Numerical Analysis and Design Strategy for Field Emission Devices

by

Yao-Joe Yang

B.S., Mechanical Engineering, National Taiwan University (1990)
M.S., Mechanical Engineering, University of California, Los Angeles (1995)
M.S., Electrical Engineering, Massachusetts Institute of Technology (1997)

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

August 1999
[September 1999]

Signature of Author ______________________________
Department of Electrical Engineering and Computer Science
August 31, 1999

Certified by ______________________________
Stephen D. Senturia
Barton L. Weller Professor of Electrical Engineering
Thesis Co-Supervisor

Certified by ______________________________
Akintunde I. Akinwande
Thesis Co-Supervisor

Accepted by ______________________________
Arthur C. Smith
Chair, Department Committee on Graduate Students

Numerical Analysis and Design Strategy for Field-Emission Devices

by

Yao-Joe Yang

Submitted to the Department of Electrical Engineering and Computer Science
on August 31, 1999 in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

ABSTRACT

An efficient *three-dimensional* (3-D) *computer-aided-design* (CAD) tool for modeling and simulating *field-emission devices/displays* (FEDs) is presented. The performance optimization of FEDs using this CAD tool and the *design of experiment* (DOE) method is demonstrated. Experimental results of emission current and spot size are in good agreement with the simulated results. The regression models created by the DOE method give excellent guidelines for both the design and operation of the FEDs.

This CAD tool uses an accelerated *boundary-element-method* (BEM) electrostatics solver and an adaptive explicit integration method. The typical CPU time for a complete electrostatics and trajectory simulation is less than 1 hour running on a SUN Ultra 30 model 295 workstation, versus a few hours/days for 3-D *finite-element-methods* (FEM) simulations. The simulated results of a single-gate *field emitter array* (FEA) are in excellent agreement with experimental results. Simulation examples of proximity focused FEDs and *integrated-focus-electrode* (IFE) focused FEDs are also presented.

A design strategy using DOE techniques is proposed. Experiments were conducted by using the CAD tool. The central composite design, which is capable of screening significant parameters and building regression models, is employed to explore device performance. The regression models created by the central composite design give excellent descriptions of the relationships between the device response and their design parameters. Design examples of proximity focused FEDs and IFE focused FEDs are presented, and their optimal design and operating condition are explored. The anode-gate separation is the most significant parameter for the proximity focused FEDs, while the ratio of the IFE voltage to the gate voltage is the most significant parameter for the IFE focused FEDs.

Thesis Supervisors:
Stephen D. Senturia
Barton L. Weller Professor of Electrical Engineering

Akintunde I. Akinwande
Associate Professor of Electrical Engineering

ACKNOWLEDGMENTS

Although this work bears only my name, it wouldn't have been accomplished without the help and support of many people. I would like to take this opportunity to acknowledge some of the specific contributions.

First and foremost, I would like to express my greatest gratitude to my research co-advisors, Professors Stephen Senturia and Tayo Akinwande for their support and guidance. Steve's feedback and suggestions were always timely and precise and kept me on the right track of research. Steve's acute insights into physical systems always inspire me to analyze problems in broad and profound ways. His sharp ability to recognize the hidden significance of my technical contributions and then nurture and encourage them to completion is highly appreciated. Also, Steve's enormous patience and encouragement always keep me working with him with great alacrity.

Tayo inspired the work in the thesis, and has been an extremely effective thesis advisor. His extensive industrial experience and physical insight have provided me with a learning experience, both technical and (more importantly) otherwise, which I consider more important and pertinent to life than what I have learned in the classroom. Tayo's patience and kindness was greatly appreciated. Especially, I would like to show my deepest appreciation to both of my co-advisors. Because of their enormous encouragement and support during the most difficult (emotionally) year of my life, I could motivate myself towards my graduation. Without their help and wisdom of life, I wouldn't have accomplished my degree.

I would also like to thank my thesis reader Prof. Duane Boning for providing numerous helpful suggestions on the experimental design in this work and for reviewing the thesis on very short notice. I must also express a tremendous gratitude to Dr. Tom Korsmeyer for sharing his knowledge and providing valuable discussion of FastLap and numerics. His kindness and patience in helping me to implement the simulation procedure are greatly appreciated.

Another great thanks must go to my former officemates, Dr. Vladimir Rabinovich and Dr. Bart Romanowicz. Dr. Rabinovish implemented this preliminary work of the thesis project. I would like to thank him for his effort on proving the concept of this work, and the time and patience on leading me into this project. Dr. Romanowicz, an expert in computer systems and numerical simulations, has helped considerably in computer system administration issues and numerical simulation issues. Also, their affable personality was an important factor in making my stay at MIT a very enjoyable one.

Thanks much also go to Ding Meng for providing experimental results. I am also thankful to David Pflug and Leonard Dvorson for valuable discussion in FED design issues. Most of all, I would like to thank DARPA/ONR (contract number N00014-96-1-0802) and ONR/NFESC (contract number N00014-97-1-0827) for supporting my research.

Of course, the members of my research group have also been an essential part of this thesis. Erik Deutsch has given me useful comments on my thesis writing. Mathew Varghese also offered help on computer administration that was necessary to keep my research running. Without their help, my research would not have been completed on time. I would like to thank my former colleagues, including Dr. Elmer Hung, Dr. Lynn Gabbay, and Dr. Jan Mehner for discussion, advice and many other contributions. Also, I would like to thank to Scotti Fuller, Carolyn Zaccaria and Carol Nicolora for their help in administrational details.

Great thanks must go to a few close friends, Ginger Wang, Chih-Wei Chou and Jason Hu, outside of my research environment. Your friendship has made me appreciate the life of Boston over the past few years.

I also would like to thank Julia Hsu. Although we separated before the completion of this thesis work, I greatly appreciate her support and care during my first three years at MIT.

Finally, but most importantly, I would like to express my greatest gratitude my parents, Chow-Chi Yang and Cho-Lan K. Yang, and my sister Sophia Yang. Their encouragement and support has always been there when I needed most through the years of my graduate study abroad. I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Flat panel display technologies have received great attention in the display industry because of an increased interest in portable computation and information systems. The field emission approach is one of the promising technologies because of its excellent performance and low power consumption [1][2]. Although a great number of corporate, university, and national institute research groups have been conducting intensive research on *field emission devices* (FED), their focus is mainly on different fabrication techniques which allow developers to make a low-cost and high-yield product [3]. The full potential and optimal performance of field emission devices has not been fully explored for different designs and approaches because fabrication is expensive and time-consuming.

If the performance of a specific design of field emitters can be predicted precisely before the field emitters are physically implemented, the number of trial-and-error fabrication runs and the overall cost of development can be significantly reduced. The typical way to analyze the performance during the design phase is to use a *computer-aided-design* (CAD) tool. Analysis of FED requires accurate electric field and robust electron trajectory calculations. The most significant challenge in FED numerical modeling is the solution of the potential problem in which the dimensional scales of the tip region and the regions around and above the gate differ by about 3 orders of magnitude. To date, the reported works regarding to FED simulation tools either apply *three-dimensional* (3-D) *finite-element-method* (FEM) approaches, or *two-dimensional* (2-D) finite-element-method/*boundary-element-method* (BEM) approaches [4]-[21]. 3-D FEM approaches typically need several hundred thousand mesh-elements, so solutions require a few hours/days on an engineering workstation, even though the matrices generated by FEM are sparse. 2-D FEM and BEM approaches require less resources, but lack the ability to simulate emitter arrays because of the assumption of azimuthal symmetry. Analytical and Semi-analytical methods proposed in [22][23] provide efficient ways to estimate emission current or spot size, but these methods do not posses the flexibility of modeling various types of devices as FEM/BEM solvers do.

