LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-700977

Characterizing Complexity of Containerized Cargo X-ray Images

G. Wang, S. Glenn, H. Martz, N. Birrer, C. Divin

August 19, 2016

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Characterizing Complexity of Containerized Cargo X-ray Images

Guangxing Wang

Hosting Site: Lawrence Livermore National Laboratory

Mentor(s): Harry Martz Jr., Steven Glenn, Charles Divin

LLNL-TR-700977

Abstract. X-ray imaging can be used to inspect cargos imported into the United States. In order to better understand the performance of X-ray inspection systems, the X-ray characteristics (density, complexity) of cargo need to be quantified. In this project, an image complexity measure called integrated power spectral density (IPSD) was studied using both DNDO engineered cargos and stream-of-commerce (SOC) cargos. A joint distribution of cargo density and complexity was obtained. A support vector machine was used to classify the SOC cargos into four categories to estimate the relative fractions.

This research was performed under an appointment to the U.S. Department of Homeland Security (DHS) Domestic Nuclear Detection Office (DNDO) Summer Internship Program, administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and DHS at the Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. ORISE is managed by Oak Ridge Associated Universities (ORAU) under DOE contract number DE-AC05-06OR23100. This work has been supported by the US Department of Homeland Security, Domestic Nuclear Detection Office, under competitively awarded contract/IAA HSHQDC-12-X-00341. This support does not constitute an express or implied endorsement on the part of the Government. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of DHS, DOE, or ORAU/ORISE.

1. Introduction

Millions of cargo containers are imported into the United States every year. In order to ensure that no radiological and nuclear (Rad/Nuc) threats are hidden in these imports, X-ray technology may be used in the inspection process. It is important to assess the performance of existing and emerging X-ray technologies applied to cargos with various characteristics. Cargo density and complexity (inhomogeneity) have been shown, in previous DNDO studies, to be highly influential on the detection performance of X-ray systems. This project aimed to characterize the complexity of imported cargos. The complexity distribution can be combined with the density distribution to give an indication of the difficulty detecting Rad/Nuc threats in the cargo.

Section 2 explores various alternative complexity measures. In particular, integrated power spectral density (IPSD) is discussed in more detail since it provides a complexity measure that closely matches human-perceived image complexity. In section 3, the application of IPSD to stream-of-commerce (SOC) cargos to obtain the cargo complexity distribution is discussed.

2. Complexity Measures

Three different complexity measures were applied to X-ray images of 18 DNDO engineered cargos. The cargos were labelled by humans as high-density, high-complexity (HH), high-density, low-complexity (HL), low-density, high-complexity (LH), or low-density, low-complexity (LL). The goal was to find a measure that closely matched human-perceived complexity classifications.

In addition to IPSD, Law's texture measure [2] and edge density were considered. Law's texture measures give a nine-dimensional vector that contains combinations of measurements of levels, edges, spots, and ripples. Additional analysis would be needed to combine elements of this vector into a single complexity metric. The edge density was by Canny's edge-counting algorithm [1]. The process of obtaining this metric is similar to IPSD for each image. The only difference is that for IPSD, the power spectral density was used in the calculation; but for edge density, the pixel values of the edge image obtained by Canny's edge-counting algorithm was used.

Among the three measures considered, IPSD is the closest to human response. Figure 1 demonstrates that IPSD can distinguish between low- and high-complexity DNDO engineered cargos similar to human classification. The other two complexity measures did not match very well to human responses. Figure 2 shows the first dimension of the Law's texture measure vs. mean equivalent steel

thickness for the DNDO cargos. It clearly misclassified water as high-complexity cargo. Moreover, it also had difficulty separating between LH and HL cargos near complexity 0.01. The method of counting edges did not show satisfactory result either, as can be seen in Figure 3, where it failed to distinguish between LL and LH cargos. It also misclassifies water as high-complexity.

