appears on the final part and was not taken into consideration by the customer when specifying the surface. Additionally, the percent of the surface that is classified as abnormal, which was specified in only one of the qualitative methods, is specified within the standard and can be modified, if desired, allowing the customer to better relay his or her requirements. These aspects of the quantitative standard allow for a clearer communication of expectations of cast surface specifications between the manufacturer and customer.

Work is ongoing by the authors to develop methods to automate the data collection and data analysis. Ultimately these techniques would be integrated into a portable scanning device that a user could enter the specified VSR values and point the scanner at the 8 by 8 centimeter surface patch in question and it would determine if the surface was acceptable. The intent is that this device would be used to assist the manual visual inspection process; however, future efforts could include this methodology in an automated inspection process.

V. Conclusions

Surface standards for metal cast surfaces help to determine the acceptability of surface quality. Implementation of the quantitative inspection standard will increase the quality of metal cast surfaces by improving communication between manufacturers and customers in the interpretation of requirements. Methods to collect and clean point cloud data for use in this standard are currently being developed to increase repeatability and reproducibility when calculating components of the VSR. This includes the development of algorithms for the underlying geometry of the scanned part. Future work includes exploring the feasibility of an automated inspection process to eliminate the need for human interaction in the process.

VI. Acknowledgements

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-033. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

VII. References

[1] Daricilar, G., 2005, “Measurement Error of Visual Casting Surface Inspection,” M.S. Thesis, Iowa State University, Ames, IA.

[2] Eubanks, P. E. et al, “Standard for the Visual Inspection of Casting Surfaces,” Steel Founders' Society of America Alloy Casting Institute, Des Plaines, IL, 1969.

[3] MSS SP-55, Quality Standard for Steel Castings for Valves, Flanges and Fittings and Other Piping Components (Visual Method), Manufacturers Standardization Society of the Valve and Fittings Industry, New York, NY, 2006, www.msshq.org

[4] A802/A802M-95, Standard Practice for Steel Castings, Surface Acceptance Standards, Visual Examination, ASTM International, West Conshohocken, PA, 2015, www.astm.org

[5] BNIF 359, Characterization of Surface Condition of Castings, Editions Techniques des Industries de la Fonderie, Sèvres, France, 1996

[6] American Society of Mechanical Engineers, 2009, "ASME B46.1 Surface Texture: Surface Roughness, Waviness and Lay," ANSI/ASME, New York, NY, 2009, pp. 67-70, www.asme.org

[7] ISO 11971, Steel and iron castings –Visual examination of surface quality, International Standards Organization, Geneva, Switzerland, 2008, www.iso.org

[8] BS EN 1370, Founding—Surface roughness inspection by visual tactile comparators, British Standards Institution: European Committee for Standardizations, London, United Kingdom, 2011, www.bsigroup.com

[9] A997-08, Standard Practice for Investment Castings, Surface Acceptance Standards, Visual Examination, ASTM International, West Conshohocken, PA, 2012, www.astm.org

[10] Voelker, M., Kemper, P., and Peters, F., "Development of a Digital Standard to Specify Surface Requirements," presented at the 69th Annual Steel Founders' Society of America—Technical & Operating Conference, Chicago, IL, Dec 10-13, 2015, Iowa State University, Ames, IA—unpublished.

CHAPTER 5: EVALUATION OF SLICING METHODOLOGIES TO DETERMINE UNDERLYING GEOMETRY OF CAST METAL SURFACES

A paper submitted to *Journal of Materials Processing Technology*.

Michelle M. Voelker[7] and Frank E. Peters[8]

Abstract

The geometry of cast metal surfaces is complex, and in reverse engineering applications, the task of identifying the true geometry of the part is challenging. This is due to the inherent roughness and surface variants, or abnormalities, caused by the nature of the casting process. In addition to abnormalities due to fusion and porosity, among others, non-uniform mold movements and metal shrinkage of the part will cause variation in the part's original geometry. This geometry of the part including shrinkage, or underlying geometry of the casting, is used in evaluating the baseline roughness and abnormality level in the Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces, and without this geometry, an accurate means of calculating surface parameters of castings does not exist. This paper outlines a slicing process to estimate the underlying geometry of castings for use in the standard. Various fitting methods for the two-dimensional slices are explored to evaluate the effect each method has on accurately representing the actual part's geometry and surface characteristics while minimizing the effects of abnormalities on the end product.

---

[7] Primary author; Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa, 50011, United States; 0000-0002-3521-694X
[8] Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa, 50011, United States; 0000-0002-8998-0062

I. Introduction

The proposed standard for Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces (Appendix A) allows for quantitative specifications for inspection [1, 2]. The standard allows point cloud data to be analyzed to calculate components of the Voelker Surface Ratio (VSR), including the baseline roughness, abnormality level, and abnormality percentage. However, without standard methods to collect and clean this data, the repeatability and reproducibility is highly variable. A process for measuring the components of the standard must be outlined so the inspection process is consistent among manufacturers. This will require methods in order to determine the underlying geometry to customize the calculations based on the actual geometry of the part after molding and post-shrinkage, which would vary from part to part and give inconsistent measurements of the surface. This article proposes algorithms to determine the underlying geometry of metal castings for use in this standard.

II. Previous Work

The evaluation of current contact and non-contact methods for evaluating surface parameters was explored. This information was used to determine the mathematical gap for the evaluation criteria of the digital standard. Methods were evaluated to explore the general proof of concept of the process.

Previous work included the evaluation of the standard parameters using alternative methods to non-contact scanning. This included comparisons of the casting to the original CAD model, contact profilometers, and non-contact profilometers.

The first method explored was to compare the point cloud data to an original solid computer model. This nearly perfect geometric model of the casting could easily be compared to a scan of the actual cast part to identify variation between the two. However, this method was not considered feasible due to the variability between the part and model after manufacturing from contraction, among other geometric variations. These variations would not allow for an accurate calculation of the surface parameters of the standard.

To evaluate contact methods, a MahrSurf SD 26 surface profilometer was used. The profilometer software is designed to calculate the roughness by eliminating waviness in the sample; however, unlike machining, the waviness and roughness are not cyclical, which makes filtering more difficult. For example, too much of the surface variation for roughness was filtered out with the waviness for rougher parts. Rougher parts also more frequently gave incomplete readings due to the curvature and height of abnormalities exceeding the range on the profilometer. Additionally, since this method only provided data for a three-inch, two-dimensional profile, it was not representative of the entire cast surface.

The Zygo Surface Profiler was also examined. This non-contact, three-dimensional profilometry method took a surface scan of a one-centimeter square. The filtering functions for roughness and waviness appeared to be similar to the contact method. Much like the contact method, this method was not feasible due to the very small surface area the sample could cover on the part and the time it took to collect data.

To evaluate current scanning methods, several cast surface samples and replications were evaluated. Three sample casting surfaces were scanned manually using a Faro Edge scan arm. The scan data was saved as a point cloud text file and imported into commercial software for evaluation. The point clouds were not subsampled or manipulated prior to running the

software. Two methods used for the initial evaluations of the surface included localized roughness (comparison of points to surrounding points) and shape fitting (comparison of points to best fit plane). The analysis was completed within the software for the proof of concept.

Figure 5.1 shows a flat sample casting evaluated based on the localized roughness and shape fitting methods. For the localized roughness (left), the abnormalities, noted by an arrow, were evident in most cases, and the shape fitting method was successful at identifying abnormalities (right). Figure 5.2 shows the D5 SCRATA comparator evaluated using both methods. The localized roughness, pictured left, clearly identifies the abnormalities. For shape fitting, right, the abnormalities are still visible to an extent; however, the color mapping shows an area on the comparator that is a greater distance (blue) from the geometric shape, a plane. This indicates the comparator is curved. This causes an inaccurate representation the roughness due to the curvature in the sample. In order to get an accurate measurement of the roughness, the surface must be compared to a surface accurately representing the underlying geometry of the scanned surface.

Figure 5.3 shows the E3 SCRATA comparator results. The abnormality, marked with an arrow, is evident and roughness consistent for the best fit plane, right; however, the localized roughness, left, is not identifying the flat surface on the top of the abnormality as abnormal. This is because the average deviation in elevation of the unit vectors of surrounding points is minimal.

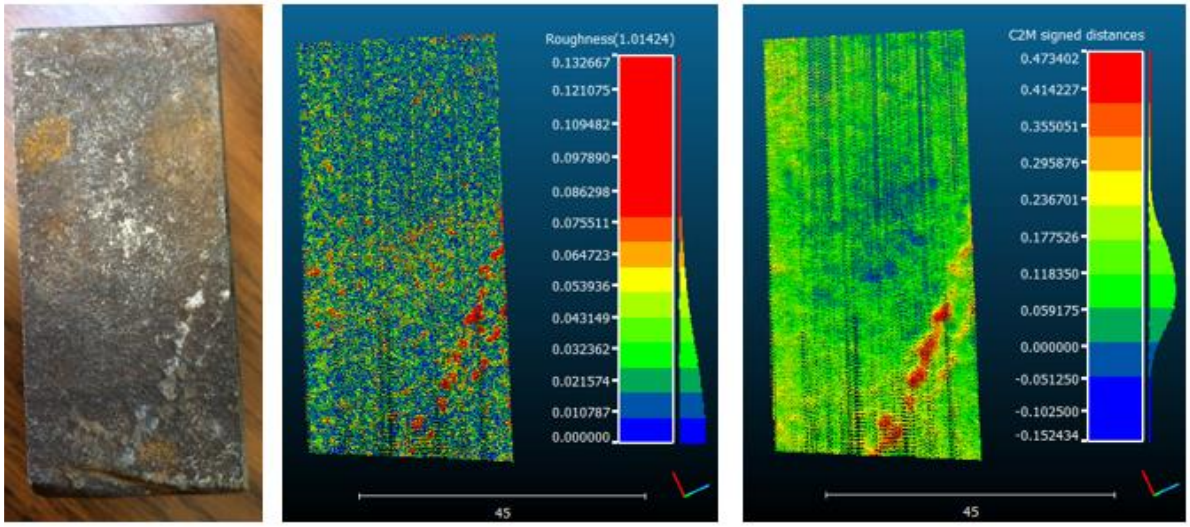


Figure 5.1—Localized roughness versus best fit plane distance on sample casting

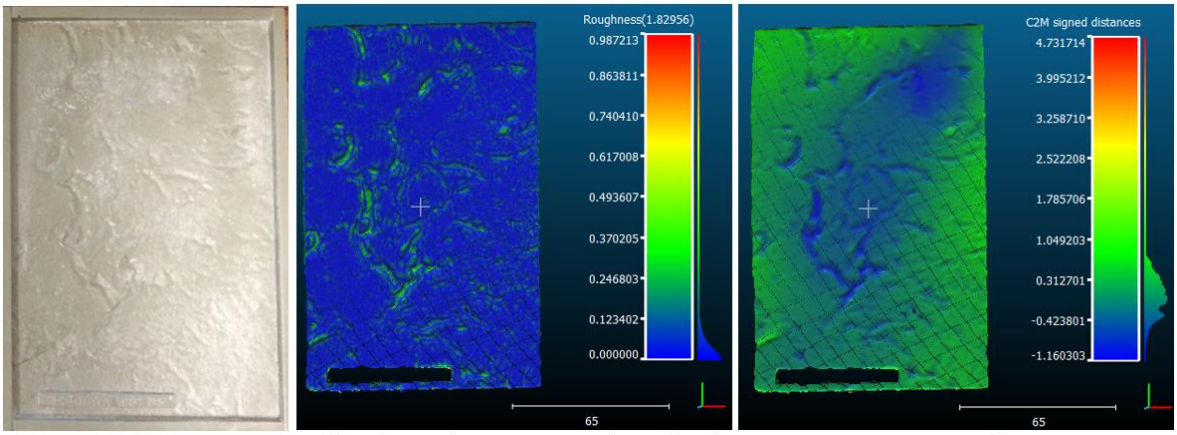


Figure 5.2—Localized roughness versus best fit plane on D5 SCRATA plate

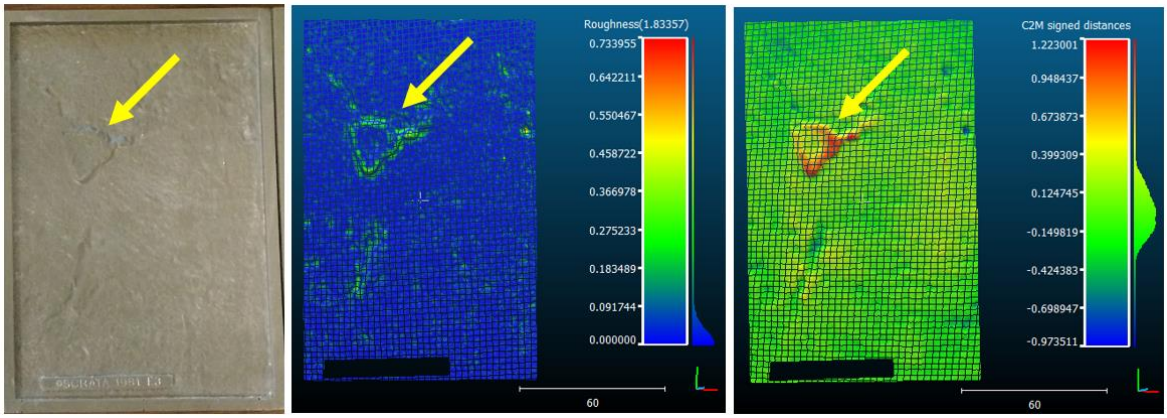


Figure 5.3—Localized roughness versus best fit plane on E3 SCRATA plate

From the results, the localized roughness did not properly identify all abnormalities, or anomalies of the surface exceeding three times the baseline roughness. In the E3 SCRATA plate, the abnormality was large and had a smooth face. In this case, only the edges of the abnormality were detected. Since the points on the top of the abnormality were at the same elevation relative to the surrounding points, the local roughness, or change in elevation, was minimal and failed to identify the specific location as abnormal. Similarly, the best fit plane fell short when determining parameters for non-planar, or non-geometric, shapes. Regardless of whether or not a part is designed to be a specific shape, the resulting manufactured part will not be identical to the design due to the nature of the casting process. For example, if a cast feature was designed to be a planar, non-uniform shrinkage during cooling may have occurred causing the geometry to stray slightly from the intended geometry. Therefore, if a plane was fit to a scan of the surface in the surface fitting method, the deviations may be skewed as seen in Figure 5.4. These variations could make the difference between whether or not a part passes or fails an inspection; in fact, the Linear Fit may fail due to the abnormality percentage exceeding that of the acceptance criteria in the example. Therefore, these methods cannot be considered appropriate for standard parameter calculations.

The localized roughness and shape fitting methods used in this case study are not sufficient to calculate the parameters of the standard. A consistent method to calculate these parameters is necessary so there is agreement among suppliers in the interpretation of the standard. In order to consistently achieve this, the underlying geometry, or true geometry of the casting, can be used as a baseline to calculate the parameters from. This will allow for inherent variation in the casting process while consistently delivering a reliable value for inspection to the standard.