Furthermore, there are several tens of design parameters for a FED structure. Although the effect of each parameter on the device performance can be predicted by using a CAD tool, finding the optimal device performance needs a wise design strategy. This strategy is required, not only to avoid an intolerable number of simulations, but also to help explore the relationship between device performance and design parameters. Currently, there are no available FED CAD systems that provide the option of using a design strategy to characterize the device performance.

A 3-D CAD tool is presented for efficient simulation and modeling of FEDs. This CAD tool, which is leveraged by the interface support of MEMCAD [24], uses a BEM solver that is *accelerated* by a so-called "fast multipole" algorithm (FastLap [25]) that effectively makes the BEM matrix sparse. The computational cost of this BEM approach is order N, where N is the number of boundary elements. Although the cost of FEM approaches is also order N [26], this BEM approach is more efficient because it has a much lower number of elements. Comparison of experimental results and simulated results are also provided.

A design strategy is proposed for characterizing and optimizing FED performance. This strategy employs the *design of experiment* (DOE) technique [27], which explores the responses of a system by conducting a series of "experiments" that consist of a structured space of design parameters. The first step of the strategy is to screen the significant design parameters and build regression models. This step might need to be iterated with less or different design parameters, until the regression models describe the device performance reasonably accurate. The second step is to optimize the device performance by studying the regression models that describe the relationship between the design parameters and the device performance. The DOE technique used in this work is called *the central composite design* (CCD), which is capable of screening significant design parameters and building quadratic response surfaces (regression models). Two design examples with different electron-beam focusing schemes are demonstrated. A few important and interesting results predicted by the models can be used as both design and operational guidelines of field emission devices.

This thesis is divided into seven chapters. CHAPTER 2 describes the background of field emission theory and technology. FED applications and comparison with other flat-panel display technologies are also presented in this chapter. A brief description of boundary integral equations and boundary element formulations used in this work is presented in CHAPTER 3. The comparison of FEM and BEM is described. The discussion of why BEM was chosen in this study is also presented. The trajectory calculation method and the preliminary simulation results of two types of FEDs are discussed in CHAPTER 4. The experimental results are also presented. CHAPTER 5 briefly introduces the design of experiment techniques that are used in this work, including the two-level factorial design, the central composite design, the regression analysis, and the response surface method. CHAPTER 6 first presents an automation system that efficiently creates and automatically executes a series of simulations defined by a DOE technique. Five sets of design results for two different types of FEDs are presented. CHAPTER 7 contains the summary and conclusion of this work.

CHAPTER 2

FIELD EMISSION THEORY AND TECHNOLOGY

In this chapter, the physics of five types of electron emission mechanisms are briefly introduced and compared. Among these mechanisms, theory and derivation of the field emission mechanism is discussed. Various types of field-emission devices and their applications are also described. Finally, a brief introduction of a few dominant display technologies is presented, and a comparison is made between these technologies and the field-emission display technology.

2.1 Electron Emission Theory

2.1.1 Electron Emission Mechanisms

For an energy barrier with or without applied voltage bias, there are five electron emission mechanisms: photo emission, thermionic emission, Schottky emission, field emission, and thermally–assisted field emission [28][29][30]. Figure 2.1.1 shows the energy band diagrams for a metal near the surface when there is no applied bias voltage. Photo emission takes place when there is an energy transfer from a photon to an electron, and the energy acquired by the electron overcomes the work function (surface barrier energy) of the metal, as shown in Figure 2.1.1(a). This energy is also called photo-ionization energy.

Thermionic emission is dependent on temperature. As temperature is elevated, the electrons that acquire thermal energy greater than the work function can overcome the barrier. More precisely, the energy distribution of electrons can be presented by the Fermi-Dirac distribution function, as shown in Figure 2.1.1(b). The distribution curves with different temperature indicate that higher temperature creates more electrons that have enough energy to overcome energy barrier. Thermionic emission is also the standard way of emitting electrons for typical *cathode-ray tubes* (CRT) [31].

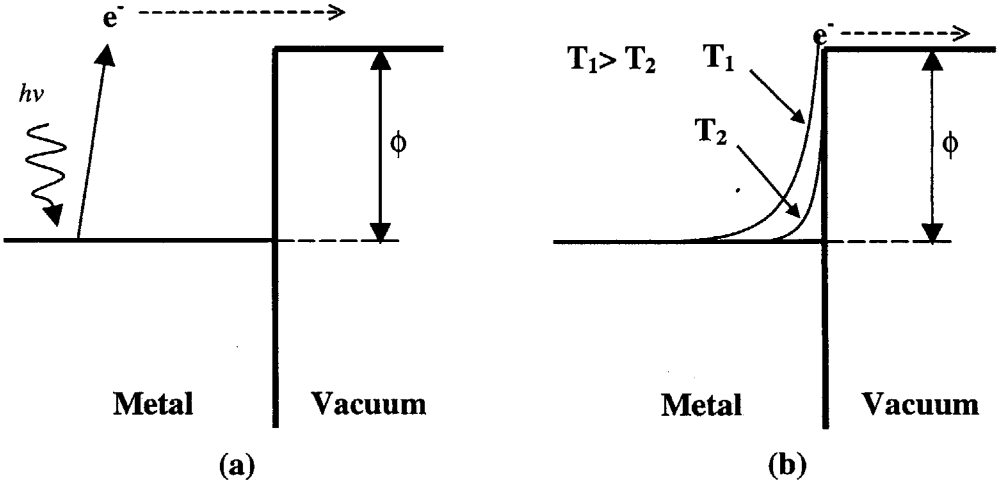


Figure 2.1.1 Electron emission mechanisms : (a) photo and (b) thermionic emissions. Both the mechanisms are under zero or light bias.

When a field is applied to a metal surface, the potential barrier should become triangular in shape; however, the barrier is also deformed and lowered by the image potential of emitted electrons as shown in Figure 2.1.2(a). The field-dependent emission that occurs due to the barrier lowering is known as the Schottky emission.

As the electric field increases, the barrier becomes lower as well as narrower. When the barrier is sufficiently narrow, electron tunneling takes place, as shown in Figure 2.1.2(b). This emission mechanism is called field emission or cold electron emission because it can occur even at extremely low temperature (e.g., 0 K). Field emission is also called the Fowler-Nordheim emission because Fowler and Nordheim first gave a thorough study to explain this emission mechanism [32].

Thermally assisted field emission is similar to field emission, except that the electrons are elevated to higher energy level due to absorption of thermal energy before they tunnel through the narrow barrier.

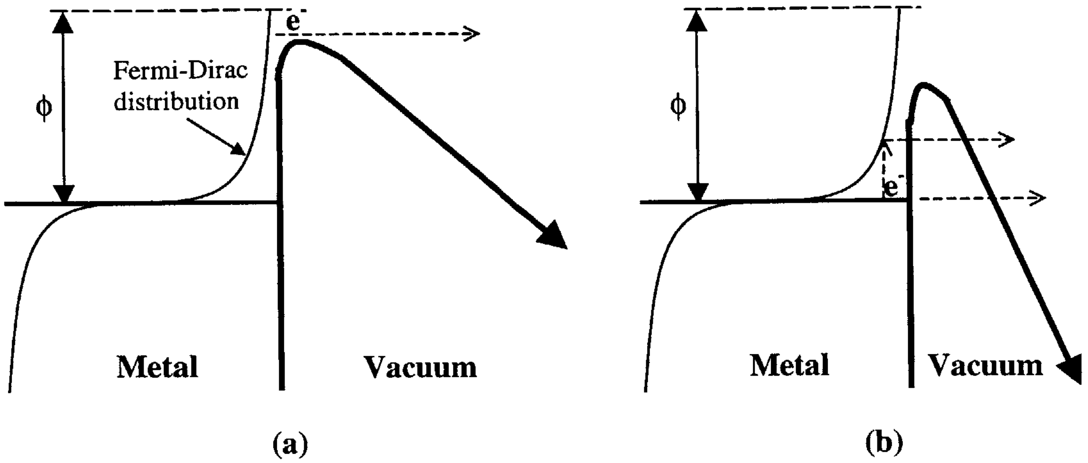


Figure 2.1.2 Electron emission mechanisms : (a) Schottky emission, and (b) field emission and thermally-assisted field emission.