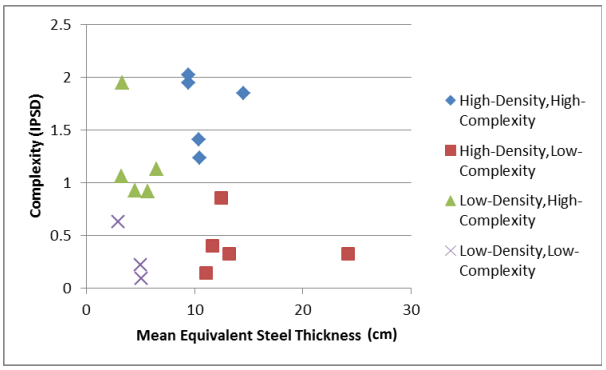

Figure 1. IPSD vs. Mean equivalent thickness.

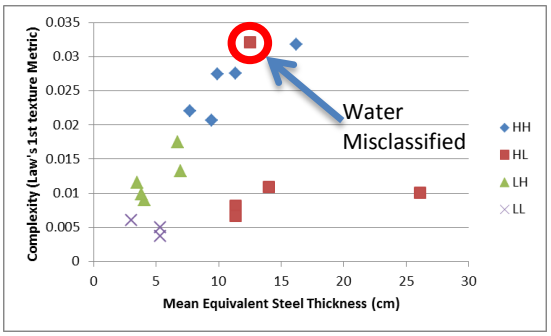

Figure 2. Law's 1st texture metric vs. Mean equivalent thickness.

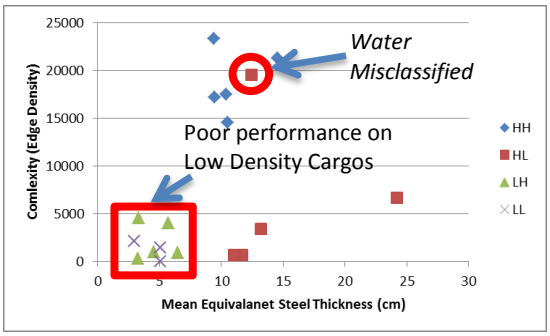

Figure 3. Edge density vs. Mean equivalent thickness.

In order to minimize the effects of different responses that X-ray scanners may have on the performance of the complexity measures, all of the analysis was done on equivalent steel thickness images converted from intensity images. Figure 4 shows two examples of DNDO engineered cargos with different human perceived complexities. Plastic cargo is labelled as low-complexity due to its homogeneity, whereas brake cargo is labelled as high-complexity due to its heterogeneity.

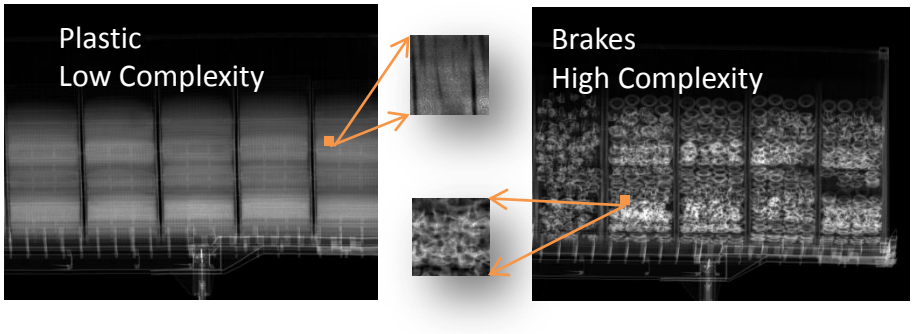

Figure 4. Examples of DNDO engineered cargo equivalent steel thickness image.

The IPSD complexity measure is constructed by sliding a 128x128 window across the image. In other words, a sub-region of 128x128 is formed at each pixel location. A Hamming window is applied to each sub-region to eliminate effects caused by discontinuities at the edges. The next step is to use Fast Fourier Transform to transform each sub-region to the spectral domain and obtain a spectral density image as shown in Figure 5. As shown in Figure 6, the DNDO cargos have different spatial frequency distributions for certain frequency range. Furthermore, high-complexity cargos tend to have more high-frequency content than low-complexity cargos in the frequency range between 0.2 cm^{-1} and 0.55 cm^{-1}. Extremely high frequencies (>0.8 cm^{-1}) are often the result of very small structures or noise in an image, which do not contribute to human perceived image complexity because they are difficult to be observed by human eye. On the other hand, extremely low frequencies are from large structures in an image, which are easy to recognise by humans, thus they should not be considered into calculating complexity. In this particular case with X-ray cargo images, large structures can be formed by the way cargos are packaged or positioned in a container; thus are not the properties of the cargo itself. As a result, the extremely low and high frequencies should be excluded when calculating the complexity measures. A grid search was performed to find lower and upper limits that best separate the high and low complexity cargos. The optimal lower (0.2 cm^{-1}) and upper (0.55 cm^{-1}) frequency limits are the left and right boundaries of the shaded region in Figure 6. The two-dimensional version can be seen as the two red circular rings in Figure 5. The horizontal axis in Figure 6 corresponds to each ring in Figure 5, and the vertical axis shows the coefficients at each frequency which are shown in Figure 5 as different darkness. The IPSD was computed as:

$$IPSD = \int_{f_x, f_y} PSD(f_x, f_y) df_x df_y$$

where

$$PSD(f_x, f_y) = \left|H(f_x, f_y)\right|^2$$

and

$$H(f_x, f_y) = \int_{x,y} e^{2\pi i x f_x} e^{2\pi i y f_y} g(x, y) d_x d_y$$

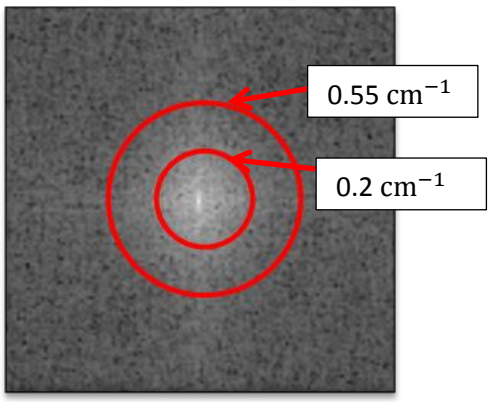

Figure 5. Spectral density image of a sub-region from brakes cargo image in Figure 7 showing annular region used for IPSD calculation. The red rings correspond to the limits shown in Figure 6.

The fourth step is to sum all the values in an annulus within the spectral density to obtain the IPSD. As the sub-region slides across the image, an IPSD value is obtained at each pixel. Figure 7 shows the comparison between the original image and its corresponding IPSD image. It can be seen that the IPSD image for brakes cargo has higher complexity than the plastic cargo.

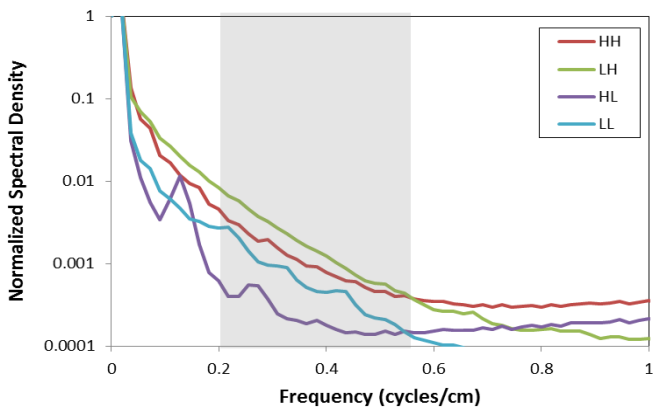

Figure 6. Azimuthally-averaged power spectral densities for DNDO cargos. The shaded region shows the optimal frequency bound for discriminating high- and low- complexity cargos.

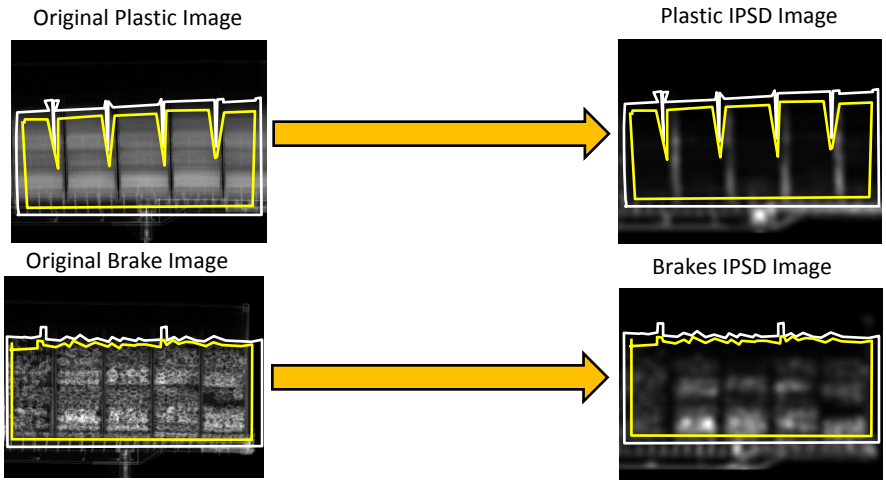

Figure 7. Examples of steel thickness (left) and IPSD (right) images. White outline is the boundary of cargos in the container, yellow outline is eroded from white outline by 64 pixels.