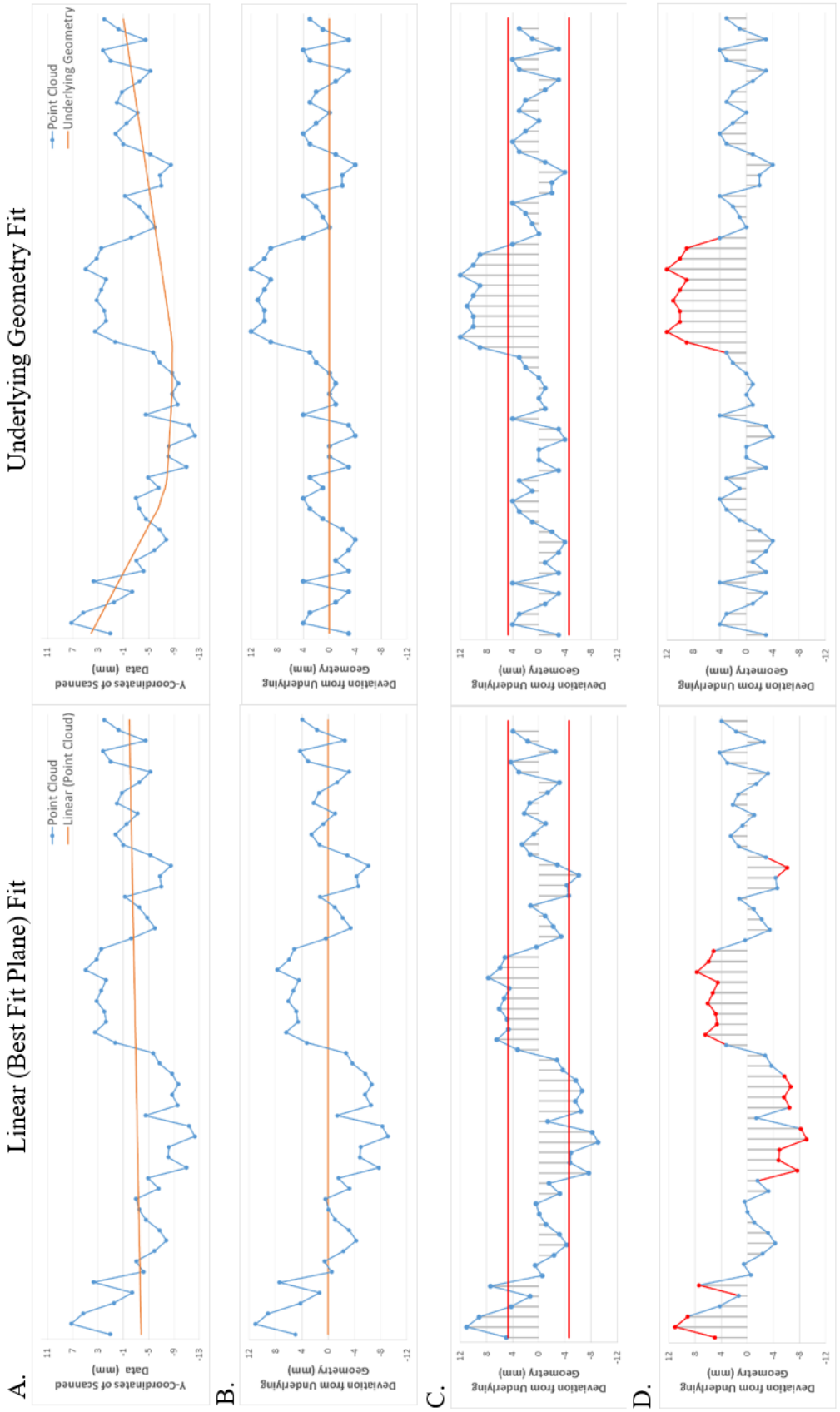


Figure 5.4—Determining parameters based on shape fitting versus underlying geometry example
A) fitting plane versus underlying geometry, B) calculating the deviations from the fit line to each individual data point, C) setting bounds on for abnormalities at twice the baseline roughness, D) identifying abnormal points greater than twice the value of the specified baseline roughness (>4.64 mm)

III. Exploration of Slicing Methodology for Underlying Geometry

The digital standard proposed by Voelker and Peters [3] quantifies surfaces based on deviations of data obtained from digital scanners from the underlying geometry. The quantifiable parameters include the baseline roughness, abnormality level, and abnormality percentage. The baseline roughness is the roughness average of the cast surface excluding abnormalities, or anomalies of the surface. The abnormality level is the maximum allowed deviation from the actual part geometry, and the abnormality percentage is the percentage of the surface contained in the region bounded by twice the baseline roughness and the abnormality level. Inspecting castings using the new digital standard can improve communication between the customer and manufacturer in addition to reducing the discrepancies between inspectors' interpretations; however, the parameters must be calculated consistently based on reference geometry to achieve repeatable results. Not only will the underlying geometry enable a consistent means of calculating the parameters of the standard, but it can also be used in other applications such as reverse engineering of castings.

The underlying geometry dictates all calculations for the components of the specification. The end product of the underlying geometry algorithms is a smooth surface that accurately represents the casting free of surface roughness and abnormalities. This surface will then be compared to the actual digital surface in order to calculate the surface roughness, abnormality level, and abnormality percentage.

As previously discussed, fitting a geometric shape using shape fitting methods or using a CAD model of the part to compare to the scan of the actual part does not allow for the inherent deviations in geometry resulting from the casting process. Additionally, profilometry software using filters for roughness and waviness do not properly accommodate for the

noncyclical variation and abnormalities in the cast surface. Determining the underlying geometry of the casting will not only aid in the identification and measurement of abnormalities, but it will provide a consistent method in calculating the actual baseline roughness.

This paper explores slicing methods to calculate the underlying geometry and was chosen based on its simplicity and calculation speed. Using the original point cloud, this method aimed to eliminate roughness and abnormalities in order to calculate the underlying geometry. The method was evaluated based on its ability to eliminate abnormalities and the baseline roughness when compared to the original point cloud as described in the following paragraphs.

Determining a best fit surface for a point cloud with complex geometry is not an easy task. In order to reduce the complexity, the surface will be examined in two-dimensional data sets. This can be compared to the process of integration of three-dimensional calculus where the double integral splits the surface into two two-dimensional parts. Rapid prototyping technologies also use this technique to simplify the construction of three-dimensional objects by only examining the cross section of the model at incremental locations. Similarly for the slicing method, the digital representation of the cast surface will be sliced in order to simplify the surface into a series of two-dimensional data. First, the point cloud will be sliced into small, unidirectional slices of points of a given width, which will be condensed into a single two-dimensional data strip. Each subsection will be examined and a curve will be fit through the set of data points to represent the underlying geometry at each slice. The slices will then be compiled and a mesh will be created across slices in order to create a composite surface to compare to the point cloud. A generalization of this process is broken down in Figure 5.5.

This process will be completed for angles of 0, 45, and 90 degrees to eliminate the effects of skewed fitting due to linear indications parallel to the slices. Once three surfaces are constructed, the intersections of the surfaces will be identified, and the medial surface will be selected for the final underlying geometry. Details of this process are shown in the flow chart in Figure 5.6.

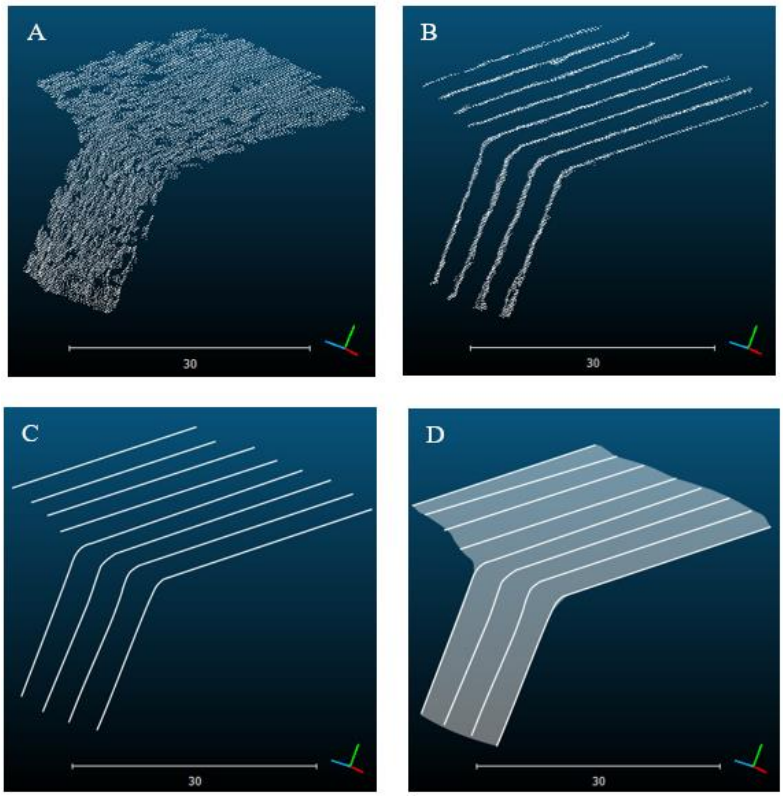


Figure 5.5—Slicing method for calculating underlying geometry A) original point cloud, B) slice and condense data, C) fit curve to 2D segments of data, D) create surface across curve

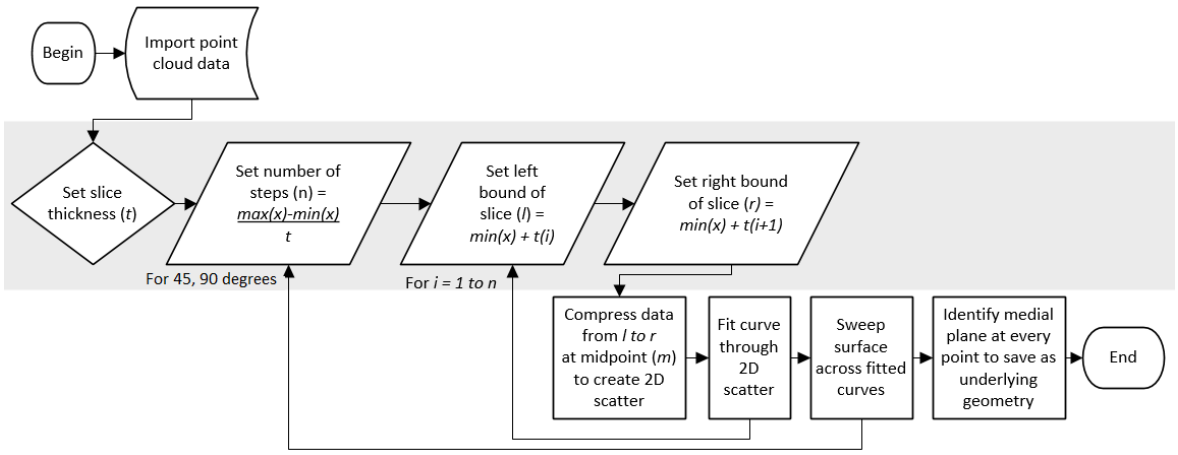


Figure 5.6—Break down of slicing algorithms including importing the point cloud, slicing the point cloud, and determining the underlying geometry through fitting curves to the slices

Based on slicing procedure studies in rapid prototyping [4], the ideal method of slicing to most accurately represent the underlying geometry would be to compress the points in each

slice to the center in order to fit a two-dimensional curve through it. In rapid prototyping applications, choosing to contact the nominal geometry at either end of the slice causes over or under sizing of the part [3]. Similarly for determining the underlying geometry, the less distance the points must travel to be compressed, the more accurate the underlying geometry. For simplicity, Figure 5.7 shows an example of the compression of two-dimensional slices into lines of data. For the three-dimensional point cloud data, slices will be made in the z-direction and the resulting slice "shape" can be seen in the x-y plot of data. Regardless of how the points are compressed, the resulting shape, or plot, will be identical with exception of the positioning relative to the original part. In the cases of two-dimensional shapes, the main body of the part will be oversized since the compression takes into account all data points in the slice. However, as the size of the slice approaches zero, the difference in the positioning of the two compression methods approaches zero.

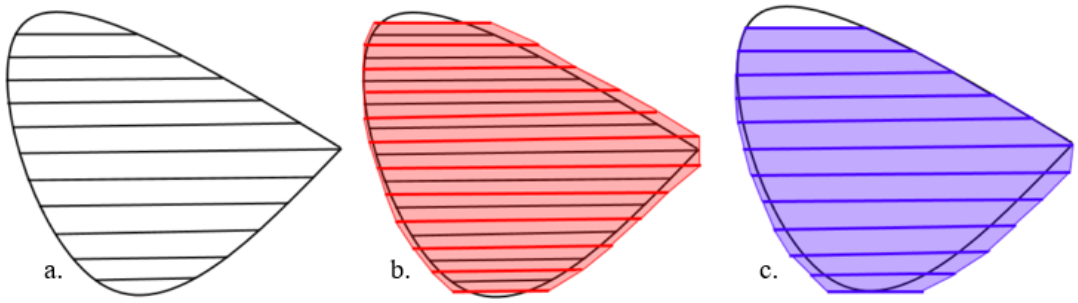


Figure 5.7—Effects on a) sliced original shape from b) central and c) extrema compression

Potential risks associated with fitting data to casting surfaces includes the effects from the presence of abnormalities and roughness variation. This is because extreme variations may cause any fitting function to be skewed when passed through an extremely rough region or abnormality, resulting in a function not accurately representing the underlying geometry of the casting. This will be explored in depth by looking at theoretical cross-sections of data to

see which curve fitting function will best accommodate for the varying levels of abnormalities on different geometries.

Three variations of three sample surface profiles were constructed, as seen in Table 5.1, to analyze effects of the geometry and abnormalities on the slicing methods. The surface profiles consisted of flat, curved, and wavy to see how each fitting technique accommodated each profile.

Table 5.1—Sample profiles

Abnormalities	Geometry: Flat	Curved	Wavy
No	Type 1	Type 2	Type 3
Yes	Type 4	Type 5	Type 6
Removed	Type 7	Type 8	Type 9

Profiles under each geometry were identical with exception to the set area where an abnormality was introduced, as seen in Figure 5.8. For a surface with no abnormality, the general surface roughness remained consistent across the abnormal region. For a surface with an abnormality, the abnormal region introduced a protrusion or depression in addition to roughness across the abnormal region. For a surface with a removed abnormality, no data was present for the abnormal region.

An additional ten flat surfaces were constructed with variations in surface profile characteristics. These characteristics include roughness, scaling, point density, abnormality height, and abnormality width.

The characteristics described in this paragraph are the default for all flat samples unless otherwise noted. Samples ranged from -1.00 to 2.55 units along the x-direction and -25.0 to 25.0 units in y-direction. The increment in the x-values was 0.05 units, which would be representative of a point cloud that was cleaned in order to reduce redundant data. The roughness profiles were randomly generated to represent the actual variation of a cast surface slice. No abnormalities were present in the roughness, scaling, and point density samples.

Exceptions to these criteria for specific flat samples will be described in detail in the following paragraphs.

