

2006-33: PROBABILITY AND IMAGE ENHANCEMENT

Maurice Aburdene, Bucknell University

MAURICE F. ABURDENE is the T. Jefferson Miers Professor of Electrical Engineering and Professor of Computer Science at Bucknell University. He has taught at Swarthmore College, the State University of New York at Oswego, and the University of Connecticut. His research areas include, parallel algorithms, simulation of dynamic systems, distributed algorithms, computer communication networks, control systems, computer-assisted laboratories, and signal processing.

Thomas Goodman, Bucknell University

THOMAS J. GOODMAN earned his B.S. degree in electrical engineering from Bucknell University and is currently pursuing a Master's degree at Bucknell, also in electrical engineering. His research interests include discrete transforms and efficient hardware implementation of transform algorithms and other operations used in digital signal processing. He will be graduating from Bucknell in May 2006 and plans to begin work as a hardware design engineer shortly thereafter. He grew up in Rochester, NY.

Probability and Image Enhancement

Abstract

We present one of five projects used in our course, **Probability with Applications in Electrical Engineering**. The course is required for all electrical engineering students and is open to third and fourth year students. The project focuses on the applications of probability to image enhancement using histogram equalization and histogram specification methods. These techniques demonstrate applications of functions of random variables, transformations of random variables, and the generation of random variables from specified distributions. We begin by introducing the continuous random variable transformation and demonstrating the process of transforming any random variable distribution to a uniform distribution through the use of the cumulative density function. We then explore the concept of histogram equalization: how it works, its effects on image contrast, and its applications in image processing and image enhancement. Finally, we generalize the histogram equalization problem by showing how the cumulative density function can be used to specify an arbitrary probability distribution and to transform the image accordingly.

Introduction

ABET evaluation criteria for electrical engineering programs state “The program must demonstrate that graduates have: knowledge of probability and statistics, including applications appropriate to the program name and objectives; and knowledge of mathematics through differential and integral calculus, basic sciences, computer science, and engineering sciences necessary to analyze and design complex electrical and electronic devices, software, and systems containing hardware and software components, as appropriate to program objectives.”(See <http://www.abet.org/criteria.html>).

We present one of five projects used in our course, **Probability with Applications in Electrical Engineering**. The course is required for all electrical engineering students and is open to third and fourth year students. We introduce a way to make this topic more appealing to students. In the latest offering, the four other projects included linear averaging,¹ computer networks and simulation,² frequency response and least-squares estimation,¹ and conditional probability and receivers in communication systems¹.

The project focuses on the applications of probability to image enhancement using both histogram equalization and histogram specification methods. The histogram equalization technique directly uses the original image pixel values to compute the enhanced image’s pixel values. Histogram equalization is widely used in medical image processing, facial recognition, radar, and photo processing software. Image enhancement techniques demonstrate applications of functions of random variables (transformations of random variables, derived random variables) and the generation of random variables from specified distributions. Image processing examples are good in the sense of yielding immediate visual feedback; in addition, students may already have had experience with image editing software.

The project requires students to apply their knowledge of probability concepts, including probability density functions (pdf), probability mass functions (pmf), and transformations of random variables to image processing.^{3,4,5,6,7} In addition to a deeper understanding of probability, students discover two important points: first, that there is no general theory of image enhancement, and second, the quality and method of the image enhancement depends on both the original image and the viewer⁸.

In class and homework, we begin by introducing continuous random variable transformation and demonstrating the process of transforming any random variable distribution to a uniform distribution through the use of the cumulative density function (cdf).^{5,6,7,8,9} We then explore the concept of image histogram equalization: how it works, its effects on image contrast, and its applications in image processing and image enhancement. Finally, we generalize the histogram equalization problem by showing how the cumulative density function can be used to specify an arbitrary probability distribution and to transform the image accordingly.

Functions of Continuous Random Variables

Given a random variable X , its pdf, $f_X(x)$, and a derived random variable $Y = g(X)$, students are asked to find the pdf and cdf of Y and compare them with the pdf and cdf of X . In addition, we look at the mean values and standard deviations of both X and Y . Students are asked to explore both analytically and using MATLAB, the following:

1. Let $Y = X + 3$.
 - a. Determine the pdf of Y if X is a uniformly distributed random variable with

$$f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Sketch $f_X(x)$ and $f_Y(y)$.