The next step is to apply the cargo mask to the image to isolate the cargo area, and apply binary erosion to shrink the mask so that the cargo edges and gaps between cargos do not contribute to the final complexity measure. This procedure is shown in Figure 7, where the yellow boxes indicate the eroded cargo mask boundaries. The final complexity measure is defined as

$$\boldsymbol{Complexity} \equiv \frac{\mathbf{100 \times Average\ within\ yellow\ boarder}\ \boldsymbol{in\ IPSD\ image}}{\mathbf{Average\ within\ yellow\ boarder\ inequivalent\ steel\ thickness\ image}} \tag{2.1}$$

where the denominator is used as a normalizer.

3. Application of IPSD on SOC cargos

IPSD was computed for 483 SOC cargo X-ray images from the NRIP Freeport data set to study the joint complexity/density distribution. Figure 8 shows a scatter plot of the cargos, where red dots are 40-foot containers and blue dots are 20-foot containers. Figure 9 shows the contour plot of the joint density estimated by kernel density estimation by a Gaussian kernel. It can be seen that few SOC cargos are high-density, high-complexity, and many are high-density, low-complexity. Figure 9 also shows that the DNDO engineered cargos span the distribution of the SOC cargos, which means that they are good representatives of SOC cargo.

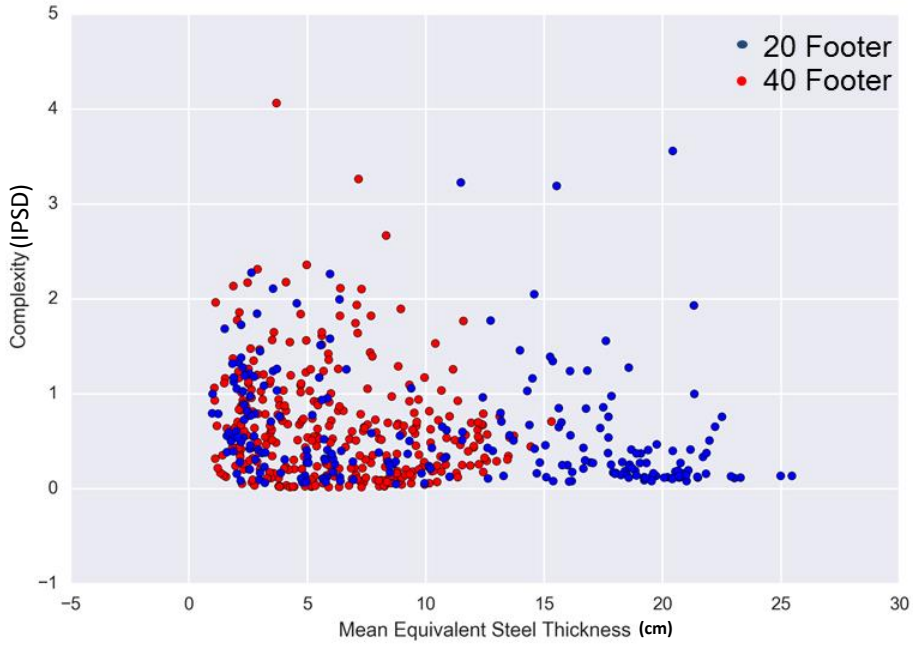

Figure 8. Scatter plot of IPSD vs. Mean equivalent steel thickness for SOC cargos.