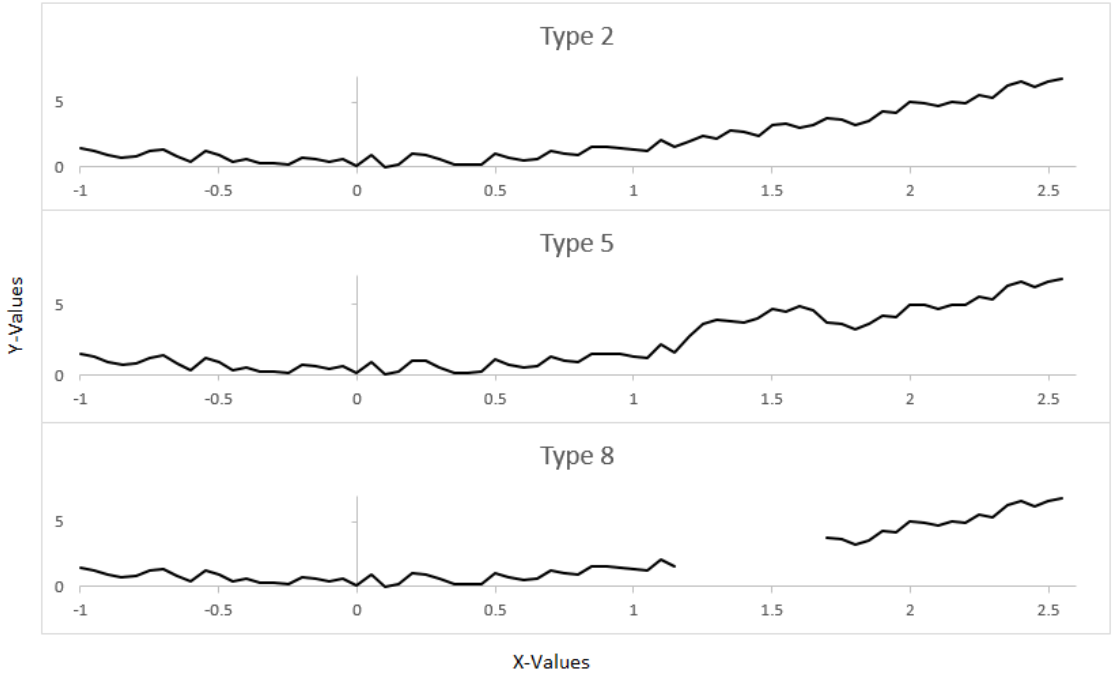


Figure 5.8—Example of curved surface profiles: Type 1) profile with no abnormalities, Type 5) abnormality present, and Type 8) abnormality removed from data set

Two roughness characteristics were explored. First, an alternating roughness profile was constructed. In this profile, the y-data was cyclical such that every other value cycled from positive to negative. This cyclical profile is similar to the surfaces seen in machining processes. The second profile had a random roughness, which is more representative of cast surfaces. These surfaces were constructed using a random number generator between -25 and 25 given the aforementioned constraints on the alternating roughness profile.

The scaling characteristic manipulated values on the y-axis. The original profile for the scaling comparison met the default requirements. The reduced profile used the exact same x-values as the normal profile; however, the y-values were scaled down by 100.

Three point density characteristics were evaluated, which manipulated the total number of data points in the data set. The 100% point density used the default requirements mentioned previously. For the 50% point density, every other x-value in the profile was eliminated, which reduced the point density to half of the original. Therefore, the increment in the x-values was 0.10 units. The point density was doubled from the original for the 200% point density with x-values at 0.025 unit increments.

For the abnormality characteristics, four new surface profiles were constructed based off of the random roughness profile from the roughness characteristic data set. This profile was also compared to the abnormality height and width characteristics since it did not contain any abnormalities. Abnormality height consisted of tall and short abnormalities. Tall abnormalities reached up to 60 units on the y-axis, and short reached up to 30 units. Abnormality width consisted of wide and narrow abnormalities. Wide abnormalities were 1.50 units on the x-axis, and narrow were 0.75 units. All abnormalities were centrally located in the sample profile. The abnormality direction, protrusion versus depression, was not explored since the fitting methods and the lack of normal vectors associated with each point could not differentiate the two.

Various methodologies were explored to analyze how each fitting or filtering method represented the underlying geometry. The following fitting methods were evaluated: segmenting, mid-point locus, polynomial, and moving average.

Each fitting method was compared to the ideal underlying geometry, which was determined by hand for each profile, using the roughness average (R_a). For the flat surface profiles, the ideal underlying geometry was at y = 0 for all x-values. The coefficient of determination, or *R-squared* value, was not used due to the high variation in roughness

between sample profiles. Since each surface profile contained unique surface roughness values, the R_a could not be compared across each sample profile directly; however, the values could be compared for a single sample profile across fitting methods. To make a direct comparison of a single fitting method across surface profiles, the R_a values were normalized. To normalize the values, the R_a of each surface profile was calculated in comparison to the ideal underlying geometry ($R_{profile}$). Then the R_a from each fitting method, given a specific surface profile ($R_{method\ |\ profile}$), was divided by $R_{profile}$ resulting in the normalized roughness ($\%R_a$) as seen in Equation 5.1.

$$\%R_a = \frac{R_{method\ |\ profile}}{R_{profile}} \quad \textbf{(Eq. 5.1)}$$

Since the R_a of any given data set cannot be less than zero, the $\%R_a$ also cannot be less than zero. If the fitting method was identical to the ideal underlying geometry, $R_{method\ |\ profile}$ would equal 0.00 resulting in a $\%R_a$ of 0.00%. A $\%R_a$ equal to 0.00% would indicate the fitting method is a perfect representation of the ideal underlying geometry. Fitting methods where the $\%R_a$ is greater than 100% indicate that the actual surface profile more accurately represents the ideal underlying geometry than the fitting method represents the ideal underlying geometry. In essence, the lower the $\%R_a$, the better the fitting method represents the ideal underlying geometry.

Segmenting

A segmented filter is a simple method to eliminate non-uniform waviness from a surface profile. This method, also known as high-pass filtering, segments the data into equal sample lengths along the x-axis [4]. Each segment is then fit to straight lines. Additionally, much like with the slice size as discussed previously, the ideal segment length must be

identified in order to increase the accuracy of the estimated underlying geometry. If the segment length is too long, it will not accurately represent smooth curvatures in the geometry, which will result in a jagged, ruled surface approximation. However, if the segments are too short, they will not cover enough data points to fit a line through in order to eliminate the variation in the surface roughness. The parameters used in this methodology included a sample length of 1/3 and 1/6 the total length of the profile, which will be referred to as *Seg 3* and *Seg 6*, respectively.

Mid-point Locus

The mid-point locus line can also eliminate non-uniform waviness from a profile. In this method, a window of a specified width is moved across the profile along the x-axis and the average height is plotted in the center of the window. Typically, the window should overlap the previous region when it is shifted, however, the shifting distance does not have to be uniform [5]. The parameter combinations used in this method including the sample width, which is the fraction of the total sample length, and overlap, which is the fraction of the width that is overlapped, can be seen in Table 5.2.

Table 5.2—Parameters for mid-point locus line

Line Name	Width	Overlap
L-1/3-1/2	1/3	1/2
L-1/6-1/4	1/6	1/4
L-1/6-1/2	1/6	1/2
L-1/6-3/4	1/6	3/4

Polynomial

Additionally, polynomial filters can be used to eliminate waviness. For short lengths of data, polynomial curves are fit to the data using least-squares method [6]; however, by continually increasing the degree of polynomial, the R_a of the line will inherently decrease.

To avoid unintentional overfitting and manual intervention, a third degree polynomial was fit to the profile, *Order 3*. Additionally, a best fit was manually selected by the user to most closely fit the ideal underlying geometry, *Order Best*.

Moving Average

Weighing functions, such as a moving average iterative weighing function, can also be used to minimize the effects of outliers [6]. In this method, the data is sorted based on its x-value, and each point is replaced by the average value of the surrounding points. The period parameter in this method represents the number of total points used in calculating the average. The parameters used in this methodology included a period 5 and period 11 moving average, *MA5* and *MA11*, respectively.

IV. Results

To evaluate the overall fitting methods, the sample statistics were calculated on the $\%R_a$ for sample profiles Type 1-9 of each fitting method. Additionally, a paired t-test was conducted to determine if the differences in the test results were statistically significant using a p-value cutoff of 0.05. The results from this analysis can be seen in Table 5.3.

As seen in the data table, the *Order Best* fitting method had the lowest mean $\%R_a$ and had a statistically significant difference in means compared to six of the nine other methods; however, it is noted this method was highly manual compared to all other methods, which were strictly calculations. It is also noted the *L-1/3-1/2* and *Order 3* fitting methods had a mean and standard deviation much greater than the other methods.

To evaluate how the geometry and presence of abnormalities affected the fitting methods, sample statistics were calculated on the $\%R_a$ across all fitting methods of each geometry and abnormality. Again, a paired t-test was conducted to determine if the differences in the test results were statistically significant using a p-value cutoff of 0.05. The results from this analysis can be seen in Table 5.4.

Table 5.3—Sample statistics on fitting methods' $\%R_a$ based on sample profiles Type 1-9 (n = 9)

Fitting Method	Mean (%)	Standard Deviation	Range
Seg 3	60.45	0.6282	(0.1059 , 2.0796)
Seg 6 *†	44.74	0.2207	(0.1648 , 0.8955)
L-1/6-1/2 *	43.91	0.2653	(0.1443 , 0.9935)
L-1/3-1/2 *‡	124.08	1.2025	(0.0775 , 3.6338)
L-1/6-3/4 *	51.05	0.3416	(0.1535 , 1.1390)
L-1/6-1/4 *	49.81	0.3916	(0.1049 , 1.1882)
Order Best †	28.10	0.3272	(0.0000 , 0.7772)
Order 3	255.82	4.1641	(0.0212 , 12.7297)
MA11 †	39.25	0.2234	(0.0894 , 0.8245)
MA5	50.55	0.1895	(0.2428 , 0.8690)

Indicates statistical significance based on paired t-test (p-value < 0.05) with *Order Best, † MA5, ‡ L-1/6-1/4

Table 5.4—Sample statistics on geometry and abnormalities' *%Ra* based on sample profiles Type 1-9 (n = 3)

	Mean (%)	Standard Deviation	Range
Geometry			
Flat	66.53	0.82	(0.0000 , 4.0642)
Curved	62.20	0.73	(0.0000 , 3.6338)
Wavy	95.68	2.35	(0.0828 , 12.7297)
Abnormalities			
No	86.69	2.33	(0.0000 , 12.7297)
Yes*	95.23	0.72	(0.4589 , 4.0642)
Removed	42.49	0.81	(0.4589 , 4.2326)

Indicates statistical significance based on paired t-test (p-value < 0.05) with *Removed

It is important to note from this data, the geometry does not have a significant impact on the $\%R_a$; however, the wavy geometry had increased values for the mean and standard deviation. The presence of abnormalities, on the other hand, did show a statistically significant difference between the profile with an abnormality and the same profile with the abnormality data omitted. It is also noted profiles without abnormalities have a higher standard deviation, and profiles with abnormalities removed have a mean at least half of that of the profiles with and without abnormalities.

The sample statistics were calculated on the $\%R_a$ for the flat sample profiles of each fitting method to evaluate the overall fitting methods. They were also calculated across all fitting methods of each flat surface profile to evaluate how the surface profile characteristics affected the fitting methods. Paired t-tests were also conducted to determine if the differences in the test results were statistically significant using a p-value cutoff of 0.05. The results from these analyses can be seen in Table 5.5 and 5.6.

Table 5.5—Sample statistics on fitting methods' *%Ra* based on flat sample profiles (n = 10)

Fitting Method	Mean (%)	Standard Deviation	Range
Seg 3 *†	42.28	0.2787	(0.0922 , 0.8615)
Seg 6 *+	25.39	0.2892	(0.0818 , 0.8914)
L-1/6-1/2 *	22.89	0.3730	(0.0922 , 1.0493)
L-1/3-1/2 *‡	26.01	0.3128	(0.0372 , 0.8078)
L-1/6-3/4 *‡	21.77	0.3019	(0.0668 , 0.8762)
L-1/6-1/4 ⅂	18.89	0.4236	(0.0647 , 1.1407)
Order Best (1) *‡	16.25	0.2743	(0.0605 , 0.7853)
Order 3 *	72.62	0.3172	(0.1192 , 0.9606)
MA11 *	25.73	0.2468	(0.1490 , 0.8415)
MA5	41.97	0.1782	(0.2582 , 0.8678)

Indicates statistical significance based on paired t-test (p-value < 0.05) with * MA5, †MA11 and Order 3, ⅂L-1/3-1/2, + Seg 3, ‡ Seg 6 and L-1/6-1/2

Table 5.6— Sample statistics on surface profile characteristics' *%Ra* based on flat sample profiles (n = 1 except for characteristics [tall, short, wide, narrow] n = 2) *(right)*

	Mean (%)	Standard Deviation	Range
Roughness*			
Alternating	13.02	0.06	(0.0475 , 0.2582)
Random	20.98	0.12	(0.0372 , 0.4739)
Scaling			
Normal	13.77	0.13	(0.0432 , 0.4835)
Reduced 100x	13.77	0.13	(0.0432 , 0.4835)
Point Density			
50%	23.01	0.15	(0.0510 , 0.5553)
100%	20.98	0.12	(0.0372 , 0.4739)
200%	18.77	0.14	(0.0583 , 0.5200)
Abnormality Height*			
Tall (+60)	77.71	0.17	(0.4587 , 1.1407)
Short (+30)	62.08	0.19	(0.2897 , 0.9682)
None	20.98	0.12	(0.0372 , 0.4739)
Abnormality Width*			
Wide (1.5)	84.61	0.12	(0.6012 , 1.1407)
Narrow (0.75)	55.18	0.14	(0.2897 , 0.8261)
None	20.98	0.12	(0.0372 , 0.4739)

*Indicates statistical significance based on paired t-test (p-value < 0.05)

Of the 45 total t-tests conducted on the difference of means between fitting methods for flat profiles, 24 were considered statistically significant. Additionally, the *Order 3* fitting method had the highest mean, overall. When comparing the result of the flat surface profiles with results from sample profiles Type 1-9, *MA11* and *Seg 6* have a lower $\%R_a$ on average than *MA5* and *Seg 3* respectively. Similarly, the mid-point locus line with the width of 1/3 has a higher $\%R_a$ than the width of 1/6.

From this data, the roughness characteristic showed a significant difference in means between the alternating and random roughness with the $\%R_a$ being lower for the alternating roughness. Scaling the data appeared to have no effect on the $\%R_a$. It is also noted although the point density did not have a statistically significant difference between the means, the $\%R_a$ appears to increase as the point density increases. For all abnormality characteristics, height and width, the p-values from the paired t-test were less than 0.0001, which indicates a very strong statistical significance the means differ. In general, as the height and width of the abnormality increases, so does the $\%R_a$. Overall, the standard deviation for the flat surface profile characteristics is much lower in comparison to the other data sets. (A graphical representation of the results tables can be seen in Appendix B.)

V. Discussion

The slicing methodology proposed in this paper created a simple process for estimating the underlying geometry. By analyzing various fitting methods for slices with different geometries and characteristics, insight was gained into how well each method estimated the underlying geometry.