- b. What is $E(X)$, the expected value of X ? What is the variance of X ?
 - c. What is $E(Y)$? What is the variance of Y ?
2. Let $Y = 5X + 1$.

- a. Determine the pdf of Y if X is a uniformly distributed random variable with

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Sketch $f_X(x)$ and $f_Y(y)$.

- b. What is $E(X)$? What is the variance of X ?
 - c. What is $E(Y)$? What is the variance of Y ?

3. Let $Y = 3X + 5$.

- Determine the pdf of Y if X is a normally (Gaussian) distributed random variable with $E(X) = 0$ and $\text{Var}(X) = 2$.
- Sketch $f_X(x)$ and $f_Y(y)$.
- What is $E(Y)$?
- What is the probability that $Y \leq 5$?

4. Let $Y = X/10$.

- Determine the pdf of Y if X is a uniformly distributed random variable with

$$f_X(x) = \begin{cases} 1/10 & 0 \leq x \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

- Sketch $f_X(x)$ and $f_Y(y)$.
- What is $E(X)$? What is the variance of X ?
- What is $E(Y)$? What is the variance of Y ?

5. Let $Y = 1 - e^{-X}$.

- Determine the pdf of Y if X is an exponentially distributed random variable with

$$f_X(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Sketch $f_X(x)$ and $f_Y(y)$.

Examples 4 and 5 show that if $g(X)$ is the cdf of X , then Y is a uniformly distributed random variable in the range $[0,1]$.

Our students use Matlab in an earlier course, Linear Systems. Students are asked to check their analytical results by using a package such as MATLAB¹¹ or Mathematica.¹² We then ask the students to use the transformation $y = F_X(x)$ and observe the distributions of Y and as mentioned earlier they note that Y has a uniform distribution over the range $0 \leq y \leq 1$. We call this process histogram equalization. Here we ask students to critically look at the characteristics of $F_X(x)$. We then ask what is the range of values for X and Y ?

Random number generation

Often we are asked to generate random numbers with a specific distribution. Suppose we would like to generate the exponential probability density function exponential

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

as shown in Figure 1 with $\lambda = 1$.

The cumulative density function for an exponential distribution is

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Since $0 \leq F_X(x) \leq 1$, if we let Z be a uniformly distributed random variable with $0 \leq z \leq 1$, then $P(Z \leq z) = P(X \leq x)$ and

$$x = F^{-1}(z) = -(1/\lambda) \log(1 - z).$$

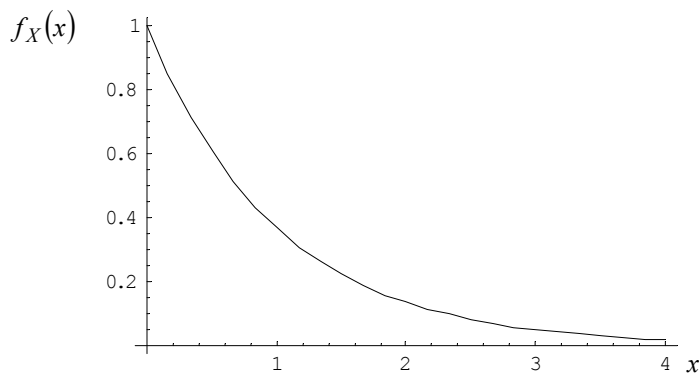


Figure 1. Desired probability density function

Application to images

Given a two-dimensional image $\{x\}_{jk}$ shown in Figure 2a, we define the random variable X as the pixel values of the image with the sample space of $X = \{x_0 = 0, x_1 = 1, x_2 = 2, \dots, x_{255} = 255\}$.

The image has a probability density function $f_X(x)$, and a histogram of the pixel values is shown in Figure 2c. We would like to transform the image to have the probability density function $f_Y(y) = (1/64)e^{-y/64}$ as shown in Figure 3.

As we have seen, the transform is performed by replacing each $\{x\}_{jk}$ with a new value $\{y\}_{jk}$, which will depend both on $f_X(x)$ and $f_Y(y)$. The sample space of $Y = \{y_0 = 0, y_1 = 1, y_2 = 2, \dots, y_{255} = 255\}$.