2.1.2 The Fowler-Nordheim Tunneling Theory

Field emission is the emission mechanism discussed in this work. In order to obtain high electric field with reasonable applied voltage, a typical approach is to fabricate structures with small radius of curvature. The width and shape of the barrier, as shown in Figure 2.1.2, and the supply of electrons are the two factors that determine the tunneling current. Intuitively, for a certain energy state, the width and shape of the barrier determines the transmission probability, and the supply of electrons can be calculated from the occupation probability.

The tunneling current density for a specific energy level is the product of the incident flux, the transmission probability per electron state, and the occupation probability of this state [33]. For a one-dimensional potential barrier, the current density is the energy integration of the product of transmission probability and the electron supply function:

$$J = e\int_0^{\infty} D(E_x)N(E_x)dE_x \tag{1}$$

where $D(E_x)$ is the transmission probability which can be calculated by the WKB method and is given by

$$D_{WKB}(E_x) = \exp\left(-2 \cdot \int_{x_1}^{x_2} \kappa(x) dx\right) \tag{2}$$

where

$$\kappa(x) = \sqrt{\frac{2m}{\hbar}\left(V(x) - E_x\right)} \tag{3}$$

and $V(x)$ is the potential barrier and $\kappa(x)$ vanishes at x_1 and x_2, which are the classical turning points. By using the WKB approximation, D_{WKB} is given by

$$D_{WKB}(E_x) = \exp\left(-\frac{4}{3e\varepsilon}\left(\frac{2m}{\hbar^2}\right)^{\frac{1}{2}}(\phi - E_x - E_F)^{\frac{3}{2}}\right) \tag{4}$$

where $N(E_x)$ is the electron supply function defined by

$$N(E_x) = \frac{4\pi m}{\hbar^3}\int_{E_x}^{\infty} f(E) dE \tag{5}$$

where $f(E)$ is the Fermi-Dirac distribution function defined as

$$f(E) = \frac{\exp\left(\frac{E_F - E}{k_b T}\right)}{1 + \exp\left(\frac{E_F - E}{k_b T}\right)} \tag{6}$$

Taking the integral in (5) results in the closed form of $N(E_x)$ given by

$$N(E_x) = \frac{4\pi m k_b T}{\hbar^3}\ln\left(1 + \exp\left(\frac{E_F - E_x}{k_b T}\right)\right) \tag{7}$$

With further approximation developed by Schottky, Fowler and Nordheim derived the famous Fowler-Nordheim equation, which describes the emission current density distribution as a function of the electric field distribution on the emitter surface, and the work function of the material.

$$J = \frac{e^3 \cdot F^2}{8\pi \cdot \hbar \cdot \phi \cdot t^2\left(\frac{\left(e^3F\right)^{\frac{1}{2}}}{\phi}\right)} \exp\left(-\frac{4\sqrt{2m} \cdot \phi^{\frac{3}{2}}}{3\hbar \cdot e \cdot F} v\left(\frac{\left(e^3F\right)^{\frac{1}{2}}}{\phi}\right)\right) \tag{8}$$

$$= \frac{e^3 \cdot F^2}{8\pi \cdot \hbar \cdot \phi \cdot t^2(y)} \exp\left(-\frac{4\sqrt{2m} \cdot \phi^{\frac{3}{2}}}{3\hbar \cdot e \cdot F} v(y)\right)$$

where $y = \frac{\left(e^3F\right)^{\frac{1}{2}}}{\phi}$, F is the electric field distribution on the tip (in V/cm), ϕ is the work function of the emitter material (in eV), e is electron charge, $\hbar$ is Planck's constant, and m is the mass of an electron.

Note that $t^2(y)$ and $v(y)$ can be approximated as

$$t^2(y) = 1.1 \tag{9}$$

$$v(y) = 0.95 - y^2 \tag{10}$$

or

$$v(y) = \cos(\frac{1}{2}\pi \cdot y) \tag{11}$$

For a typical material used to fabricate emitter tips, such as Mo and Si, the work function is in the range of 4.0 eV to 4.5 eV, corresponding to a field strength of 10^9 V/m that is necessary for field emission.

A simplified form of the Fowler-Nordheim equation is

$$J = A \cdot \frac{F^2}{\phi \cdot t^2(y)} \exp\left(-B \cdot \frac{\phi^{\frac{3}{2}}}{F} v(y) \right) \tag{12}$$

where J is in A/cm^2, $A = 1.54 \times 10^{-6}$, $B = 6.87 \times 10^7$ and $y = 3.79 \times 10^{-4} \cdot F^{1/2} / \phi$.

The I-V characteristic of field emission is more useful than the J-E characteristic. The emission current density and the field distribution on the tip can be represented as:

$$J = I / \alpha \tag{13}$$

$$E = \beta \cdot V \tag{14}$$

where α is the effective emission area (cm^2) , β is the field enhancement factor (cm^{-1}) , I is the total emission current (A), and V is applied voltage on the gate (V).

Equation (12) can be rewritten as

$$I = \alpha \cdot A \cdot \frac{\beta^2 \cdot V^2}{\phi \cdot t^2(y)} \exp\left(-B \cdot \frac{\phi^{\frac{3}{2}}}{\beta \cdot V} v(y) \right) \tag{15}$$

or

$$I = a_{FN} \cdot V^2 \cdot \exp(-b_{FN} / V) \tag{16}$$

where

$$a_{FN} = \frac{\alpha \cdot A \cdot \beta^2}{1.1 \cdot \phi} \exp\left(\frac{B \cdot 1.44 \times 10^{-7}}{\phi^{1/2}} \right) \tag{17}$$

and

$$b_{FN} = 0.95 \cdot B \cdot \phi^{3/2} / \beta \tag{18}$$

Note that the derivation of the Fowler-Nordheim equation starts with the assumption of a flat emission surface (one-dimensional barrier). The equation is valid when the radius of curvature

(ROC) of the emitter tip is substantially greater than the barrier width so the tip surface can be considered "flat". The typical barrier width at which tunneling occurs is 1 nm. For the devices with sharp tip ROC close to the barrier width, the emission current density calculation can be performed by applying modified Fowler-Nordheim equations [34].

2.1.3 The Fowler-Nordheim Plot

Equation (16) can be rewritten as

$$\ln\left(\frac{I}{V^2}\right) = \ln(a_{FN}) - b_{FN} \cdot \left(\frac{1}{V}\right) \tag{19}$$

Plotting this equation in the form of $\ln\left(\frac{I}{V^2}\right)$ vs. $\left(\frac{1}{V}\right)$, a straight line of slope of $-b_{FN}$ is obtained with an intercept of $\ln(a_{FN})$ at $\left(\frac{1}{V}\right) = 0$. In the field emission study, the experimental I-V data is often plotted in the form of $\ln\left(\frac{I}{V^2}\right)$ vs. $\left(\frac{1}{V}\right)$ so that a_{FN} and b_{FN} can be easily extracted. This type of plot is called Fowler-Nordhein plot. Figure 2.1.3 is an example of a Fowler-Nordheim plot for an emitter array.