In general, the geometry of the part does not influence the effectiveness of the fitting method. On the Type 3 sample profile, the irregularity of the wavy surface curvature made it difficult for the various methods to fit to the surface. For the *Order 3* fitting method, the profile deviated significantly from the actual profile, as seen in Figure 5.9, which increased the mean and standard deviations for wavy profiles and the *Order 3* fitting method. However, if using a segmenting method, specific attention must be given to the bend radius of the part. When

fitting straight lines to curved surfaces, much like the triangulation process in rapid prototyping, error due to the chordal deviation between the curved and flat surfaces exist. This error will increase for any curved surface if the segment increases.

By comparing the $\%R_a$ within a specific method, ideal parameters can be determined. The results between the sample statistics on fitting methods for flat surface profiles and surface profiles Types 1-9 were compared to determine if specific experimental parameters yield consistently lower $\%R_a$ values. For the segmenting method, the greater number of segments led to a lower $\%R_a$. Logically, this could be attributed to smaller segment lengths, which more tightly fit to the actual curvature of the surface profile; however, as mentioned previously, the individual data points will have a greater impact on the segment if there are too few of them, which could skew the estimated underlying geometry in areas with sparse data or fail to eliminate any of the surface roughness. The mid-point locus line method had two parameters to evaluate: the width interval and the overlap of the window. As seen in the sample statistics for the fitting methods, the overlap did not significantly impact the $\%R_a$; however, the smaller width of the interval tended to yield better results in regards to the representation of the ideal underlying geometry. Finally, the moving average method tended to improve the fit to the ideal underlying geometry when the period was increased. Although these generalities regarding surface parameters hold true for the tested values, changes should be carefully considered prior to use as extreme changes in these parameters could have

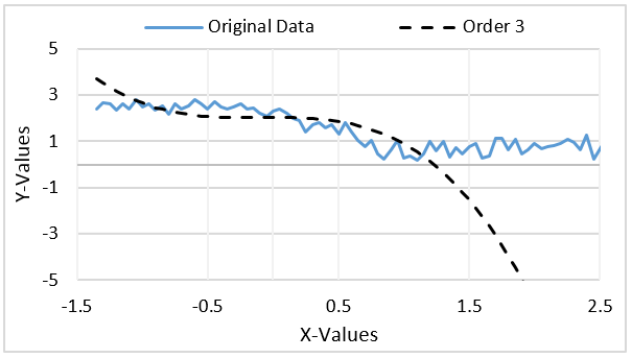


Figure 5.9—Third order polynomial fit to a wavy surface profile

detrimental results. Additionally, fitting methods may vary based on the slice length. To avoid poorly estimated underlying geometry, one must attempt to achieve slices over the entire surface that are similar in length and test the parameters of the desired fitting method on a sample profile in order to verify its acceptability.

End effects must be considered when using these fitting methods. Since the mid-point locus line method plots only data in the center of a sampling window, no data exists toward the ends of the surface profile. For the moving average methods, the true period of data at the end of the surface profiles is less than what was originally specified. This is because the average is taken using an equal number of points on each side of a specific data point, resulting in a one-sided average if there are only points on one side of the specific data point. It is common practice to fit a straight line to the ends of the data sets in order to minimize these effects [5].

The presence of abnormalities also strongly impacted the ability of the fitting method to estimate the ideal underlying geometry. When comparing sample statistics on the $\%R_a$ of abnormalities, surface profiles where abnormalities were removed had a mean over half of those without abnormalities. This can be attributed to the absence of data across the abnormal region, which reduces the variation across the sample profile falsely reducing the $\%R_a$. A drawback exists in the segmenting method, however, when a segment begins on the edge of a removed abnormality. If the slope of the segment is calculated from only a few data points that are continually increasing, it may result in a false feature, which will dramatically increase the $\%R_a$. An example of this from the Type 7 surface profile can be seen in Figure 5.10 where the segment was set based off of two data points. Besides this exception, generally surfaces

with the abnormalities removed more closely estimated the ideal underlying geometry than those with abnormalities. Similarly for flat surfaces, increasing the width and the height of the abnormality significantly increased the $\%R_a$. This is due to the fitting methods being skewed toward the outlier data in the abnormality.

Recall when evaluating the roughness profile the alternating roughness had a mean $\%R_a$ statistically lower than the random roughness profile. The cyclical effect caused when alternating between positive and negative values allowed fitting methods, such as the segmenting, midpoint locus, and moving average methods, to centrally focus since they were not skewed one direction or the other due to an increased number of data points falling on one side of the underlying geometry, also known as an abnormality. Based on this observation, fitting methods, such as the moving average method, may not seem desirable; however, if one implements multiple iterations of the moving average method, the effects of abnormalities on the estimated underlying geometry will be minimized. A drawback to the multiple iterations exists when a standard is not present compromising the repeatability and reproducibility of this method; this is similar to smoothing operations seen in commercial software, which can be highly manual and difficult to replicate.

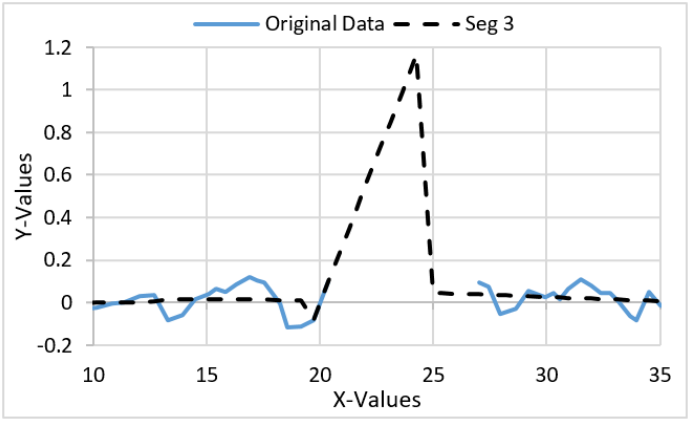


Figure 5.10— Segment set based on two data points near the removal of an abnormality in the abnormal region

In order to successfully implement slicing methods to estimate the underlying geometry, the user must not only keep in mind the benefits and drawbacks to each fitting

methodology, but they must realize the limitations of the methodology as a whole. These include the geometry orientation, the overfitting of data, and the presence of abnormalities.

Geometry Orientation

The orientation of the surface is important when slicing data for manipulation. In this study, slices were taken along the z-axis, and the cross-sectional area on the x-y plane was analyzed based on ordering the points by their x-values. Since the data enters the system as a point cloud, the order of the points is random, so ordering the points is necessary to complete the specific techniques covered in this paper that treat the data set as a function of points. The use of normal vectors would be beneficial in determining this orientation. In rapid prototyping, the interior of a slice is determined by observing the order of points in polygonal chains or using line crossing algorithms, which would allow for easy calculation of the normal vectors to the surface; however, often times surface scans consist of partial scans of the surface making it difficult to determine the normal vectors of the surface, since this data is typically not obtained through simple scans. An example of this can be seen in Figure 5.11. By viewing the point cloud in Figure 5.11a, one can assume the order of the points and geometry of the surface, Figure 5.11b. However, when choosing a default ordering technique, such as ordering by the x-values, a different geometry is portrayed, Figure 5.11c.

In addition, the orientation of the geometry must minimize the difference in slice profile length across the entire surface. For example, a square surface rotated at 45 degrees, or a diamond, will have high variation in slice length comparing the corners to the tangential. Since the corners will have less data in each slice, in some cases only a few data points, the

edge data may have to be discarded in order to maintain integrity of the estimated underlying geometry from the other slicing orientations.

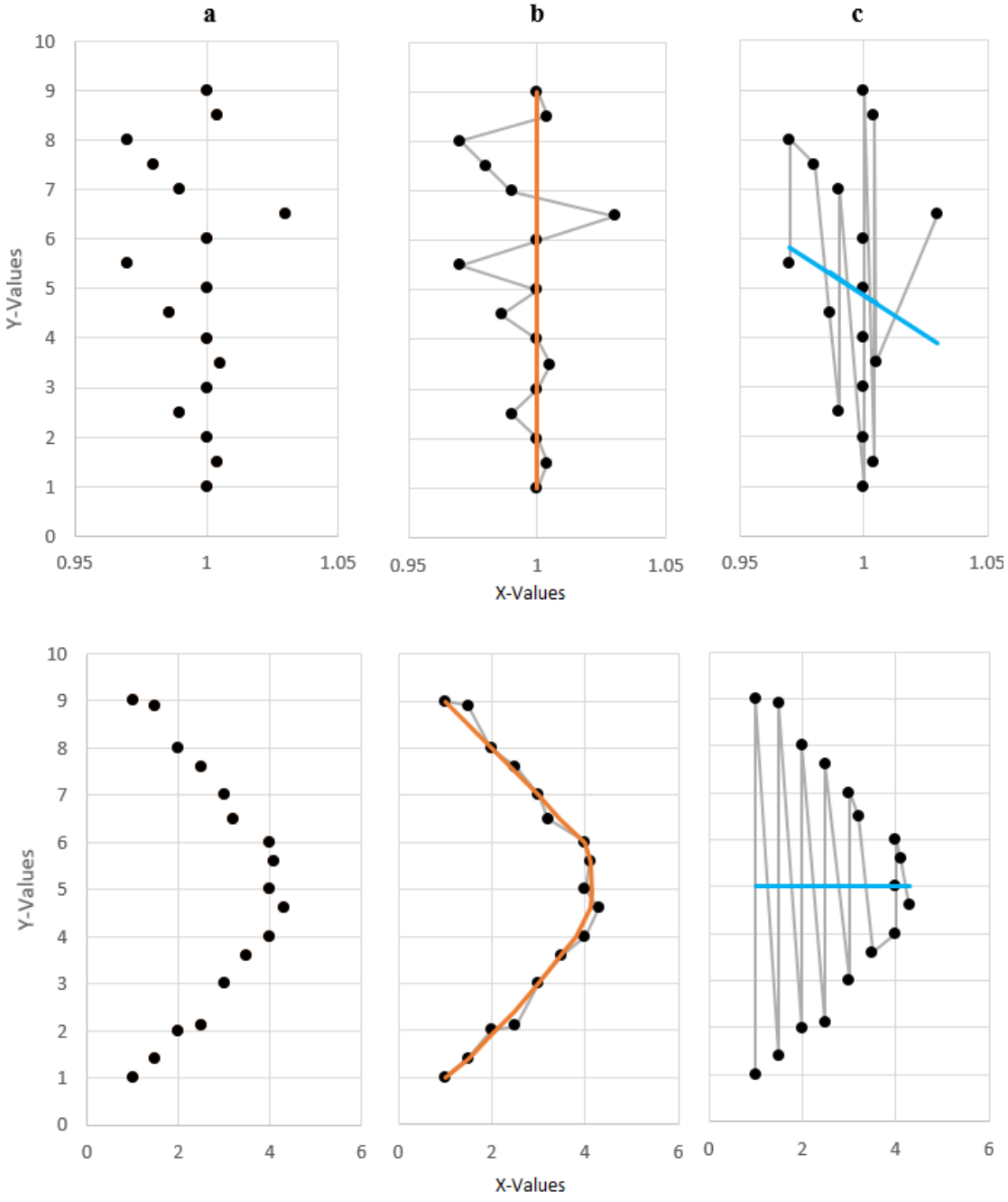


Figure 5.11—Effects of orientation on slices of the x-y plane a) original point cloud, b) fit based on desired geometry, c) fit based on orientation

Overfitting

A risk of the inherently rough surface profiles of cast surfaces is overfitting data. The random roughness and presence of abnormalities can skew the estimated underlying geometry calculated through various fitting methods. An example of overfitting can be seen in Figure 5.12. The moving average filter is easily influenced by series of abnormal points, which pull the estimated underlying geometry toward the point cluster. This effect can also be seen by fitting an increasing order of polynomial or a spline to the surface profile. If overfitting occurs, abnormalities may not be correctly identified causing unnecessary rework (false alarms) or unacceptable surface quality (misses) of the castings.

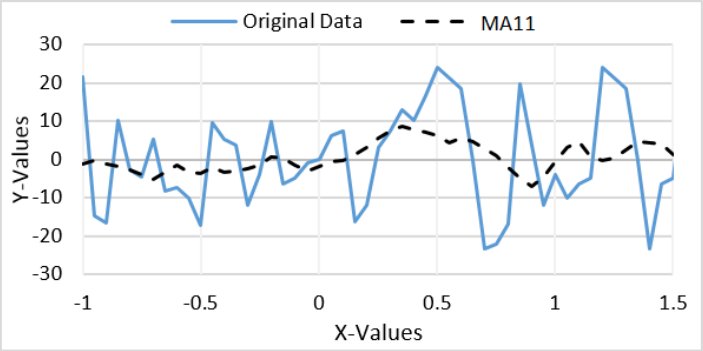


Figure 5.12—Overfitting of the *MA11* estimated underlying geometry from roughness and outliers

Abnormalities

As seen in the sample statistics for both Type 1-9 and flat sample profiles, the presence of abnormalities increased the mean $\%R_a$, regardless of fitting method. This indicates a single iteration of the proposed methods is not enough to reduce the effects of abnormalities, or clusters of outlier points, on the estimated underlying geometry. Therefore, it would be ideal if the surface analyzed was absent of abnormalities, or if abnormalities are present, they are removed. As demonstrated previously, the absence of data may cause some fitting methods' effectiveness in estimating the underlying geometry to decrease, as was demonstrated in Figure 5.10. Methods to identify abnormalities for removal prior to estimating the underlying

geometry, as well as to fill in the absent data with a curved surface representative of the actual surface, would allow for a more accurate representation of the ideal underlying geometry.

VI. Conclusions

The underlying geometry algorithms discussed in this paper will aid in the calculation of parameters for use in the digital surface inspection standard for castings. This paper has identified the benefits and drawbacks of the slicing methods presented. Utilizing slicing methods simplifies the surface into two-dimensional parts in order to reduce the complexity of calculations required in estimating the underlying geometry. This estimated underlying geometry will aid in the identification and measurement of abnormalities, and it will provide a consistent method in calculating the actual baseline roughness.

Future work of this research includes refining the slicing process, exploring alternative methods to estimate the underlying geometry, and identifying methods to eliminate abnormalities. To refine the slicing process, the tri-direction slicing method will be explored to analyze the effects of selecting the medial surface for estimated underlying geometry in addition to methods to orient the point cloud effectively for manipulation. A subsampling method to estimate the underlying geometry will also be explored. Additionally, various means of eliminating abnormalities and approximating absent data will be explored to improve the effectiveness of the fitting algorithms.

VII. Acknowledgements

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-033. The views and conclusions

contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

VIII. References

[1] Voelker, Michelle, Paul Kemper, and Frank Peters. *Development of a Digital Standard to Specify Surface Requirements.* Proc. of Steel Founders' Society of America—Technical & Operating Conference, The Drake, Chicago. 2015. 1-22. Print.

[2] Voelker, Michelle M., and Frank E. Peters. "Development of a Digital Standard to Specify Surface Requirements of Cast Metal Surfaces." *ASTM Materials Performance and Characterization Special Issue on Surface Texturing* (2016): (in review).

[3] Liao, Y.-S., and Y.-Y. Chiu. "A New Slicing Procedure for Rapid Prototyping Systems." *The International Journal of Advanced Manufacturing Technology* 18.8 (2001): 579-85.