Assuming a continuous random variable Y , the desired cumulative density function is given by

$$F_Y(y) = \int_0^y f_Y(\omega) d\omega = 1 - e^{-y/64}.$$

It follows that

$$F_Y^{-1}(z) = -64 \log(1 - 0.989z),$$

where z is treated as a uniformly distributed random number with $0 \leq z \leq 1$. Thus the image transformation we require is given by

$$\{y\}_{jk} = -64 \log(1 - 0.989 F_X(\{x\}_{jk})).$$

The image that results from applying this operation to the sample image in Figure 2a is shown in Figure 2b, along with the histogram of pixel values in Figure 2d.



(a)



(b)

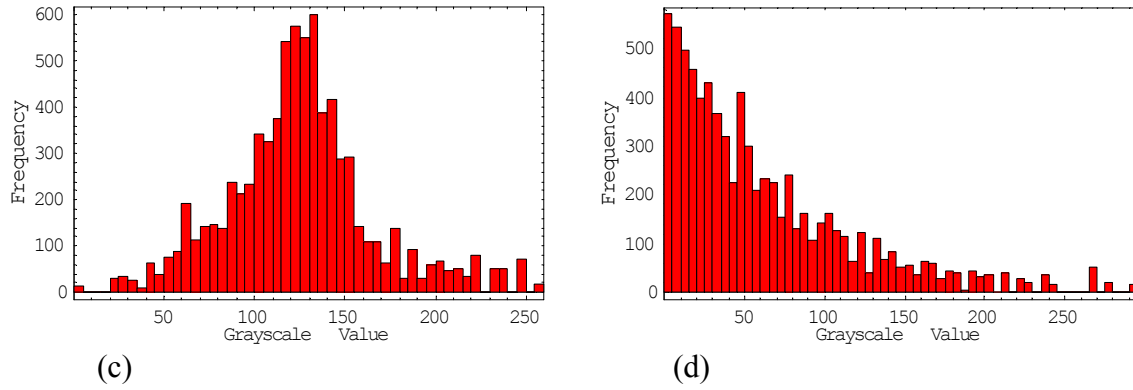


Figure 2. (a) Sample image; (b) transformed image; (c) histogram of pixel values from sample image; (d) histogram of pixel values from transformed image.

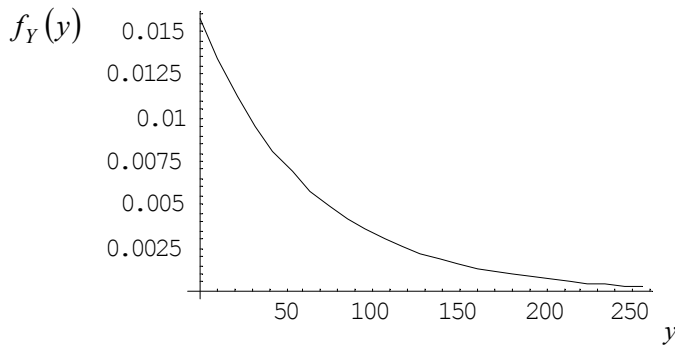


Figure 3. Desired pmf of the transformed image

Histogram Equalization

As mentioned earlier, histogram equalization is a transformation g of a random variable X such that $Y = g(X)$ has a uniform distribution, $f_Y(y)$ [6]. Given a random variable X representing the pixel values of a grayscale image $\{x\}_{jk}$ for $1 \leq j \leq N$ and $1 \leq k \leq M$, we can examine the distribution of X and perform a transformation on the pixel values to produce an enhanced image $Y = g(X) = \{y\}_{jk}$. If we let $f_X(x)$ and $f_Y(y)$ be the pmfs of X and Y respectively, then

$$\sum_{i=0}^{255} f_X(x_i) = \sum_{i=0}^{255} f_Y(y_i) = 1.$$

If we assume that Y is uniformly distributed, then its cdf is

$$F_Y(y) = \sum_{i=0}^y f_Y(y_i) = \sum_{i=0}^y \frac{1}{256} = \frac{y+1}{256}$$

and

$$y_k = 255 \sum_{i=0}^k f_X(x_i) = 255 F_X(x_k).$$

In the context of image enhancement, histogram equalization is a process by which each grayscale pixel value x_i is replaced by $F_X(x_i)$, where $F_X(x)$ is the cdf function of the original image pixel values x . An algorithm for histogram equalization is:

1. Find the pmf of the image by counting the number of times each grayscale value occurs in the image and dividing by the number of pixels.
2. Form the cdf of the image. If the pixel values of the original image range from 0-255 in one-unit increments, the cdf is given explicitly by

$$F_X(x) = \sum_{k=0}^x f_X(k).$$

3. Replace all pixels $\{x\}_{jk}$ in the image by $255 F_X(\{x\}_{jk})$.

Usually, histogram equalization transforms the distribution of the pixels in an image to increase overall contrast. In this project, we ask students to look at the pmf of both X and Y and note that the values that the two functions take on are identical, but are distributed differently. Note that the $\{y\}_{jk}$ range is spread to enhance image contrast. However, if we have a large image and we look at the histogram of the $\{y\}_{jk}$ values, we note that it looks like a “uniform distribution,” using a small number of bins in the histogram.