Because of the requirement of high electric field, micromachined tips are typically used as emitter tips. The radius of curvature of the tiny tips, however, are usually too small to be measured accurately. The extraction of a_{FN} and b_{FN} from experimental Fowler-Nordheim plots enables the extraction of the field factor and the effective emission area, and thus the ROC of emitter tips can be reasonably estimated. In this study, the emission area can be accurately estimated from numerical models, so the effective emission area, which is usually used for analyzing experimental data, will not be discussed. Detailed discussions about the effective emission area can be found in [12][35][36].

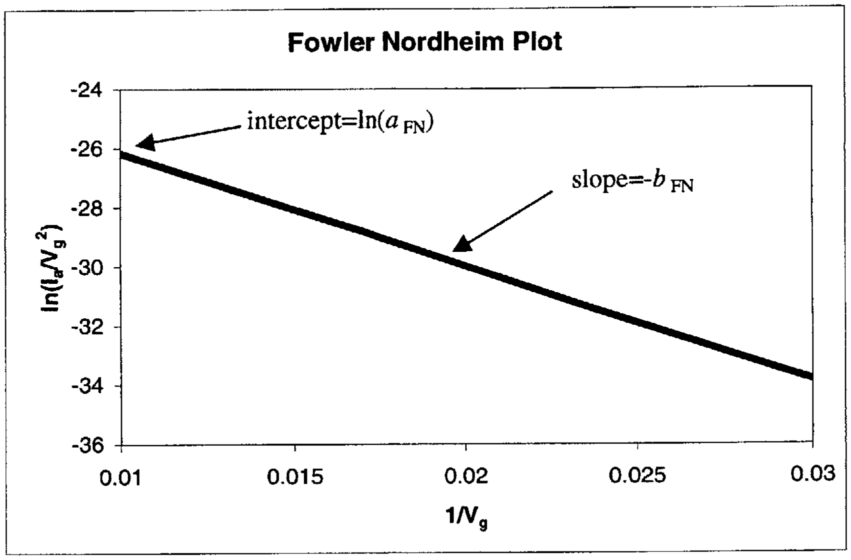


Figure 2.1.3 Example of a Fowler-Nordheim plot.

2.2 Field Emission Display and Technologies

2.2.1 Introduction

Field emission devices have received great attention recently. Among various applications of field emission devices, the field emission display has attracted most activity. In this section, the fundamental principles of the field emission displays are addressed based on the theory discussed in the previous sections, and different types of technologies used in field emission displays are discussed [37].

Figure 2.2.1 shows a side-view schematic of a field emission display. A typical device consists of arrays of emitter structures. Each emitter structure is composed of a cone with a very sharp tip and a gate with an annular aperture surrounding the cone tip. A typical device has a third electrode for collecting the emitted electrons. When a voltage is applied to the gate, a high electrostatic field exists at the tip surface allowing electrons to tunnel out the tip. The emitted electrons are collected by the anode that is biased at higher voltage than the gate. In essence, the gate modulates the anode current.

The emitted electrons diverge from the cone axis as they move towards the anode because of their off-axis velocity on leaving the gate region. The electrons can be focused by adding an additional electrode (the focus) between the gate and the anode. The focus, which is biased at a lower voltage than the gate, repels electrons towards the cone axis thus making them converge.

A typical field emission display (FED) consists of two plates: (a) the base plate, and (b) the face plate. These two plates are separated by dielectric spacers in a vacuum envelope. The base plate consists of a two-dimensional array matrix of addressable electron sources arranged in row and column fashion. The emitters (cones) are connected to the rows and the gates are connected to the columns. A voltage applied to the row and another voltage applied to the column results in a voltage difference between the gate and the emitter, leading to electron smission by the cones.

The face plate consists of indium-tin oxide (ITO) covered glass with a two dimensional array of phosphor dots that correspond to the anodes for individual electron source on the base plate. The phosphor dot and the electron source together form a pixel of the display.

A pixel is activated by the application of a row select voltage to the row connected to the pixel, and a data voltage to the column connected to the pixel, resulting in sufficient voltage difference between the gate and emitter for field emission. The electrons are accelerated towards the phosphor screen where the energetic electrons lead to photon emission from the pixel. All non-activated pixels do not have sufficient voltage differences between their gates and emitters for field emission, hence the pixels remain dark.

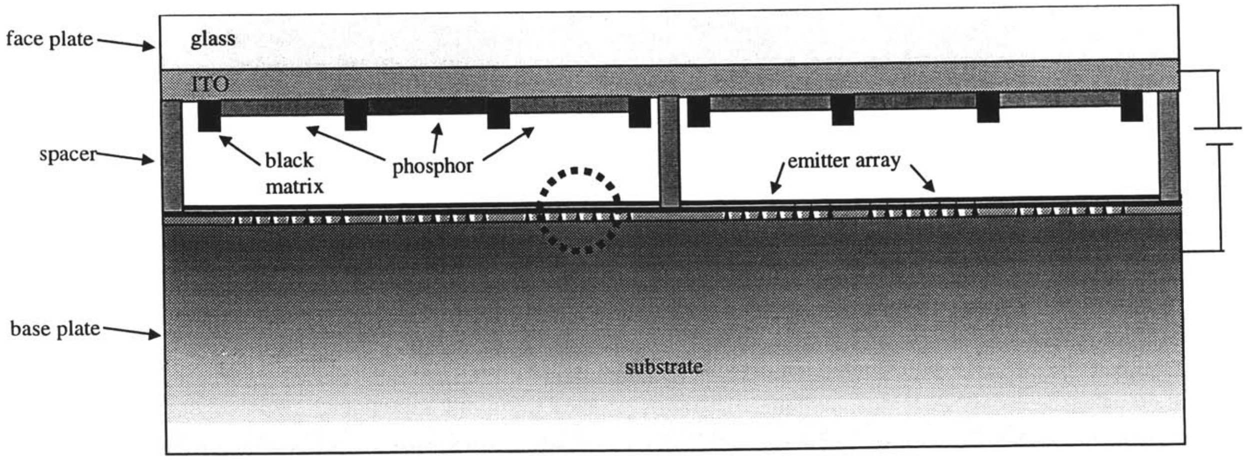


Figure 2.2.1 Schematic of a field emission display. A detailed schematic of the emitter array (circled) is shown in Figure 2.2.2.