[4] Thomas, T. R. *Rough Surfaces*. London: Longman, 1999. 122-74. Print.

[5] Whitehouse, David. "Characterization." *Handbook of Surface and Nanometrology*. 2nd ed. Bristol: Institute of Physics Pub., 2001. 12-21. Print.

[6] Bashiri, Mahdi, and Amir Moslemi. "A Robust Moving Average Iterative Weighting Method to Analyze the Effect of Outliers on the Response Surface Design." *IJIEC International Journal of Industrial Engineering Computations* 2.4 (2011): 851-62. Web.

CHAPTER 6: AN EVALUATION OF SUBSAMPLING METHODOLOGIES FOR UNDERLYING GEOMETRY APPLICATIONS IN CAST METAL SURFACES

A paper submitted to *International Journal of Cast Metals Research*.

Michelle M. Voelker[9] and Frank E. Peters[10]

Abstract

Due to the nature of the casting process, cast surfaces are inherently complex, and in reverse engineering applications, it is often difficult to model the actual geometry in absence of surface roughness, abnormalities, and shrinkage after the part is cast causing the actual geometry to differ from that of the original part model. This geometry, known as the underlying geometry, is used to evaluate parameters of the Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces: the baseline roughness, abnormality level, and abnormality percentage. Without the underlying geometry, a consistent manner to calculate these surface parameters does not exist. This paper evaluates subsampling methods as a potential means to estimate the underlying geometry of castings for use in the standard. Various considerations in the use of these methods to minimize the effects of abnormalities and roughness, while capturing the actual geometry and surface characteristics of the casting, are detailed.

[9] Primary author; Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa, 50011, United States; 0000-0002-3521-694X
[10] Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa, 50011, United States; 0000-0002-8998-0062

I. Introduction

The Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces was developed to quantify specifications for inspection [1, 2, 3]. In this standard, a digital representation of the cast surface is analyzed to calculate components of the Voelker Surface Ratio (VSR), including the baseline roughness, abnormality level, and abnormality percentage. Standard methods to clean the data and calculate these values are essential in reducing variation in this process. By laying out a process to calculate these values, the inspection process can be both repeatable and reproducible, so the variation among inspectors from the customer and manufacturer is minimized. This process requires a standardized means of calculating the underlying geometry, which is the actual geometry of the part following the casting process. This article explores the limitations and benefits of subsampling methodologies to determine the underlying geometry of metal castings for use in this standard.

II. Previous Work

The evaluation of slicing algorithms [4] to simplify the complexity of the cast surface was explored. In this work, point clouds were sliced into two-dimensional data sets at 0, 45, and 90 degrees to reduce the effects of linear defects. Four different methods with varying parameters were used to evaluate their effectiveness of estimating the true underlying geometry. These included segmenting, a mid-point locus, polynomial, and moving average fitting methods of the original data. Two to four variations of each were evaluated using different parameters. For the segmenting method, different segment lengths were evaluated. For the mid-point locus, different interval widths and overlaps were evaluated. For the

polynomial fit, a set order polynomial and a best fit were evaluated. Finally, for the moving average method, different periods were evaluated.

Surface profiles with a hand calculated ideal underlying geometry were used for evaluation. The first set of surface profiles included various geometries (flat, curved, and wavy surface profiles) and abnormalities (abnormalities not present, present, and removed). The second set of profiles were all flat and evaluated the effects of roughness, scaling, point density, abnormality height, and abnormality width on the effectiveness of the fitting methods.

Each fitting method was then evaluated against each surface profile. Then, the roughness average was calculated to compare the actual profile to the fitting method. Since the roughness varied from part to part, the roughness averages could not be compared across parts; however, the roughness averages of each method for the same part could be compared. To compare them across samples to identify an overall preferred method, a normalization of the roughness averages had to occur. To normalize these values, each roughness average was divided by the roughness average obtained from comparing the ideal underlying geometry to the actual surface profile. This resulted in a normalized roughness average in comparison to the ideal underlying geometry. For a normalized roughness average of 0.0 indicates the fitting method used is a perfect representation of the ideal underlying geometry; however, a normalized roughness greater than 1.0 indicates the fitting method is beyond the roughness average of the surface profile itself compared to the ideal underlying geometry. These values were averaged for each category of evaluation. The normalized values were then compared to each other in a paired t-test using a p-value of 0.05 to identify if there was a statistical significance between means.

The results of the study showed significant differences between certain surface profiles; however, an overall best fitting method was not apparent. Of the various surface profile characteristics, the following did not show a statistical difference between means: geometry, scaling, and point density. When evaluating roughness, an alternating, or cyclical, surface profile, often seen from machining operations, will allow for a better approximation of the underlying geometry compared to a random roughness profile, but this would not be feasible for estimating a casting surface due to the inherent variation in the surface due to the molding process. Abnormality height and width also showed a significant difference among samples. In general as the abnormality height and width increase, the less accurately fitting methods could represent the ideal underlying geometry. Additionally for the surfaces with varying geometry, it was found that removing abnormalities had a significant impact on the overall effectiveness of the fitting methods.

Additionally, several conclusions about the fitting methods could be drawn from this data. First for a polynomial method, it was not ideal to fit a set order of polynomial to the surface, since it resulted in a large quantity of error from the actual surface profile and ideal underlying geometry. Ideally, a polynomial could be manually fit to the data, but this manual intervention would be time consuming in a manufacturing environment. Next for the segmenting method, a larger number of segments led to a more accurate representation of the underlying geometry; however, it is important to note that if the number of data points in the segment approaches two, the minimum number of data points to form a segment, overfitting can occur. Likewise, with a moving average filter, overfitting can occur if the period is too short. Finally, for both the moving average and the mid-point locus fitting methods, end effects could occur skewing the ends of the data set. This could occur from the decreasing

period to one side of the moving average method or the start of the mid-point locus estimated line beginning and ending at the mid-point of the observation window. These parameters should vary based on the length of the data slice, so if a constant parameter is desired for a specified surface, all slices must be approximately the same length. All of these parameters must be monitored for use of these fitting methods to assure the estimated underlying geometry accurately represents the ideal.

From these results, conclusions were drawn about the overall effectiveness of slicing methods. These include the geometry orientation, overfitting of data, and the presence of abnormalities. Since point cloud data sets of surfaces do not have a preassigned order, the point cloud must be appropriately oriented such that the two-dimension slice on the x-y plane is a function. If this is not the case, the fitting methods could misinterpret this data, which would result in poor estimations of the underlying geometry. Second, overfitting the estimated underlying geometry to the original data could result in unacceptable abnormalities passing inspection. Finally, the presence of abnormalities skews these fitting methods so that the estimated underlying geometry is pulled toward the peak of the abnormality. To overcome these issues, abnormalities could be removed in order to reduce their effects on the estimated underlying geometry.

III. Methodology

The underlying geometry is used in calculating all components of the VSR. The ideal underlying geometry would remove all roughness and abnormalities of the surface resulting in a smooth model to accurately represent the complex geometry of the cast surface. This

surface will be compared to the original point cloud to calculate the components of the VSR: baseline roughness, abnormality level, and abnormality percentage.

The slicing methods provided one means in estimating the underlying geometry; however, they depended on the orientation of the geometry, which requires processing the data prior to the execution of these methods. Additionally, the slicing method required three slicing directions in order to cancel out the effects of linear indications that may fall parallel to the slice, which would result in skewed data due to the large width of the abnormality in this situation. Finding a method less reliant on pre-processing and requiring less repetitive computational analysis would be ideal for estimating the underlying geometry in a manufacturing environment.

This method explores subsampling methods to estimate the underlying geometry. This method was chosen based on its usability in comparison to the slicing method. This method aims to eliminate roughness and abnormalities from the original point cloud in order to estimate the underlying geometry. Evaluation of this method was based on its ability to remove roughness and abnormalities in comparison to the original point cloud.

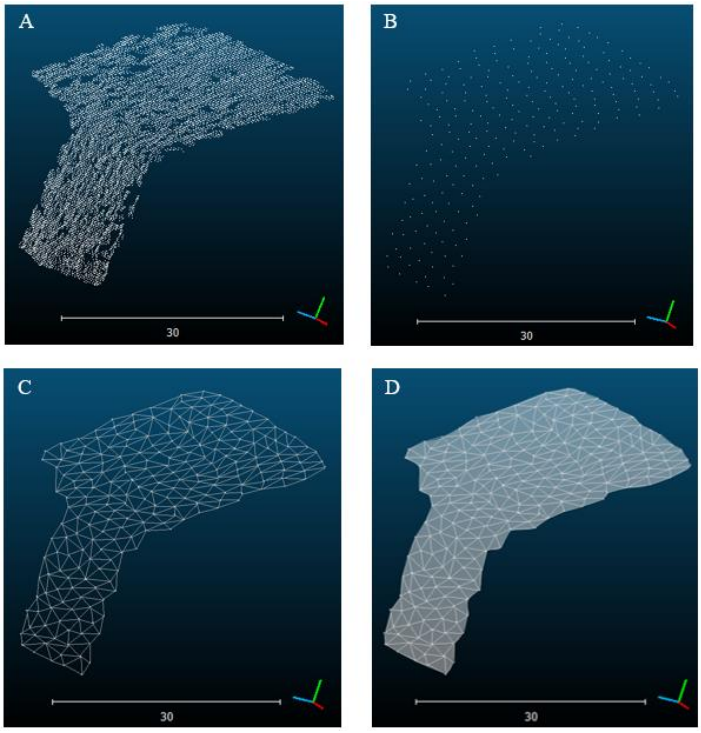


Figure 6.1—Subsampling method for calculating underlying geometry A) original point cloud, B) subsample point cloud, C) combine subsampled regions via triangulation, etc., D) create composite mesh by combining subsampled regions

The subsampling process breaks down the geometry of a part into smaller regions. A generalization of this process is broken down in Figure 6.1. First, the original point cloud is subsampled according to the desired method. Next, the subsampled regions are evaluated to provide an estimate of the underlying geometry in a specific region. Next, the underlying geometries from each region are combined to create a mesh of the entire surface.

Table 6.1—Part type based on geometry and abnormalities

		Flat Geometry	
		Yes	**No**
Abnormalities	**No**	Type 0	Type 1
	Yes	Type 2	Type 3

The surfaces used to evaluate subsampling fitting methods are from actual castings or plastic replications of castings. Five types of castings were evaluating, designated Types 0 through 4, with three castings for each type. Type 0 consisted of castings that were generally flat with no abnormalities. Type 1 consisted of unique geometry castings with no abnormalities. Type 2 consisted of generally flat castings with abnormalities. Type 3 consisted of castings that had unique geometries and linear abnormalities. Type 4 were industry casting examples with no specific geometry and abnormality requirements. These types are displayed in Table 6.1 based off of their geometry and abnormalities. Each part was classified by the type designation followed by the assigned part number from the set of three parts (*ie*: T0P2 is the second part of the Type 0 classification).

Since the Quantitative Inspection Acceptance Criteria for Cast Metal Surface calls for an inspected surface of 8 centimeters by 8 centimeters, these dimensions were used to tape off a square area on the surface of the part. The taped off area could not have any edges of the part, as an edge would indicate a boundary between two surfaces needing to be inspected separately. Once the cast surface was prepared, they were scanned manually using a Faro

Edge scan arm. Scans were taken on each part three times by one operator. This was to assure if incomplete data was collected, the other samples could serve as a back-up. The operator was required to take two complete scans of the part orthogonally to each other in the same data file to assure the density of data was high enough at all points to evaluate the data and to assure no holes were missing in the data. The scan data was saved as a point cloud text file and the most complete scan was chosen from the three samples for each part. The point clouds were then subsampled to a point density of 0.2 millimeters to eliminate redundant data and trimmed to remove areas that were scanned beyond the taped off region.

Various subsampling methodologies were explored to evaluate how each estimated the underlying geometry. The following subsampling methods were evaluated: point sampling, strategic sampling, geometric shape sampling, and grid sampling. These methods were chosen as a simple means of evaluating the effects of the roughness and abnormalities on the cast surface. The methods were treated as if they were the ideal underlying geometry of the casting, and the original point cloud of the cast surface was then compared to the underlying geometry as in the digital standard. The differences between the original point cloud and underlying geometry were calculated, along with the areal roughness average (S_a). The S_a is similar to the roughness average (R_a) in that an average of the absolute value of each data point from the underlying geometry is taken; however this calculation is conducted over the entire surface as opposed to a single, two-dimensional profile. Scalars were assigned to each point on the original point cloud to represent this difference, which were color coded for easier visualization. These representations will be referred to as “color-mappings” of the samples. These scalars include the following: blue for scalar < 0.0, green for scalar = 0.0, and red for scalar > 0.0 millimeters from the underlying geometry. From the physical casting,

areas considered abnormal or irregularities in surface roughness by visual inspection were noted. An attempt was made to identify these areas in each of the color-mapped samples to evaluate if the method was successful at detecting these surface characteristics. In theory, if these irregularities can be identified by a visual inspection, which is the current industry standard, they will need to be identifiable in the digital representation. Areas not identified in the visual inspection process but can be seen in the digital representation would not be considered crucially identifiable, and therefore, would be considered as unnecessary noise.

Point Sampling

For the point sampling method, the point cloud representation of the cast surface was reduced in order to smoothen the surface. First, the point cloud was be subsampled in order to reduce the number of data points. This intends to reduce the surface roughness from the data in order to give an approximate representation of the underlying geometry. Once the point cloud had been subsampled, the points will be transformed into a mesh creating a surface to compare to the original point cloud using two methods: triangulation and Poisson surface reconstruction.

Triangulation was utilized at point densities of 2.0, 5.0, and 10.0 millimeters. The original point density of 0.2 was also evaluated as a baseline to compare the reduced clouds. When color-mapped, the scalars assigned to this 0.2 millimeter cloud would be zero; since all the original points were used as vertices in the triangulation, there would be no difference between the original point cloud and estimated underlying geometry. Uniform point densities were used in order to create a watertight, triangulated mesh of the surface in commercial software. In this software, Delaunay triangulation was used as a global fitting method for the

point cloud. The Delaunay method uses a circle criterion such that “the circumcircle through the vertices of any triangle…will not include any other vertices” [5]. However, one limitation to this method is all points are used as part of the mesh, which can cause a rough surface representation since the vertices of the triangulated surface are rigid [6].

With the potential for rough surfaces in triangulation, screened Poisson surface reconstruction was also evaluated. This method used the same point clouds from the triangulation method in order to make a direct comparison of the color-mappings for evaluation. Prior to subsampling the point clouds, the normal vectors for each data point were estimated using commercial software. Unlike the triangulation methods discussed, the Poisson method uses both global and local fitting methods. As opposed to connecting the data points with straight lines, like the triangulated method, a b-spline function is fit between points based on their normal vectors to create a smooth transition between points, which give the surface approximation between points a more realistic representation of the cast surface geometry [7]. Additionally, not all the data points from the subsampled region are used in the final underlying geometry estimation. This is because the Poisson function utilized in this method weighs the individual data points in comparison to the surface gradients associated with the normal vectors to smoothen the surface; this smoothing identifies a single point and represents the surrounding points as a local plane [7].