When using this project, the teacher should be careful in extending the transformation of a random variable from the continuous case to the discrete case. We recognize that applying continuous histogram equalization methods to image processing might be challenging as an application of probability density transformation at the undergraduate level. In fact, students discover that image histogram equalization does not lead to uniform distributions due to the discrete nature of the problem.

Image Enhancement

Here we note that the uniformly distributed pixel values greatly enhance image quality. Let us demonstrate by performing histogram equalization on the 323x250 image shown in Figure 4a. The image has low contrast and appears “washed out,” and a histogram of its grayscale pixel values, shown in Figure 4c, confirms this observation. The distribution has a mean of 109 and a variance of 622; clearly it is not a uniform distribution.

We can perform an accumulative sum of the pmf of X to obtain the cdf of the image. The accumulation can be done recursively, i.e.

$$\begin{aligned} F_X(0) &= 0 \\ F_X(n) &= F_X(n-1) + f_X(n) \quad 1 \leq n \leq 255 \end{aligned}$$

where $f_X(n)$ is the pmf evaluated at n . Figure 5 shows the cdf for this image.

Next we will perform histogram equalization. We will replace each pixel $\{x\}_{jk}$ in the image by $255 \cdot F_X(\{x\}_{jk})$. The resulting image is depicted in Figure 4b, and a histogram of its pixel values is shown in Figure 4d. This distribution more closely resembles a uniform one (and based on the theoretical framework of the technique, the distribution is *as close as one can get* to a uniform distribution).

The mean of the new distribution is 129 and the variance is 5340. The mean has moved towards the theoretical mean of a uniform distribution of values between 0 and 255. The variance has also increased, which we expect due to the spread of pixel values.

How do the original and equalized images compare? Figure 4 shows a side-by-side comparison. The equalized image, shown in 4b, appears brighter and has better contrast. From a visual standpoint, image quality has improved significantly.