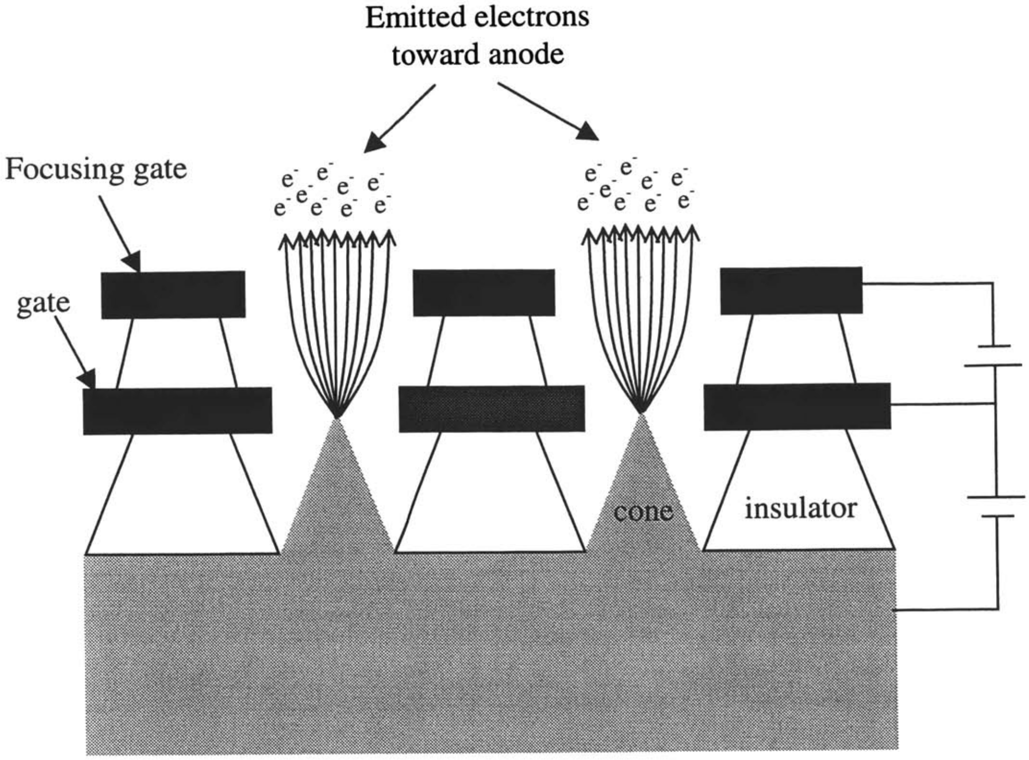


Figure 2.2.2 A detailed view of emitters in an array.

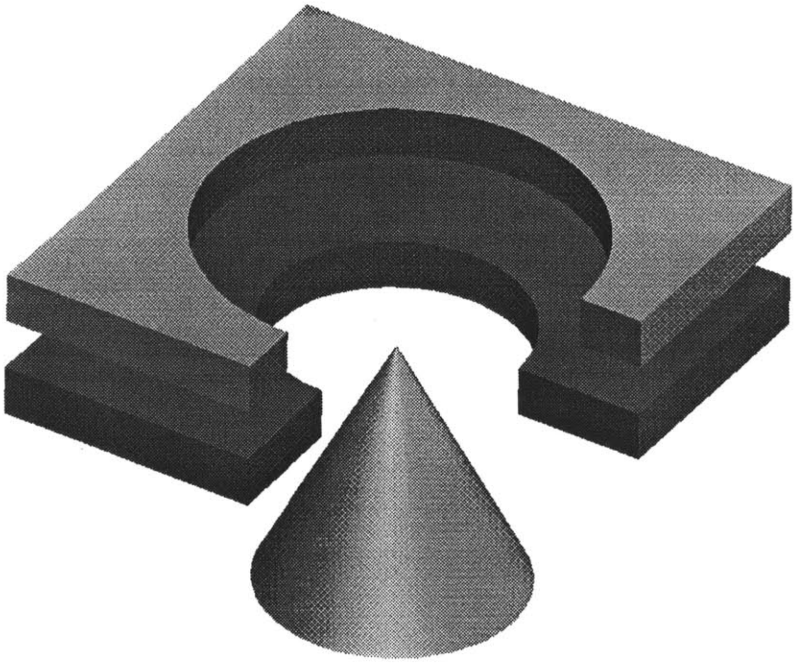


Figure 2.2.3 Three-dimensional schematic of an emitter structure.

The field emission display is essentially a flat CRT display, and consequently it has all the characteristics and advantages of the cathode ray tube such as:

- High spot and screen brightness
- High luminous efficiency
- Temperature and radiation insensitivity
- Wide viewing angle
- Dynamic range of 10 bits of gray scale/color

It combines the positive attributes of CRTs with the positive attributes of flat panel display technologies such as:

- Thin profile
- Matrix addressing : higher screen brightness
- Light weight
- Low power

2.2.2 Other Types of Emitter Structures

In addition to the cone-shaped emitters described previously, there are a few other types of emitters that have been proposed and fabricated. Figure 2.2.4 and Figure 2.2.5 are the schematics of a ridge-shaped emitter and a thin-film edge emitter. In [38], the maximum electrical field on the tip of a cone emitter can be approximated as

$$F_{\max} = \frac{V_g}{r} + \frac{V_g}{d-r} \tag{20}$$

where V_g is the gate voltage, d is the radius of the gate aperture, and r is the tip radius of curvature. Similarly, the maximum electric field obtained by the ridge-shaped emitters is approximated as [20]

$$F_{\max} = \frac{V_g}{r \cdot \ln\left(\frac{r+d}{r}\right)} \tag{21}$$

where r is the radius of curvature of the edge, and d is the distance from the gate to the edge.

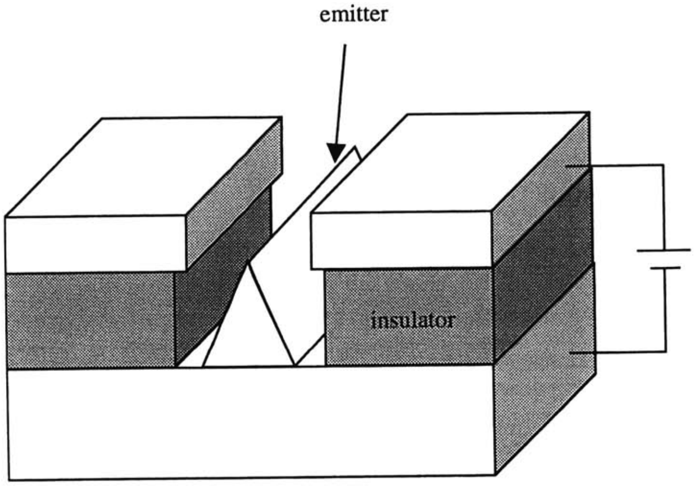


Figure 2.2.4 Schematic of a ridge-shaped emitter.

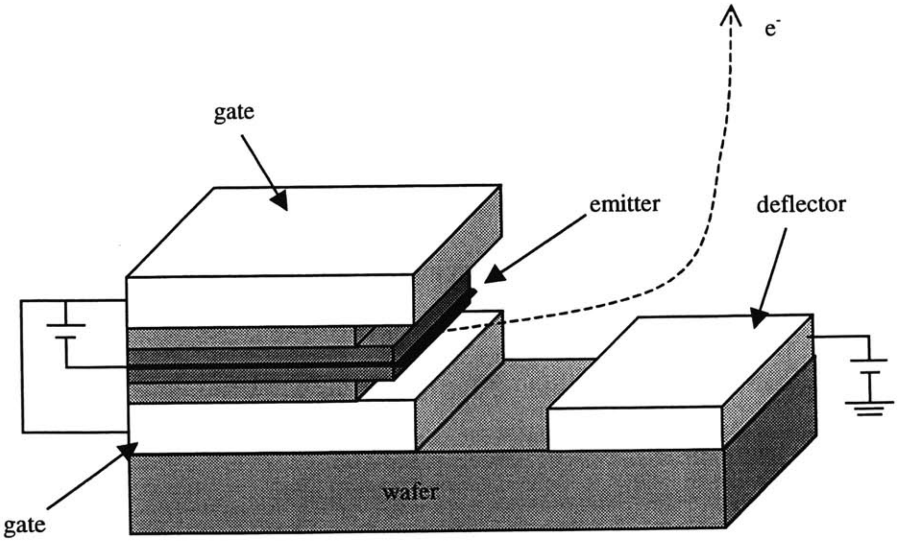


Figure 2.2.5 Schematic of a thin-film edge emitter.

From equations (20) and (21), it can be observed that the maximum electric field from the ridge-shaped emitters is smaller than that from the cone-shaped emitters, assuming that both radii of curvature are equal. Therefore, higher operating voltage is needed for the ridge-shaped emitters.

The thin-film edge emitter is a lateral device fabricated by surface micromachining techniques. With a deflection electrode, the emitted electron beam can be deflected into the vertical direction. This device has proved to be a good candidate for backlighting lamp for LCDs [39].