Strategic Sampling

This method is similar to the point sampling method; however, obvious abnormalities were manually removed prior to subsampling. This method was only evaluated on Types 2 and 3 (parts with abnormalities). In order to locate the abnormalities, the original cast surface

was observed and perceived abnormalities were identified. These abnormalities were then located on the digital representation of the surface, and the data points associated with these abnormalities were removed. After the abnormalities were removed, the same methods used for point sampling were followed to determine if this had any effect on estimating the underlying geometry and identifying abnormalities.

Geometric Shape Sampling

The previous work discussed in the slicing methodologies [4] fit a single geometric shape to the entire surface. In this method, the surface will be fit to numerous geometric shapes using random sample consensus (RANSAC), which fits geometric shapes to noisy data even with many outliers [9]. This process is summarized in the following paragraph.

The process begins with a set quantity of geometric shapes; these include plane, cylinder, sphere, cone, and torus geometries. Then, the parameter for minimum points to be sampled in order for a geometric shape to be considered was set arbitrarily at 1000. This means a geometric shape must be able to account for at least 1000 points in the cloud to be selected for use. The RANSAC algorithm then randomly samples the point cloud assigning an estimated normal for each point. These normal vectors are then compared to the various geometric shapes and scored based on the ability of the shape to fit the points. By manipulating the parameters of each shape (*ie*: radius of the sphere), the angle between each point and the geometry is minimized. Through each iteration, the highest scoring shape is selected, and the points used to determine this shape are eliminated from the next iteration. This process is continued until all data points have been explained by a geometric shape. By

comparing each set of points to the geometric shape they explain, a color-map can be created, and all subsampled regions can be combined to create a composite of the surface.

Grid Sampling

Similarly to the geometric shape sampling method, grid sampling identifies a best fit geometric shape (plane) to represent the data; however, small patches on the surface are analyzed as opposed to the entire surface. This method breaks the surface into a grid and analyzes each grid patch as a separate surface. The grid parameters evaluated over the surface were 4 by 4 units (2.0 square centimeter patches), 8 by 8 units (1.0 square centimeter patches), and 16 by 16 units (0.5 square centimeter patches). Similarly to a proposed method for die castings [10], a plane is fit using the least-squares method to each patch on the grid.

IV. Results and Discussion

Comparisons of each of the original point clouds to the estimated underlying geometry were analyzed to determine the benefits and drawbacks of each subsampling method. S_a values were not used to draw conclusions about how well the estimated underlying geometry compared to the ideal underlying could not be made, since the ideal underlying geometry was not known. In fact, these values provided no insight into the comparison of the methods since low values did not necessarily mean the estimated underlying geometry was a good representation of the ideal underlying geometry, but instead low values indicated if the estimated underlying geometry was a good representation of the original point cloud itself. Therefore, the visual comparisons were used to evaluate the effectiveness of the methods to identify abnormalities and roughness.

As discussed previously, visual inspection was used to identify areas of interest, including roughness and abnormalities, of the physical cast surface. Go and no-go criteria were assigned based on the whether or not the subsampling method could detect each area of interest noted from visual inspection. Areas were also noted where indications appeared on the color-mappings but were not identified through visual inspection. An example of this can be seen in Figure 6.2. In each case, the cause behind why the area of interest did or did not appear on the color-mapping was identified through further investigation. This was done for both the areas of interest identified and not identified through visual inspection.

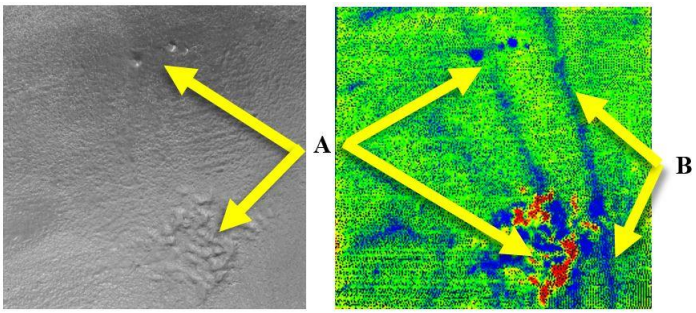


Figure 6.2—Example of areas of interest A) identified through visual inspection and B) not identified through visual inspection on a color-mapping of T2P3 5.0 millimeter Poisson sample

Subsampling reduces the complexity of the cast surface; however, some risks and limitations exist based on the method chosen. These include the point density, abnormalities, and sampling method of the cast surface.

Point Density

An obvious risk to the point sampling method is determining the ideal amount the point cloud should be reduced. If it is reduced too much, there may not be enough data points left to construct an accurate estimate of the underlying geometry, especially for complex geometries. A color-mapping of a part with complex geometry, a fairly smooth surface, and one abnormality can be seen in Figure 6.3. As the point density increases, the ability of the estimated underlying geometry to define the actual geometry of the part improves. The red

and blue curved regions on the upper and left areas of the color-mapping indicate the estimated underlying

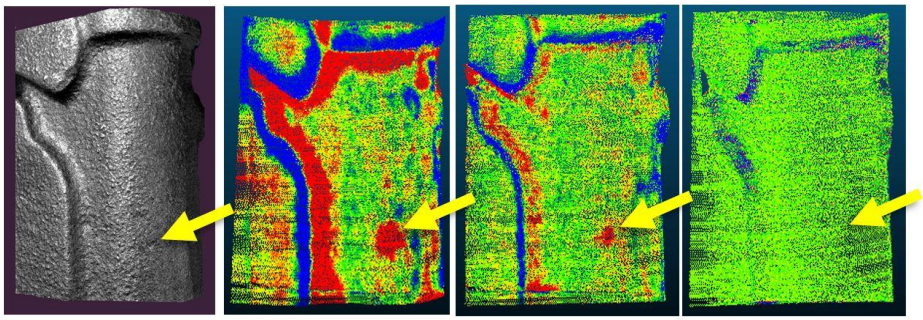


Figure 6.3—Color-mapping of complex geometry part (T4P2 Poisson sample) at point densities of 5.0, 2.0, and 0.2 millimeter (left to right) with arrow indicating the location of a potential abnormality

geometry is not accurately representing the sharp curvature of the part; however, the gradual curvature of seen on the right area of the part appear to be fitting the estimated underlying geometry better since there is less red and blue in this region. In contrast, if the cloud is not reduced enough, there will still be inherent variation in the mesh, which is not representative of the underlying geometry and will prevent abnormalities from being identified, as indicated with the arrows in the figure.

Increased noise can occur in the color-mappings of surfaces with lower point densities. As seen in Figure 6.4 moving from the original point density to a lesser one, more surface indications of potential abnormalities appear. This is because as the point density decreases, interpolation of the surface must occur between points; for example with the triangulation method, the interpolation is based off of the facets between data points. This could be accommodated for by pursuing non-uniform subsampling methods with varied triangle sizes to more accurately represent the actual surface geometry [9]. On the other hand moving from a low point density to the original, less indications are present; in fact on the original cloud, no indications are present. This is because the estimated underlying geometry used all of the original points to form a surface. Therefore, the point cloud must be reduced to a point density

less than that of the original point cloud, or the mesh constructed based off of these points will show no difference compared to the original point cloud.

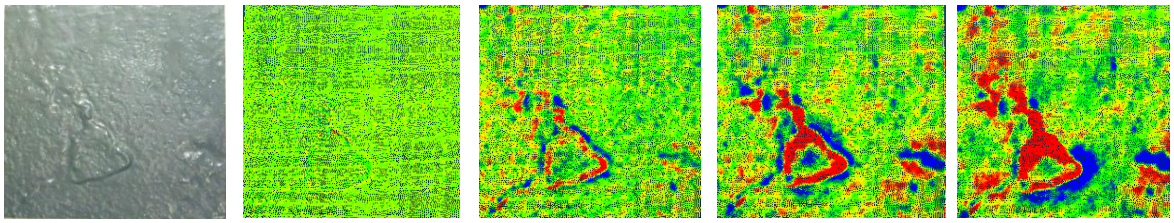


Figure 6.4—Picture of T2P1 casting (left) and color-mappings of triangulated underlying geometry estimates at point densities 0.2, 2.0, 5.0, and 10.0 millimeters (left to right)

Abnormalities

There are inherent risks associated with estimating the underlying geometry using subsampling methods when abnormalities are present. If a sampled point is located on an abnormality, the estimate of the underlying geometry will be skewed toward it, and the abnormality will not appear as severe as it really is, if it appears at all. An example is shown in Figure 6.5. When the abnormality was present during the estimation of the underlying geometry, points on the abnormality were chosen as part of the point sampling process. When these points were used during surface reconstruction of the underlying geometry, they were skewed toward the abnormality, which caused the difference between the original point and the estimated geometry to be minimal and made the abnormality appear less severe as seen in the left color-mapping. When points on the abnormality were not included in the sampled points, the abnormality was more visible and its actual geometry was more accurately represented in the right color-mapping. This shows that in order to use subsampling methods effectively, the

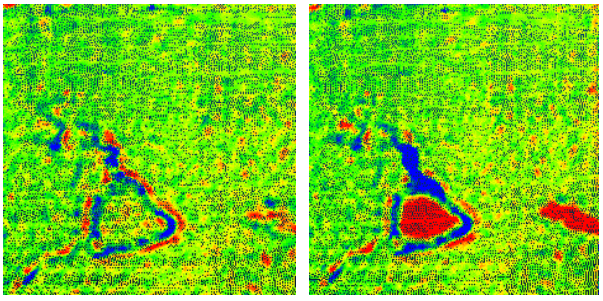


Figure 6.5—Example of point sampling of T2P1 Poisson subsampled at 2.0 millimeters with abnormalities present (left) and abnormalities removed (right)

abnormalities must be removed or avoided when subsampling prior to estimating the underlying geometry.

Sampling Method

Limitations exist with each subsampling method in regards to the random selection of points on a surface. These points play an important role in the estimation of the underlying geometry, and are key factors in the elimination of roughness and abnormalities.

The selection of points in the sampling process is important to get an accurate estimate of the underlying geometry. Local minima and maxima on the surface can skew the underlying geometry estimate if they are selected though point sampling. These local extrema can cause high variation in the underlying geometry, which gives inaccurate results when calculating the baseline roughness. Additionally, if only minima are selected (or maxima), the variation in the underlying geometry will be minimized; however, sampled points that are not minima (or maxima) will appear as if they are outliers. An example of this can be seen in Figure 6.6. A majority of the surface appears blue on the color-mapping indicating the scalars are less than 0.0. This is because when the surface was subsampled, only the high points on the surface were selected, which caused the rest of the surface to be negatively offset from the estimated underlying geometry. This could be avoided by taking local averages of points on the surface to reduce the effects of these local extrema. Also as mentioned previously, if subsampled points are located on an abnormality, it will skew the underlying geometry for that location and cause issues when calculating abnormality levels since they will more likely be considered a part feature as opposed to an abnormality. To avoid this, sampled points must not exist on the abnormalities, or the abnormalities must be removed prior to sampling.

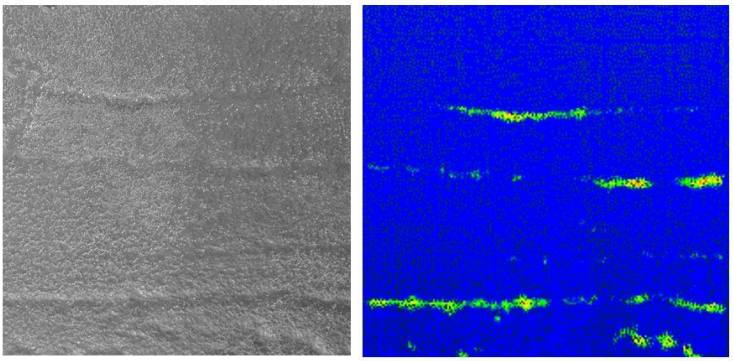


Figure 6.6—Example of locally extreme points being selected on the color-mapping of T3P3 Poisson with 10.0 millimeter point density and abnormalities removed

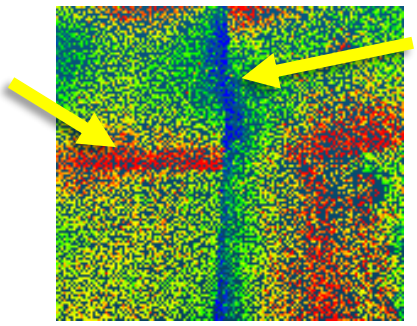


Figure 6.7—Example of discontinuities between grid patches on color-mapping of T1P1with two square centimeter patches

Similarly to the sampling process issue, the grid sampling method will not eliminate the effects of abnormalities from the individual grid patches. This is due to the method's inability to differentiate abnormalities from surface features. Additionally, the grid size was shown to have a major effect on the estimated underlying geometry. First, the grid size needed to be maximized to decrease the effects the roughness had on the plane. The smaller the patches are, the more sensitive the best fit plane would be to outliers. Second, the grid size needed to be minimized to more accurately represent the curvature of the cast surface geometry. To accommodate for the complex surfaces of the castings, b-spline surfaces could be utilized to smoothen the overall estimated underlying geometry model; however, it is important to note b-spline surfaces may not eliminate all of the noise caused by the surface roughness due to the order of b-spline used for fitting (high order is less sensitive to noise) and the size of the patch (larger patch increases the number of control points on the b-spline) [11]. A study on parametric surface reconstruction [12] analyzed surface patches in this manner using Gaussian processes to model in Euclidean space, which allows for data interpolation for surface computations, much like the subsampled surfaces discussed in this paper. This process used "isometric mapping" in order to essentially unroll the shape to reduce

its complexity without changing the spatial relationship between points [12]. This method was not previously explored since it cannot accommodate for extreme changes in the surface, such as abnormalities; however, these method could be explored further if abnormalities are removed prior to data manipulation. Finally, continuity between surfaces on the grid is ideal to prevent inconsistencies at grid boundaries, as seen in Figure 6.7. This could be corrected by taking a more global approach through an analysis of surrounding surface patches. For instance in a study of surface reconstruction of bone point clouds, the patches used tangential vectors along the boundaries to achieve C^1 continuity for b-spline patches [13].

For the RANSAC method, the surface was subsampled by regions of similar shape. This method proved to be highly variable as a random sampling of points in all samples resulted in different geometric shapes each time. Additionally, as discovered in previous work where only one geometric shape was selected [4], the complexity of the surface curvature was not able to be accurately described by these shapes. The parameter for the minimum required points to fit a shape to was modified; however, an ideal setting was unable to be determined to give an accurate estimate of the underlying geometry. When the number of points was increased, less shapes were fit to the surface and the results approached one shape as studied in the previous work [4]; however, when the number of points was decreased, many shapes were fit to the surface preventing identification of the roughness and abnormalities. These algorithms are intended to explain a point cloud as a series of geometric shapes in the presence of noisy data [8]; however, this benefit to reverse engineering parts to an model similar to the original computer model does not take into account the complex geometry of the surface and shrinkage post molding.

VI. Conclusions

The evaluation of subsampling methods discussed in this paper will aid in the advancement of underlying geometry estimation of cast metal surfaces. This paper has identified the benefits and drawbacks of the subsampling methods presented. Subsampling methods can be used reduce the complexity of calculations required in estimating the underlying geometry and provide a standard means of identifying and measuring abnormalities and the actual roughness.