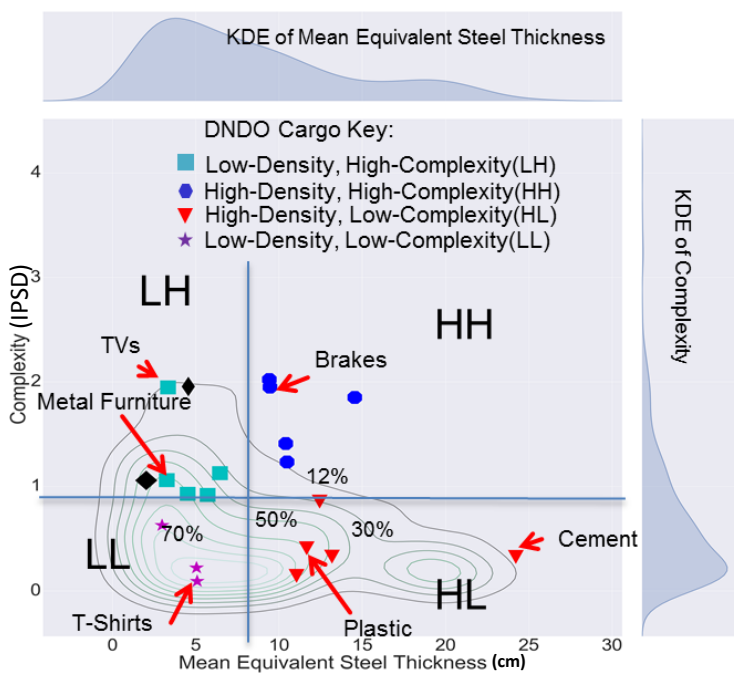

Figure 9. Kernel Density Estimated Contours for SOC cargos. The data points are for DNDO cargos. The blue lines serve as a guide that shows the DNDO cargos can be separated well by IPSD and Mean equivalent steel thickness.

In order to estimate the fractions of SOC cargos that are HH, HL, LH, and LL, support vector machine (SVM) was used for classification. The training data set is the DNDO engineered cargos. The SVM SOC cargo class are shown in Figure 10. Among the 483 Freeport SOC cargos, SVM classified 5.3% as HH, 44% as HL, 27.1% as LH, and 23.6% as LL. Based on the classification by SVM, of the high density-high complexity SOC cargos 36.7% are 20-footers, of the high density-low complexity cargos 50.4% are 20-footers, of the low density-high complexity cargos 33.3% are 20-footers, and of the low density-low complexity cargos 33.1% are 20-footers.