2.2.3 Focusing Schemes of FEDs

The cone-shaped emitters are the most accepted design in the industry for the following reasons:

- Electrons are emitted vertically and the electric field is symmetrical about the cone axis, so the majority of electrons go to the anode without additional deflecting electrodes.
- Additional electrodes can be easily added to improve spot size and luminous efficiency.

In general, there are several types of focusing schemes for the cone-shaped emitters:

- Proximity focusing
- Local in-plane integrated focusing using an integrated-focus electrode (IFE) [40]
- Local out-of-plane integrated focusing [41][42]
- Global in-plane integrated focusing [43]
- Global out-of-plane integrated focusing [44]
- External focusing grid [45]

A detailed survey about those focusing schemes can be found in [46].

In terms of structural complexity, the simplest scheme is the 'proximity focusing'. This method keeps the separation between the anode and the gate as small as 0.2 mm, and thus the emitted electrons can be collected by the anode before diverging. Pixel crosstalk is also eliminated. Since no additional focusing components are needed, this scheme is also referred to as a passive focusing scheme. Although the proximity focusing scheme is simple to fabricate and gives reasonable resolution, the screen voltage that can be applied is limited by the narrow gate-anode distance, and thus requires the use of low-voltage phosphors. The low-voltage phosphors typically require only 200-800 volts of anode voltage vs. 3000-5000 volts for high-voltage phosphors. Unfortunately, low-voltage phosphors have two serious drawbacks: low luminous efficiency and aging problems, which make the performance of FEDs using low-voltage phosphors unsatisfactory.

On the other hand, high-voltage phosphor technology is mature and has been widely used in CRT displays for decades. In order to use high-voltage phosphors, however, the gate-anode distance usually needs to be about 1 mm or larger. Without any focusing scheme, the electron beam spreads out and resolution decreases. A few other focusing schemes have been proposed to solve the trade-off between the use of high-voltage phosphors and resolution. Since these schemes require extra components (integrated-focusing electrodes or external grids) to create focusing effect, they are also called the active focusing schemes.

For the local integrated focusing schemes, a focusing electrode is integrated on each emitter structure. For the global integrated focusing schemes, a focusing electrode is integrated on each emitter array. In general, the local integrated focusing schemes provide better beam collimation, but are more difficult to fabricate.

The local and global integrated focusing schemes can be classified into two categories: in-plane integrated focusing and out-of-plane integrated focusing. The focusing electrode for the in-plane focusing scheme is located on the same plane of the gate(s), but surrounds the gate(s), as shown in Figure 2.2.6 and Figure 2.2.7. In general, the local or global in-plane integrated focusing schemes require a larger area because the integrated-focusing electrodes and the gates are on the same plane, and hence reduce the density of emitters (packing density).

For the out-of-plane integrated focusing schemes, the focusing electrode is located above the gate, as shown in Figure 2.2.2 (local out-of-plane scheme only). This provides the most effective focusing effect and highest packing density, but is also the most difficult to fabricate because of the additional layers. Also, since the integrated-focusing electrodes are close to the tips, the voltage drop between the gate and the integrated-focusing electrode gives rise to a reduction in emission current. Fortunately, this can be easily overcome by increasing the gate voltage.

The external-focusing-grid scheme uses a focus grid inserted between the cathode and the anode. One opening on the grid is for each emitter array (pixel). The focusing grid not only actively converges the electron beams from emitter arrays, but it also intercepts stray electrons and shields the cathode from the bombardment of the positive ions created at the anode.

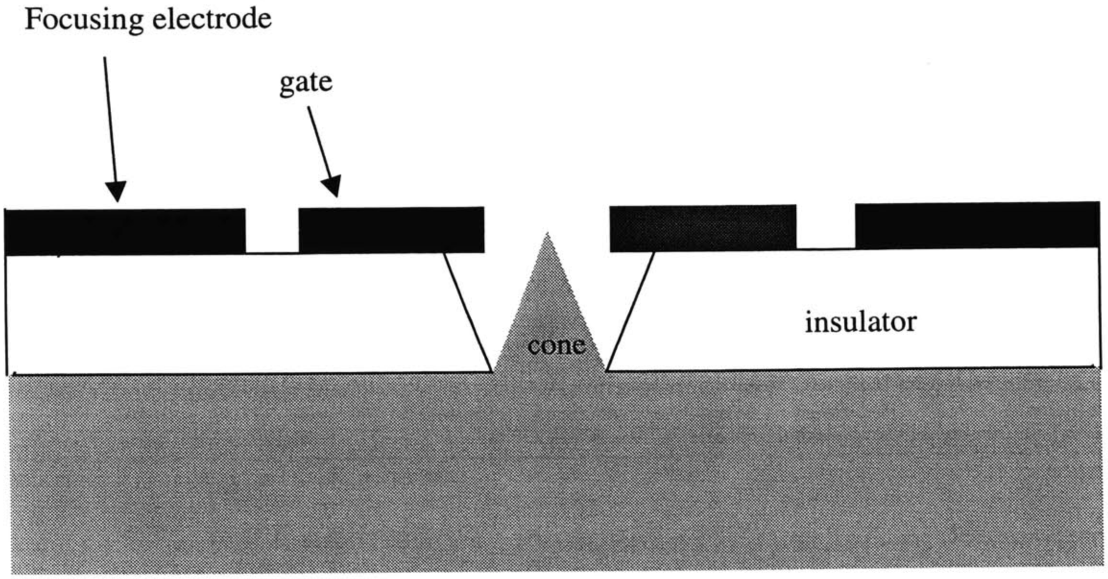


Figure 2.2.6 local in-plane focusing scheme.

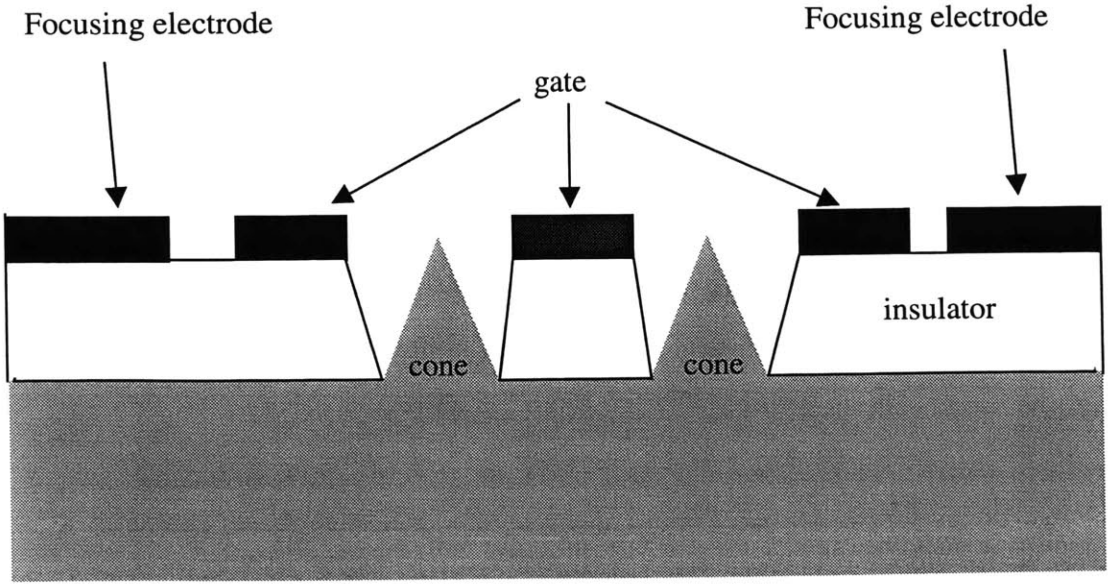


Figure 2.2.7 global in-plane focusing scheme.