Future work of this research includes refining the subsampling process, identifying methods to eliminate abnormalities, and exploring the feasibility of automating underlying geometry algorithms. To refine the subsampling process, traditional methods for smoothing rough surfaces will be incorporated to reduce variability from randomly selecting local maxima or minima. From there, a complete surface evaluation process will be completed to compare to the slicing methodologies [4]. Additionally, various means of eliminating abnormalities and approximating absent data will be explored to improve the effectiveness of the fitting algorithms. A semi-automated inspection process to enhance the visual inspection qualification process and verify acceptance of visually inspected surfaces will be explored in coordination with a software program to calculate the VSR parameters in order to give feedback to the user in order to aid in selection of surface criteria, identification of rework areas, and calibration of inspectors.

VII. Acknowledgements

Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-033. The views and conclusions

contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

VIII. References

[1] ASTM International. Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces. West Conshohocken: ASTM Int'l, 2016. *Appendix 1.*

[2] Voelker, Michelle, Paul Kemper, and Frank Peters. *Development of a Digital Standard to Specify Surface Requirements*. Proc. of Steel Founders' Society of America—Technical & Operating Conference, The Drake, Chicago. 2015. 1-22. Print.

[3] Voelker, Michelle M., and Frank E. Peters. "Development of a Digital Standard to Specify Surface Requirements of Cast Metal Surfaces." *ASTM Materials Performance and Characterization Special Issue on Surface Texturing* (2016): (in review).

[4] Voelker, Michelle M., and Frank E. Peters. "Evaluation of Slicing Methodologies to Determine Underlying Geometry of Cast Metal Surfaces." (2016): (in review)

[5] Loze, Michael K., and R. Saunders. "Two Simple Algorithms for Constructing a Two-dimensional Constrained Delaunay Triangulation." *Applied Numerical Mathematics* 11.5 (1993): 403-18. Web.

[6] Kazhdan, Michael, Matthew Bolitho, and Hugues Hoppe. *Poisson Surface Reconstruction.* Proc. of Eurographics Symposium on Geometry Processing, Sardinia. 2006. 1-10. Print.

[7] Kazhdan, Michael, and Hugues Hoppe. "Screened Poisson Surface Reconstruction." TOG ACM Trans. Graph. *ACM Transactions on Graphics* 32.3 (2013): 1-13. Web.

[8] Schnabel, R., R. Wahl, and R. Klein. "Efficient RANSAC for Point-Cloud Shape Detection." *Computer Graphics Forum* 26.2 (2007): 214-26. Web.

[9] Scheidegger, Carlos E., Shachar Fleishman, and Claudio T. Silva. *Triangulating Point Set Surfaces with Bounded Error*. Proc. of Eurographics Symposium on Geometry Processing, Austrian Academy of Sciences, Vienna. U of Utah, 2005. Print.

[10] Bračun, Drago, Valter Gruden, and Janez Možina. “A Method for Surface Quality Assessment of Die-castings Based on Laser Triangulation.” *Measurement Science and Technology* 19.4 (2008): Web.

[11] Zeid, Ibrahim. *Mastering CAD/CAM*. Boston: McGraw-Hill Higher Education, 2005. 207-16. Print.

[12] Castillo, Enrique Del, Bianca M. Colosimo, and Sam Davanloo Tajbakhsh. “Geodesic Gaussian Processes for the Parametric Reconstruction of a Free-Form Surface.” *Technometrics* 57.1 (2014): 87-99. Web.

[13] Yoo, Dong-Jin. “Three-dimensional Surface Reconstruction of Human Bone Using a B-spline Based Interpolation Approach.” *Computer-Aided Design* 43 (2011): 934-47. Elsevier. Web. 1 Jan. 2015.

CHAPTER 7: GENERAL CONCLUSIONS

General Discussion

Inspecting castings to verify the quality of a part is critical for foundries to maintain a high level of customer confidence. Current methods to evaluate cast metal surfaces require an inspector to visually and tactilely inspect a part to identify surface abnormalities, such as porosity and inclusions, and determine if they are acceptable. The manufacturer must correctly interpret the inspection criteria set by the customer in order to meet the design specifications. Current visual inspection standards are qualitative and make it difficult to interpret these standards. Often times, the interpretation of what is acceptable differs from the customer to manufacturer and even from inspector to inspector. In this thesis, the visual inspection of cast metal surfaces was explored in depth, and improvements to current methods were proposed.

Understanding sources of error in the visual inspection process allows for management to improve and monitor environmental and human factors with the most impact; however, there still exists a high variation among inspectors due to the subjectivity of the standards. In Chapter 3, the risk of Type I and II errors were evaluated based off of varying states of environmental and human factors in the inspection process. Human capabilities and environmental factors had the greatest impact on the overall errors; however, each individual factor contributing to these areas has little impact by itself. In the worst case scenario, the probability of a Type I error is 92.9%, while the probability of a Type II error is 82.3%. Even under the best case scenario, the Type I and II errors are 17.8% and 29.8%, respectively. This signals a need for a more quantitative standard to evaluate the surface of a casting.

In Chapter 4 digital standard is proposed, which specifies three parameters to allow the customer to communicate their exact needs in regards to surface finish to the

manufacturer: the baseline roughness, the abnormality level, and the abnormality percentage. These parameters are calculated based off of a part's true geometry post shrinkage in absence of surface roughness and abnormalities, or underlying geometry.

Since the underlying geometry differs from the part's intended geometry, or computer modelled geometry, methods were explored in Chapter 5 to estimate the underlying geometry from a point cloud of the part's surface. In the slicing method, geometry orientation was an important factor to improve the accuracy of the fitting method to the ideal underlying geometry. Additionally, the risk of overfitting to the point cloud was identified, and the presence of abnormalities proved to dramatically skew the data, which hindered the fitting process. The proposed methods were compared and contrasted in order to identify which approach should be explored further to calculate the ideal underlying geometry. Once an ideal method is identified, it will be used as a standard method to calculate the underlying geometry in order to create consistency among inspectors at both the customer and manufacturer.

The work completed in this thesis will raise awareness of the risk associated with current visual inspection methods for cast metal surfaces. The new, digital standard will reduce the variation in this inspection process allowing greater confidence in the parts leaving the manufacturer. Additionally, the standard will allow the customer to improve communication with the manufacturer in order to achieve the quality of surface required for their specific needs.

Recommendations for Future Research

Three main focuses exist for future research in the development of the digital standard for cast metal surface inspection. These areas include completing of the standard for

implementation in industry, refining methods to estimate the underlying geometry, and developing user-friendly tools to help the customer and manufacturer understand the specification criteria.

First, the standard must be completed and introduced to industry partners. Information regarding the ideal point density, or resolution, required to identify areas of interest on varying levels of surface roughness are currently being explored to include as guidelines in the standard, seen in Appendix A. In addition, industry samples from different molding processes are being evaluated in order to provide suggested baseline roughness criteria for customers to specify. The Steel Founders Society of America (SFSA) has been working in partnership with the research team at Iowa State University. Feedback from industry partners at SFSA meetings and conferences allowed the development of the standard to be collaborative in order to increase buy-in from the standard's users and improve the chances of a successful implementation into industry.

Next, methods to estimate the underlying geometry need to be refined. A standardized process, independent of the digital standard, can be developed in order to reduce the variation resulting from the use of different methods to determine the underlying geometry. This most likely will include the exploration of methods to remove abnormalities prior to calculating the underlying geometry, since the presence of abnormalities has proven to skew the estimated underlying geometry, or an iterative process to reduce the effects of abnormalities on the underlying geometry. Once the ideal method is established, other standards can be converted to the digital standard notation in order to ease companies' transitions to the digital standard. A place holder for these conversions can be seen in the Quantitative Inspection Acceptance Criteria for Cast Metal Surfaces in Appendix A.

Finally in order for these algorithms to be feasible in a manufacturing environment, two tools should be developed. The first tool is a software that will aid the customer in selecting a specification criteria for their parts. The second tool will be used by the customer and manufacturer for training and verification purposes.

First, a software program can be developed for customers to see how their specification criteria can be interpreted and to train inspectors on the new standard. The software will take a point cloud, determine the underlying geometry, and compare the original point cloud to the underlying geometry in order to calculate the parameters of the VSR. This information will then be used with the new quantitative surface inspection in order to provide outputs of roughness and abnormality levels. The program should have the ability for the user to change the VSR input values to allow them to visually see the differences between specification criteria on a sample scanned surface. Ideally, the parameter adjustments would color code the parameters using three colors to identify the following data points: 1) points considered the baseline roughness, 2) points considered abnormalities, and 3) points outside the acceptable range considered unacceptable based on the specified criteria. The proposed tool is outlined in Appendix C.

Second, a scanning tool could be developed in order to instantaneously scan a surface of a part in question and notify the user if the part is acceptable or not under specific criteria. The handheld tool will take a small sample scan of a part surface and compare it to a given specification criteria set by the user. After scanning the part, the device will tell the user whether or not the area is considered acceptable in terms of the abnormality level and baseline roughness. Alternatively, the user could scan part of the surface without abnormalities to check and set the baseline roughness, and then scan an area with an abnormality to determine

what the maximum abnormal point is to determine if the abnormality needs to be reworked in order to be deemed acceptable under the acceptance criteria.

This future research will improve the inspection process for cast metal surfaces. Once the digital standard and ideal method to determine the underlying geometry is complete, implementation will begin at various companies. As with any new product, method, or idea, there will be early adopters who will highly influence the future of the standard, and others may choose to use current methods until the improvement in results at other companies is noticed. To ease implementation, suggested conversions from the current standards and roughness values based on the molding process will be provided. Additionally, the development of user-friendly tools will allow for the user to accurately assign the acceptance criteria they desire, act as a referee tool in training visual inspection operators, and provide a way of quickly checking the part for acceptability. All of these research areas will aid in achieving the main goals of decreasing the variability in cast metal surface inspection and improving communication between the customer and manufacturer.

APPENDIX A: QUANTITATIVE INSPECTION ACCEPTANCE CRITERIA FOR CAST METAL SURFACES

Designation: AXXX -16

1. Scope

1.1 This standard covers the quantitative inspection acceptance criteria for the surface of metal casting.

1.2 The acceptance criteria is based on three main classification levels: the baseline surface roughness, abnormality level, and abnormality percentage.

1.3 This standard includes additional requirements in annexes:

Suggested Acceptance Levels for Cast Surfaces	Annex A1

1.4 This standard includes additional requirements in appendixes:

Corresponding Conversions from C-9 Standard (ANSI/ASME B46.1)	Appendix X1
Corresponding Conversions from SCRATA Standard (A802/A802M)	Appendix X2

1.5 Descriptions of terms related to this standard are in Section 2.

1.6 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and/or health practices to determine the applicability of regulatory limitations prior to use*.

2. Terminology

2.1 *Definitions for terms specific to this standard:*

2.1.1 *abnormality, n*—any anomaly of the surface, not part of random variation due to surface roughness, exceeding the baseline roughness. Any point on the surface of the part exceeding twice baseline roughness is considered abnormal.

2.1.2 *abnormality percentage, n*—the maximum percentage of the specified surface area that contains abnormalities falling within specification.

2.1.3 *baseline roughness, n*—the roughness average of the cast surface disregarding any form of abnormality present.

2.1.4 *underlying geometry, n*—the surface geometry in absence of random variation due to surface roughness and abnormalities. The underlying geometry may differ from the intended geometry due to contraction and other dimensional changes.

2.1.5 *VSR, n*—the Voelker Surface Ratio is the specified ratio of baseline roughness to abnormality level to abnormality percentage for cast metal surfaces. The VSR specification on a print shall be noted, "VSR [baseline roughness] – [abnormality level] – [abnormality percentage]."

3. Ordering Information

3.1 The inquiry and order should specify the following information:

3.1.1 *Acceptance Level*—A single acceptance level can be specified for the entire casting surface, or multiple acceptance levels may be specified for different locations on a single part based on part function.

3.1.1.1 *Baseline Roughness*—The overall surface roughness shall be specified, in millimeters, by its maximum acceptable value for use.

3.1.1.2 *Abnormality Level*—The abnormality level represents the absolute distance of an abnormality from the underlying geometry. Abnormality levels shall be specified by its maximum acceptable value for use. If an abnormality level is not specified, a default value of two times the baseline roughness, over the entire specified surface, shall be assigned.

3.1.1.3 *Abnormality Percentage*—The abnormality percentage shall be specified by the purchaser as a percentage noted by a number between 0 and 100. If an abnormality percentage is not specified, a default value of 5, over the entire specified surface, shall be assigned.

3.2 The specification shall be noted with the VSR value. *Example: VSR 0.03 – 0.08 –2 would indicate a baseline roughness of 0.03 millimeters, and abnormalities up to 0.08*

millimeters across a maximum of 2% of the inspected surface.

3.2.1 If the abnormality percentage and/or the abnormality level are not specified, the default values of 5 and twice the baseline roughness shall be assigned, respectively, and they may be omitted from the specification notation. *Example: To indicate a baseline roughness of 0.01 millimeters, and abnormalities up to 0.02 millimeters across a maximum of the default value of 5% of the inspected surface, any of the following notations are acceptable:*

VSR 0.01 – 0.02 – 5
VSR 0.01 – 0.02
VSR 0.01

4. Dimensions, Mass and Permissible Variations

4.1 Dimensional accuracy, resulting from contractions and other variables, is not addressed in this standard.

4.2 This standard only considers the surface roughness that is calculated from actual underlying geometry of the casting.

4.2.1 Baseline roughness shall be calculated with the underlying geometry considered nominal.

4.2.2 Abnormalities shall be measured based on the deviation from the underlying geometry.

5. Acceptance Standards

5.1 Suggested levels of acceptance for this standard are found in Annex A1.

5.2 *Point densities for scanned surfaces are being evaluated in order to require a minimum point density for the standard for consistency purposes.*

5.3 Surface criteria exceeding those covered by this standard shall be a matter of agreement between the purchaser and the manufacturer.

6. Conversion from Other Surface Inspection Standards

6.1 The conversion to other standards is approximate and includes the C-9 Cast Microfinish Comparator and the Steel Castings Research and Trade Association (SCRATA) Comparator Plates.

6.1.1 Conversions for the C-9 Standard are listed in Appendix X1.

6.1.2 Conversions for the SCRATA Standard are listed in Appendix X2.

7. Keywords

7.1 castings; inspection standards; quantitative

8. Supplemental Requirements

The following supplementary requirements shall apply only when specified by the purchaser in the purchase order or contract.

8.1 Unusual visual conditions are not addressed by this standard.

8.1.1 Unusual visual conditions include discoloration and evidence of rework.

8.1.2 Unusual visual conditions shall be a matter of agreement between the purchaser and the manufacturer.

8.2 Chaplets, or inserts, are not permissible under this standard.

8.2.1 Requirements pertaining to chaplets or inserts shall be a matter of agreement between the purchaser and the manufacturer.

ANNEXES

(Mandatory Information)

A1. Suggested Acceptance Levels for Cast Surfaces

Industry samples from different molding processes are being evaluated in order to provide suggested baseline roughness criteria for customers to specify.

APPENDIXES

(Non-mandatory Information)

X1. CORRESPONDING CONVERSIONS FROM C-9 STANDARD (ANSI/ASME B46.1)

X1.1 The following gives approximate conversions corresponding to the C-9 Microfinish Comparator.