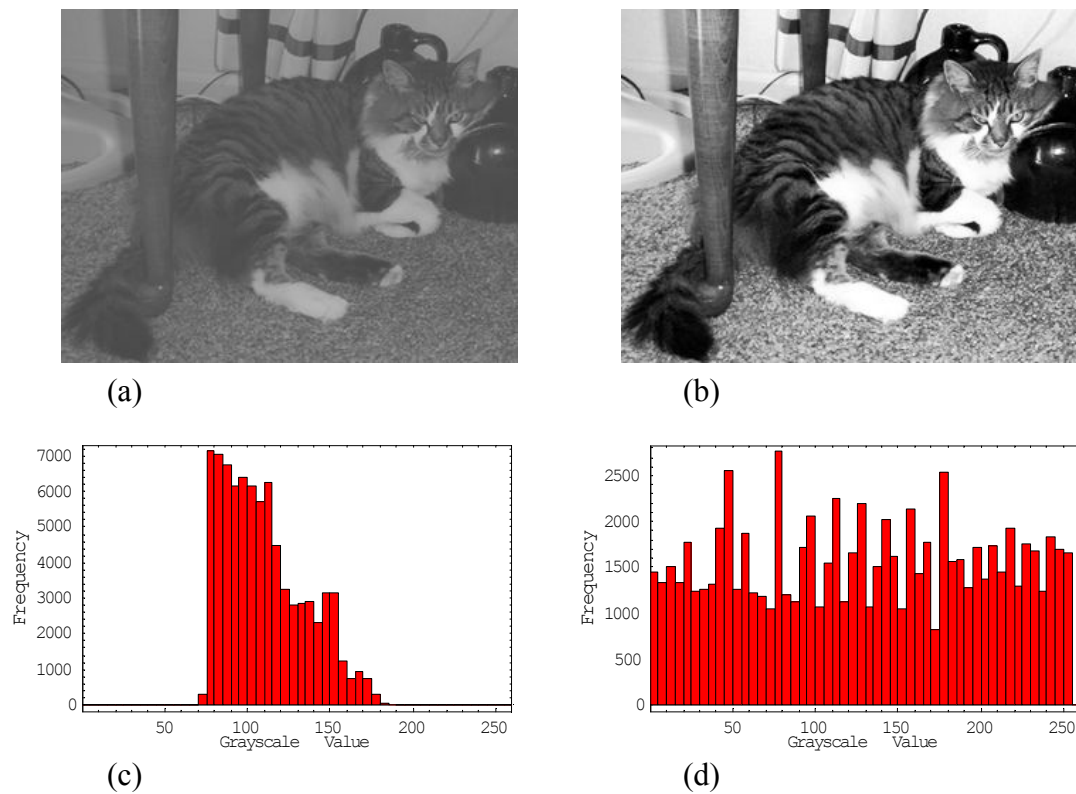


Figure 4. (a) Original image; (b) equalized image; (c) histogram of pixel values from original image; (d) histogram of pixel values from equalized image.

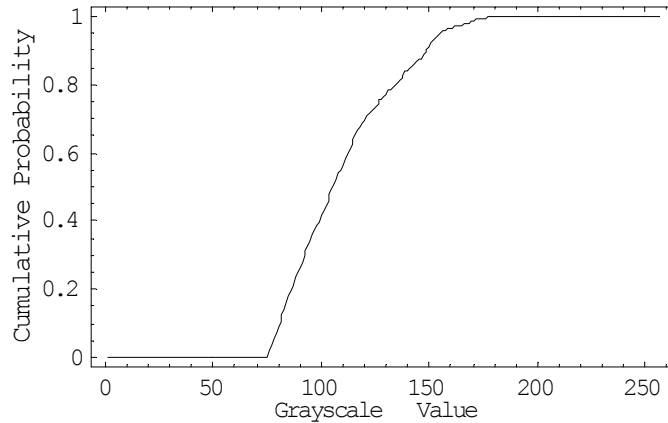


Figure 5. Cumulative distribution function (cdf) of the image in Fig. 4a

Once we notice that applying histogram equalization to this image has greatly improved overall quality, we may be tempted to apply it a second time. However, image quality will not be improved again because the distribution is already uniform. In other words, the cdf of the equalized image should be the function $F_Y(y) = (y+1)/256$. We can verify this by computing the cdf of the equalized image, shown in Figure 6. We can see that the cdf closely approximates the function $F_Y(y) = (y+1)/256$, as predicted. As a result, we predict that performing histogram equalization on this new image will have no effect, but we will perform the experiment to confirm. Figure 7 shows the resulting histogram, which is identical to the one we obtained for the first equalized image, Y .

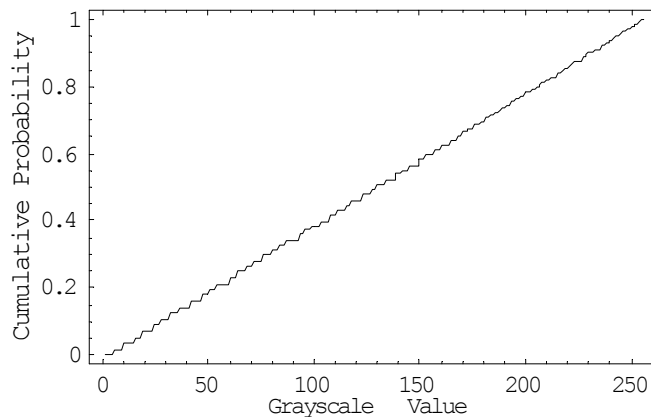


Figure 6. Cumulative probability mass function of the equalized image.

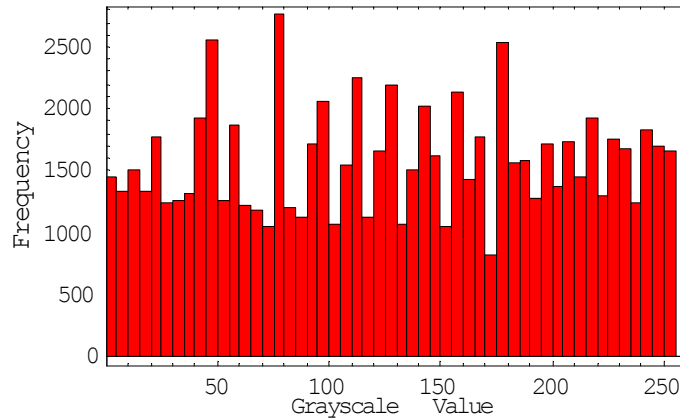


Figure 7. Histogram of the image Z, the result of applying histogram equalization to X twice.