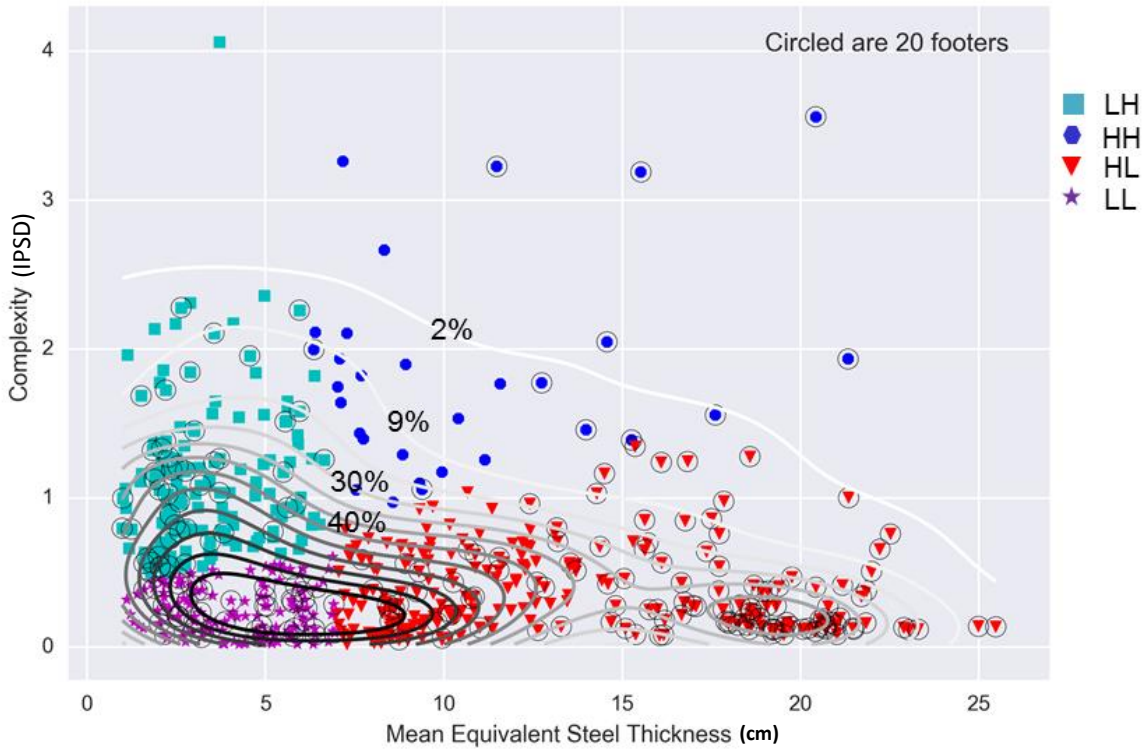

Figure 10. Classification of SOC cargos by Support Vector Machine.

IPSD complexity of SOC cargos was also studied on the basis of tariff codes (HS2). Table 1 shows the fraction of cargos that are HH, HL, LH, and LL for each HS2. As expected, plastics, wood and paper products were mostly low complexity. Iron articles span various complexities and densities. Vehicles, electrical equipment, and mechanical equipment are mostly high complexity. Instruments are mostly low density, but vary in complexity. The majority of base metal articles and furnishings are also low complexity.

Table 1 Density/Complexity breakdown for selected cargos based on HS2 codes

	Plastics (39)	Wood (44)	Paper (48)	Iron (73)	Vehicles (87)	Base Metals (83)	Electrical Equipment (85)	Mechanical Equipment (84)	Instruments (90)	Furnishings (94)
HH	0	0	0	16.7%	7.3%	12.5%	9.3%	3.8%	0	0
LH	6.8%	0	0	33.3%	82.9%	12.5%	67.4%	67.3%	66.7%	27.3%
HL	61.4%	100%	72.7%	37.5%	2.4%	50%	16.3%	7.7%	0	18.2%
LL	31.8%	0	27.3%	12.5%	7.3%	25%	7%	21.2%	33.3%	54.5%
ALL	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

4. Discussion

Three image complexity metrics were applied to DNDO engineered cargos. Integrated power spectral density (IPSD) performed the best in terms of matching human-perceived complexity. The joint distribution between IPSD and mean equivalent steel thickness for the SOC cargos was estimated by a Gaussian kernel density estimation. The SOC cargos were also classified using a support vector machine trained on the DNDO cargos. The results show that 5.3% SOC cargos are high-density, high-complexity, and 44% are high-density, low-complexity. Of the high-density, low-complexity cargos more than half are held in 20-foot containers. The complexity distributions were also analysed for each major tariff codes, where the fractions of HH, HL, LH and LL of different kinds of imports were estimated. The complexity distributions among different kinds of products were as expected. In this project, only the support vector machine was used in the classification because of time constraints. In the future, more sophisticated machine learning techniques may be applied, for example, ensemble methods, to achieve more accurate classification.

5. Impact of internship on my career

This internship has been not only a valuable opportunity to learn a new set of skills, but also a great experience to collaborate with people with different backgrounds and expertise. The project that I have worked on introduced me to the topic of image processing, and it is fascinating to see the importance and broad application of this topic in real world situations. As a result, I plan to expand my research, when I go back to school, to include topics on image analysis.

The DHS lecture series provided interesting talks, which increased my awareness of cutting-edge research that is conducted in the national labs. The tours of state-of-the art facilities and many social events organized by the lab helped me to see what it was like to work in the lab. The overall positive experience of my internship made working at the lab one of the possibilities in my list after graduation.

6. Acknowledgments

I would like to thank everyone who made this internship possible, and special thanks to my project mentors who provided great support and valuable advice on this research project. I would also like to thank my office mate, Nat Birrer, who helped me on programing and physics questions.

7. References

[1] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, **(6), 679-698.**

[2] Laws, K. I. (1980). *Textured image segmentation* (No. USCIPI-940). University of Southern California Los Angeles Image Processing INST.

[3] Stoica, P., & Moses, R. L. (1997). *Introduction to spectral analysis* (Vol. 1, pp. 3-4). Upper Saddle River: Prentice hall.