X1.2 Listed equivalents are approximate and are based on averages of three profilometer readings from three different locations on the available standard comparator. These values will be set once a standardized procedure for determining the underlying geometry is finalized.

C-9 Comparator Number	Standard Equivalent
20	VSR 0.76 – 1.52 – 0
60	VSR 1.19 – 2.38 – 0
120	VSR 1.37 – 2.74 – 0
200	VSR 1.41 – 2.82 – 0
300	VSR 2.32 – 4.64 – 0
420	VSR 2.15 – 4.30 – 0
560	VSR 1.88 – 3.76 – 0
720	VSR 2.20 – 4.40 – 0
900	VSR 2.08 – 4.16 – 0

X2. CORRESPONDING CONVERSIONS FROM SCRATA STANDARD (A802/A802M)

X2.1 The following gives approximate conversions corresponding to the Steel Castings Research and Trade Association (SCRATA) Comparator Plates.

X2.2 Listed equivalents are approximate and are based on averages of three profilometer readings from three different locations on the available standard comparators. These values will be set once a standardized procedure for determining the underlying geometry is finalized.

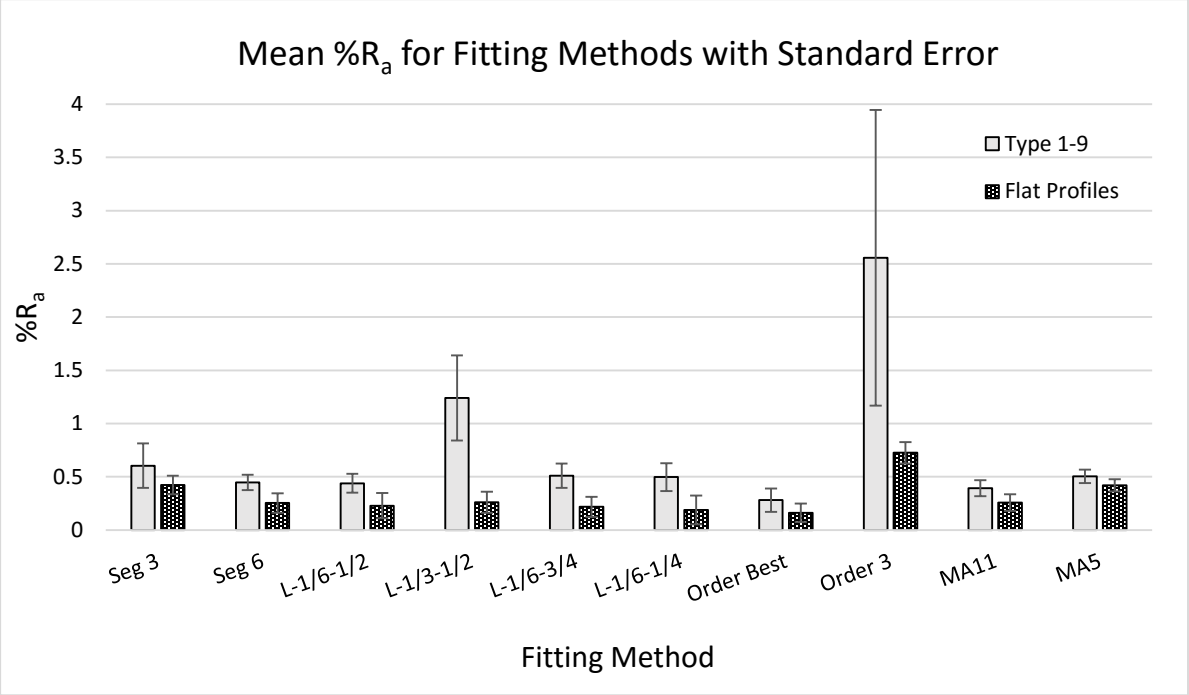
X2.3 Abnormality levels take into consideration the point with the maximum deviation from the underlying geometry, either peak or valley, regardless of whether or not the abnormality is that of the stated "surface feature" of the corresponding standard. *Example: Maximum deviation for comparator J1 results from a nonmetallic inclusion, not from a weld.*

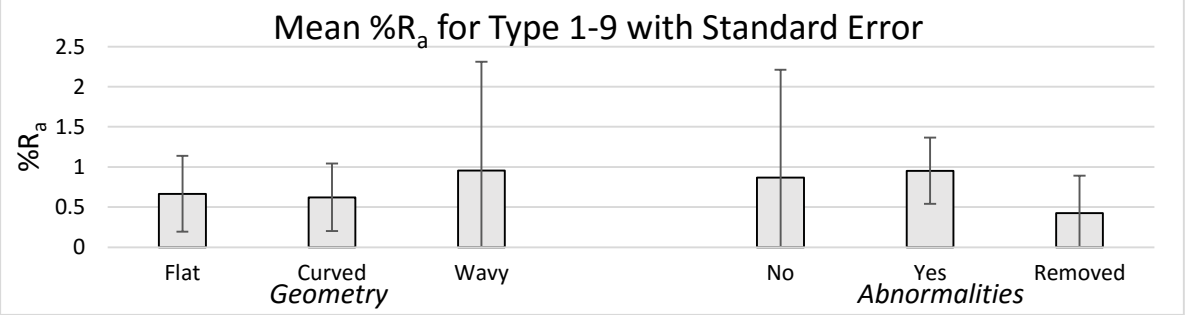
X2.4 The abnormality percentage is the surface area of all abnormalities falling within specification divided by the total surface area of the specified surface.

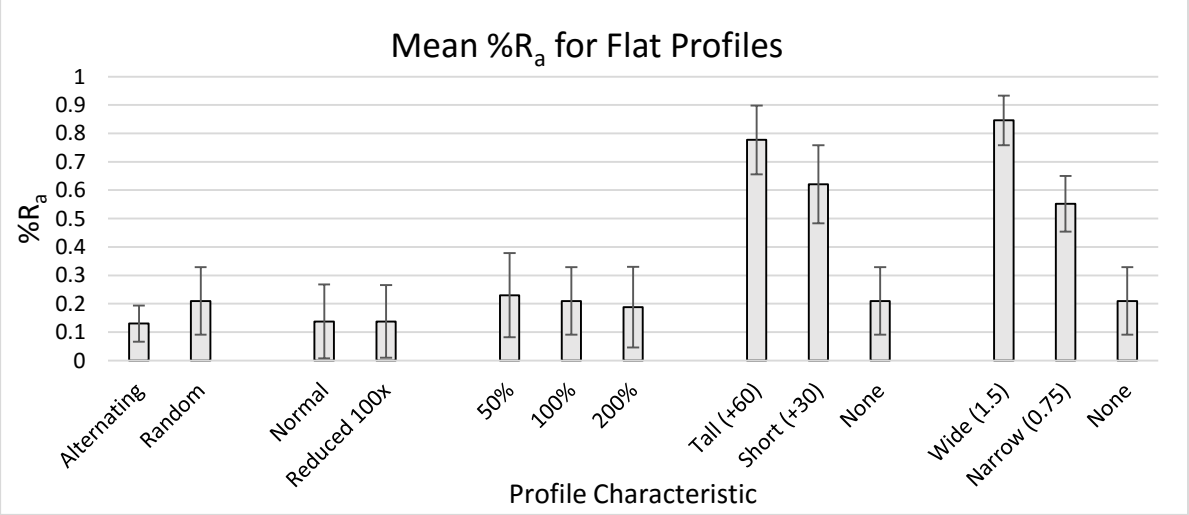
Surface Feature	SCRATA Comparator	Standard Equivalent
Surface Texture	A1	VSR 0.03 – 0.30 – 0
	A2	VSR 0.03 – 0.30 – 0
	A3	VSR 0.05 – 0.50 – 0
	A4	VSR 0.05 – 0.50 – 0
Non-metallic Inclusions	B1	VSR 0.04 – 0.40 – 0
	B2	VSR 0.03 – 0.11 – 1
	B4	VSR 0.02 – 0.12 –4
	B5	VSR 0.03 – 0.29 –6
Gas Porosity	C1	VSR 0.02 – 0.09 –0
	C2	VSR 0.03 – 0.27 –1
	C3	VSR 0.05 – 0.35 –1
	C4	VSR 0.04 – 0.36 –4
Fusion	D1	VSR 0.03 – 0.09 – 5
	D2	VSR 0.02 – 0.08 – 11
	D5	VSR 0.03 – 0.15 – 20

Surface Feature	SCRATA Comparator	Standard Equivalent
Expansion	E3	VSR 0.02 – 0.18 – 5
	E5	VSR 0.03 – 0.16 – 20
Inserts	F1	Not Applicable
	F3	Not Applicable
Metal Removal-Thermal	G1	VSR 0.01 – 0.09 – 2
	G2	VSR 0.02 – 0.10 – 23
	G3	VSR 0.02 – 0.10 – 25
	G5	VSR 0.03 – 0.58 – 52
Metal Removal-Mechanical	H1	VSR 0.02 – 0.03 – 0
	H3	VSR 0.02 – 0.05 – 1
	H4	VSR 0.04 – 0.12 – 6
	H5	VSR 0.04 – 0.31 – 12
Metal Removal-Welds	J1	VSR 0.01 – 0.06 – 0
	J2	VSR 0.01 – 0.05 – 3
	J3	VSR 0.01 – 0.10 – 22
	J5	VSR 0.01 – 0.41 – 20

APPENDIX B: GRAPHICAL REPRESENTATION OF SLICING METHODOLOGY RESULTS







APPENDIX C: PROPOSED SOFTWARE PROGRAM TO EVALUATE THE VOELKER SURFACE RATIO PARAMETERS

This appendix outlines the proposed software program to evaluate the Voelker Surface Ratio (VSR) parameters. The proposed program (tool) will use the standard process for calculating the underlying geometry once an ideal method is determined.

Purpose:
This tool will be developed for customers to see how their specification criteria can be interpreted and to train inspectors on the new standard.

Benefits:
Benefits of this tool include the following:

- Customers can use this tool to evaluate the criteria they want to specify.
- Manufacturers can use this tool to help train and calibrate inspectors.
- Surfaces can be checked against the VSR using go/no-go criteria.
- A visualization of areas falling outside the VSR criteria is provided to identify defects or potential rework areas.
- Companies can use scanning devices they already own for use with this program.

Overview:
This tool will take a point cloud, determine the underlying geometry, and compare the original point cloud to the underlying geometry in order to calculate the parameters of the VSR. This information will then be used with the new quantitative surface inspection in order to provide outputs of roughness and abnormality levels. The tool is interactive allowing users to visually see the differences between VSR specification criteria on a sample scanned surface.

Process:

1. Determine underlying geometry
2. Calculate the deviation of each data point from underlying geometry
3. Plot each point (original point cloud) and assign a scalar to the point to represent the deviation from underlying geometry
4. Input each of the VSR parameters
5. Color points according to their classification based on the VSR inputs
    a. Baseline roughness points: blue
    b. Abnormalities: yellow
    c. Defect (points falling outside of the specified criteria): red
6. Provide a Yes/No determination on each parameter so the user knows which parameter is failing
7. Provide an overall Yes/No so the user knows if the part is considered acceptable overall based on the specified VSR

A sample user interface is provided (Figures 1 and 2) on the following page as an example.

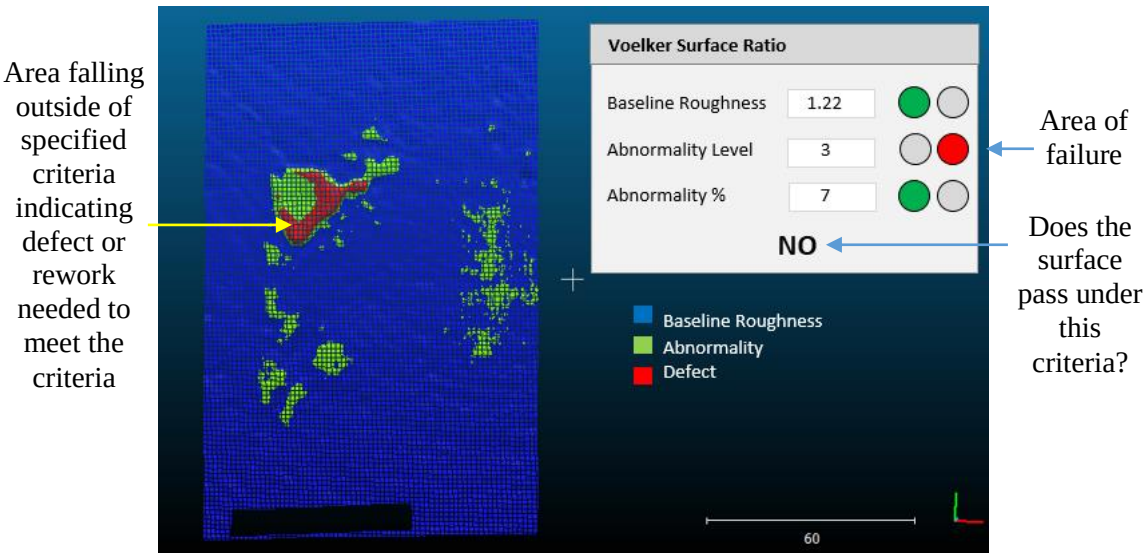


Figure 1—Initial VSR inputs resulting in an unacceptable part based on the acceptance criteria

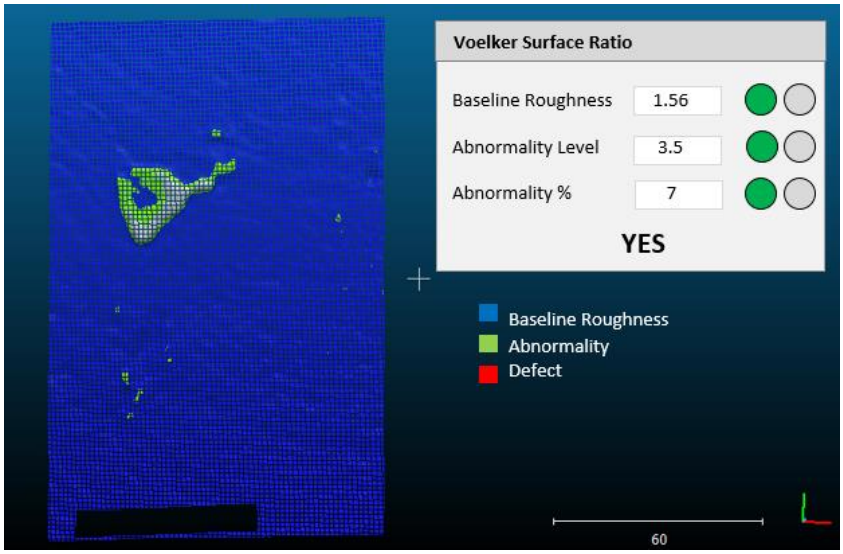


Figure 2—Modified VSR inputs (increased abnormality level) resulting in an acceptable part based on the acceptance criteria

Integration:

There are many potential options for integrating this program into already existing software. This includes creating a plugin for commercial software, such as Geomagic, or an open source software, such as CloudCompare. Alternatively, the program could be developed as a standalone interface with an opportunity for expansion with the development of the handheld scanning tool to verify these parameters.