2.2.4 Technical Issues of FEDs

Finally, there are several key technology areas that need to be solved to successfully commercialize field emission displays [47][48][49]. They are: Pixel alignment and focusing, phosphor aging and efficiency, small area uniformity, consistency in cathode performance, obtaining consistently low drive voltage of less than 10 V, making spacers invisible, acceptable yield with regard to row outs and column outs, cost effective assembly, sealing and packaging, maintaining high vacuum over time, and demonstration of 256 gray levels per primary. For the time being, no company has completely overcome these issues completely yet.

2.3 Discussion of Dominant Display Technologies

In this section, we will give a brief introduction to the display technologies [50][51][52][53][54][55] that are widely accepted in the market, and compare these technologies with the field emission technology (Table 2.3.1 [56]).

2.3.1 Cathode-Ray-Tube (CRT) Displays / Projection TV

Cathode-ray Tubes can be considered antique devices. The invention of cathode-ray tubes was as early as the late 1800s, but they are most commonly used in television sets and computer displays, and are the dominant force in the display industry today. The cathode-ray tube is a large vacuum tube with a single thermionic emission source. Electrons emitted from the heated emission source travel across the depth of the tube and finally hit the phosphor screen on the front. The screen images are "written" by horizontal and vertical deflection coils [31].

Projection TVs are essentially the same as CRT displays. A high quality and small CRT is used to generate an image, which is magnified through an optical lens (and a mirror) and projected onto a large screen. Because of the special arrangement of the CRT, the lens, the mirror and the screen, the projection TVs are thinner than regular CRTs with the same screen sizes.

The principal disadvantage of the CRT is its bulky and heavy glass tube, which has to be deep enough to allow the electron beam to be scanned on the entire screen, and also needs to be strong enough to prevent ambient pressure from crushing the vacuum tube. This disadvantage also makes CRTs completely excluded in portable systems. Despite the disadvantage of its bulky size and heavy weight, the CRT technology is the mainstream display technology because of:

- Simplicity of electronics due to the use of single serial data input
- Sufficient screen brightness, fast refresh rate and full color

2.3.2 Liquid-Crystal Displays (LCD)

Liquid-crystal displays (LCD), the most dominant flat-panel display technology, use addressable "light valves" to modulate the color/gray-scale of an image. The light valves consist of layers of a rear polarizer, an ITO pixel electrode, a polarizable liquid-crystal material, filters, color filters and a front polarizer. The layer of the "light valve" is on top of a backlight/diffuser. An electric field across the liquid-crystal material can rotate the orientation of molecules in the liquid crystal. The intensity of the light allowed through the liquid crystal is based on the rotational angle of the liquid crystal. The absence or presence of the field at any particular point of the liquid crystal layer determines whether and how much light passes through the "light valve" at that point. In a typical backlight LCD, only about 5 % of the original light intensity actually reaches the screen. *Active-matrix liquid-crystal displays* (AMLCD) improve the performance (dynamic range and non-linear response) by using a transistor on each pixel (see Figure 2.3.1).

The recent decrease in the price of AMLCDs boosts the sales of portable computers, but they still have many drawbacks:

- The requirement of a uniform and controlled light source (backlight)
- The low transmission of the "light valve" results in inefficiency in power consumption in portable systems.
- Slow refresh rate due to slow response of liquid crystal
- Sensitivity to temperature and pressure

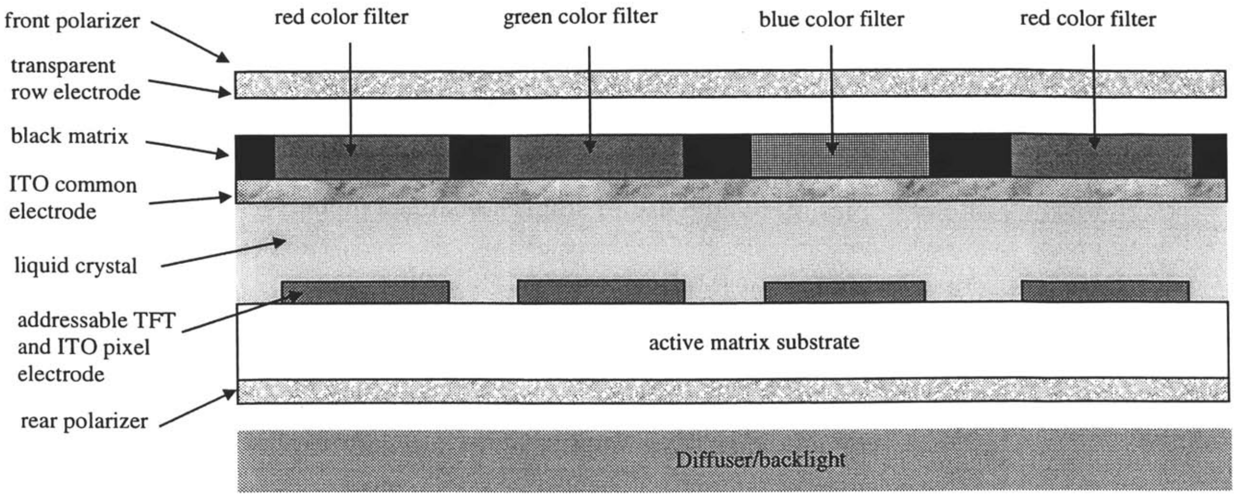


Figure 2.3.1 Schematic of an AMLCD.

2.3.3 Plasma Display Panels (PDP)

A *plasma display panel* (PDP) is a two-dimensional array of tiny neon lamps. The light is generated when a gas is rendered electrically conducting. A PDP consists of two glass substrates separated by a dielectric spacer to form a chamber. A noble gas, such as Ne, is sealed inside the chamber. Transparent conductors patterned on the two glass substrates are orthogonal to each other and face the gas chamber, and thus form a row and column addressing matrix. Phosphor dots are also coated on one of the glass substrate between conductor lines. When a high voltage is applied between the row and column electrodes, a plasma is produced at the intersection of the two electrodes generating UV emissions. Light is emitted from the phosphor that is excited by the UV radiation generated by the plasma. Figure 2.3.2 is a schematic of a plasma display.

The disadvantage of PDP is that a sufficient volume of gas must be present to create a satisfactory intensity of light, which places a limitation on the resolution and portability. The property of omni-directional light emission also leads to another limitation on resolution.

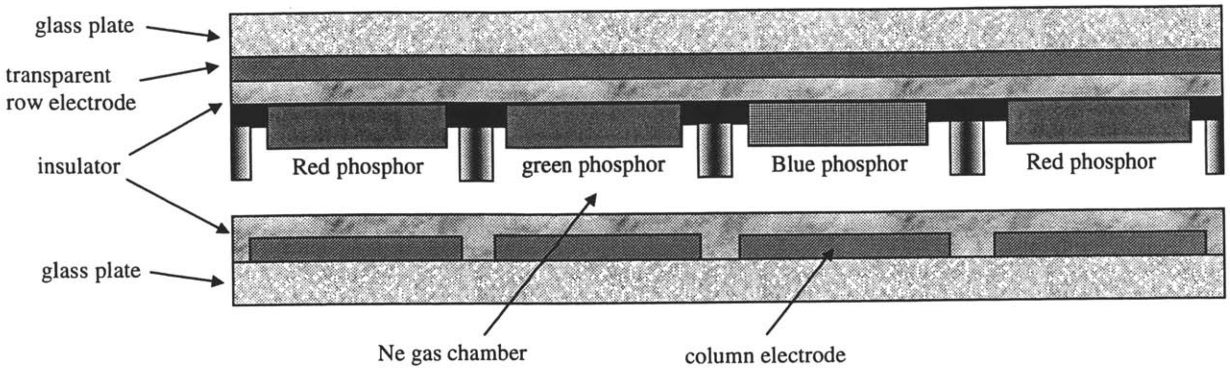


Figure 2.3.2 Schematic of a plasma display panel.