Project Assessment

This project has been used for three years. Table I shows the project reports grades for the three years. Student interest is demonstrated by the range of grades.

Table I: Project report grades

Year	Average	Standard Deviation	High	Low
2003	8.8	0.8	10	7
2004	8.86	0.81	10	7
2005	9.92	0.28	10	9

After the projects reports are submitted and returned, the students are examined on the fundamental concepts as well as the ability to perform the histogram equalization on a small size image. Table II shows the project exam grades.

Table II: Project exam grades

Year	Possible Points	Average	Standard Deviation	High	Low
2003	40	27.6	7.06	40	14
2004	50	39.5	9.22	50	13
2005	50	43.17	9.17	50	20

The course evaluation form has the question, **what aspects of the course did you like most?** Some student comments that relate to this project include:

- Projects and HW were from real life examples.
- Relating it to electrical concepts.
- The project format and the sometimes unguided approach to problem solving.
- Image enhancement.

- Projects and final project.
- The projects were good, varied and interesting with applicable (to EE) topics.
- I liked the projects because they were very diverse and interesting.
- The material is linked to real life application. Learning about image equalization, network, queuing was interesting.
- I liked the Matlab projects. They really helped gain an understanding of the material.
- I think how the course related everyday technology to the probability we learned about.
- I liked the image processing project.
- It is related to real world problems. Matlab made it more hands on and easy to understand.
- The projects because they gave me a workable physical example of problems that I could understand and be interested in.
- I enjoyed the material on image processing and decisions rules.
- I liked the projects, especially the term project because we can pick a topic that is interesting to us.
- Simulating everything in Matlab. Doing hands on work in class.
- The projects helped me apply the concepts.
- Application of probability to real life.
- Projects were generally interesting to do.
- Applying Matlab really helped to enforce principles taught in class.

Summary

We presented one project that we use to demonstrate the applications of probability in electrical engineering. The application of both histogram specification and histogram equalization to images appeal to undergraduates. Students realize that histogram equalization is a powerful tool for enhancing image quality.

Acknowledgments

The authors would like to acknowledge the help of Professor Richard Kozick.

References

[1] Maurice F. Aburdene and Richard J. Kozick, "A project-oriented course in probability and statistics for undergraduate electrical engineering students", *Proceedings of Frontiers in Education Conference*, Vol. 2, 1997, pp. 598-603.

[2] Maurice F. Aburdene and Thomas J. Goodman, "Probability, Computer Networks, and Simulation", with **Thomas J. Goodman (undergraduate student)**, *Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition*, Portland, Oregon, June 12-15, 2005, Session 1432.

- [3] Roy D. Yates and David J. Goodman *Probability and Stochastic Processes: A Friendly Introduction for Electrical & Computer Engineers*, John Wiley Publishers, 2005. Chapter 3, section 3.7.
- [4] Rodger E. Ziemer, *Elements of Engineering Probability and Statistics*, Prentice-Hall, 1997. Chapter 3, section 3-4.
- [5] Roy D. Yates and David J. Goodman *Probability and Stochastic Processes: A Friendly Introduction for Electrical & Computer Engineers*, John Wiley Publishers, 2005. Chapter 3, section 3.7.
- [6] Rodger E. Ziemer, *Elements of Engineering Probability and Statistics*, Prentice-Hall, 1997. Chapter 3, section 3-4.
- [7] T.T. Soong, *Fundamentals of Probability and Statistics For Engineers*, John Wiley & Sons, 2004. Chapter 5, pp. 119-134.
- [8] Charles W. Therrien and Murali Tummala, *Probability for Electrical and Computer Engineers*, CRC Press, 2004.
- [9] A. H. Haddad, *Probabilistic Systems and Random Signals*, Pearson Education, Inc., 2006.
- [10] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison- Wesley, 1992.
- [11] Duane C. Hanselman, Bruce L. Littlefield, *Mastering MATLAB 7*, Prentice Hall, 2004.
- [12] Stephen Wolfram, *The Mathematica Book*, Fifth Edition, Wolfram Media, 2003.