2.3.4 Electroluminescent Displays (ELD)

Similar to plasma displays, *electroluminescent displays* (ELD) are composed of a two-dimensional array of light-emitting regions. Phosphor dots are placed between conductors. Figure 2.3.3 shows a schematic of an ELD. Electroluminescent displays are the thinnest among all flat-panel displays. When the potential applied across the phosphor is high enough to cause a breakdown, hot electrons created by the breakdown excite the phosphor and generate light. The main advantage of ELDs is the non-linear response (sharp turn-on threshold) required for better performance. However, ELDs also suffer from:

- Low luminous efficiency
- Requiring a delicate circuit because of high driving voltage
- High capacitance leading to low refresh rate
- Difficulty of making full-color ELDs due to the lack of good blue phosphor

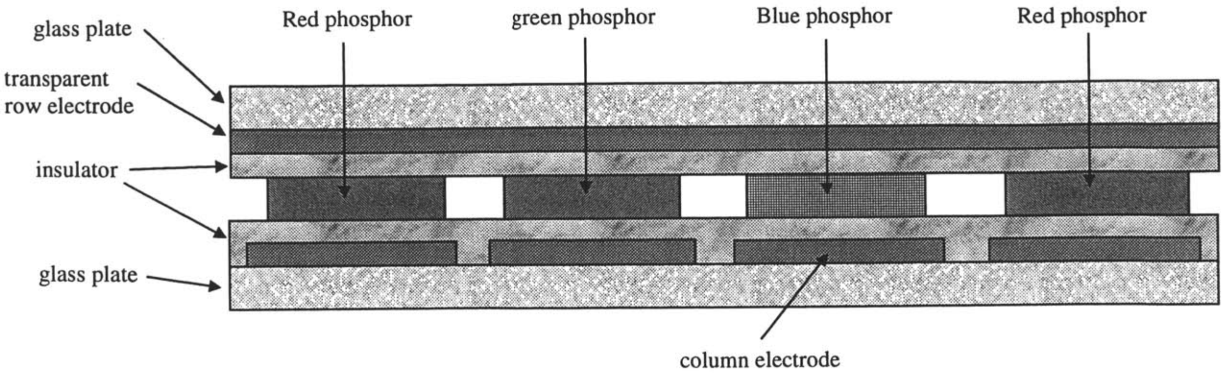


Figure 2.3.3 Schematic of a electroluminescent display.

2.3.5 Vacuum Fluorescent Displays (VFD)

Vacuum fluorescent displays (VFD) are similar to CRT because of the same photon emission mechanism (cathodoluminescence emission) for exciting phosphors. A large-area thermionic source of electrons, which consists of hot wires, is essentially a large-area cathode (see Figure 2.2.4). Electrons emitted from the cathode are modulated by a series of x-y addressable metal grid. Light is generated if electrons are accelerated through the grid. Otherwise the electrons are repelled by the grid. The main problems with VFDs are:

- Lack of low-voltage phosphors making full-color display impossible
- Unnecessary power consumption due to persistent emission of electrons during operation.

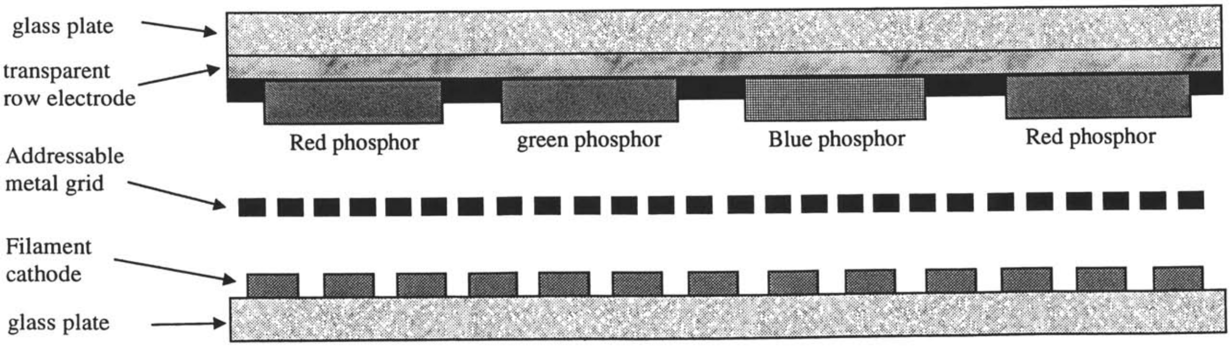


Figure 2.3.4 Schematic of a vacuum fluorescent display.

Performance	**CRT**	**LCD**	**ELD**	**VFD**	**PDP**	**FED***
Power consumption	200 W	100 W	40-50 W	60 W	60-80 W	15 W
Contrast ratio	Excellent	Good	Good	Good	Good	Excellent
Viewing Angle	Excellent	Excellent	Good	Excellent	Excellent	Excellent
Luminance	350 Cd/m^2	350 Cd/m^2	100 Cd/m^2	70 Cd/m^2	70 Cd/m^2	700 Cd/m^2
Color	Best	Full	Green/Yellow	Full	Full	Best
Resolution	Excellent	Excellent	Excellent	Satisfactory	Satisfactory	Excellent
High-ambient light readability	Almost sunlight viewable	Sunlight viewable	Sunlight viewable	Poor	Satisfactory	Sunlight viewable
Frame rate	60 Hz	60 Hz	60 Hz	60 Hz	60 Hz	60 Hz
Pixel matrix	2048 x 2048	2560 x 2048	1280 x 1024	400 x 600	1280 x1024	1280 x 1024
Resistance to:						
Temperature	Excellent	Poor	Satisfactory	Satisfactory	Satisfactory	Excellent
Humidity	Satisfactory	Poor	Satisfactory	Satisfactory	Satisfactory	Excellent
Shock & Vibration	Satisfactory	Excellent	Excellent	Satisfactory	Excellent	Excellent
Luminous Efficiency	0.16 lm/W	0.32 lm/W	0.15 lm/W	0.11 lm/W	0.09 lm/W	4.25 lm/W
Display Depth	14 in	5 in	2 in	2 in	2 in	2 in
Figure of Merit (FOM)	1.14×10^{-2}	6.40×10^{-2}	7.50×10^{-2}	5.50×10^{-2}	4.50×10^{-2}	2.13
Normalized FOM	1	5.6	6.6	4.8	3.9	187

Table 2.3.1 Comparison of display technologies [56]. The definition of the FOM in this table is the ratio of luminous efficiency to display depth.

2.4 Other Applications of Microtips

In addition to display applications, the microtips used in the field emission displays can be applied to numerous applications [57]. In this section, we will briefly describe some of the important and interesting applications of microtips.

- *Radio Frequency (RF) and microwave applications* [58][59]: Because of the advantage of vacuum as the transport medium of electrons, the field emitter arrays have excellent performance in the microwave or higher frequency range. A typical device in RF/microwave applications is a *traveling wave tube* (TWT). In a TWT, a field emitter array can be used as an electron source. The size and weight of a microwave tube can be significantly reduced due to the high intensity and fine electron beam generated from the emitter array. The field emitter array can also be used as a triode [60]. Spindt et. al. [61] has demonstrated a power amplification at 1GHz by using a low capacitance *field-emission arrays* (FEA).

- *Sensors (accelerometers, pressure sensors, infrared detectors, and etc.)* [62][63]: Microtips have been proven as sensitive transducers to detect a small variation in distance. When a voltage is applied between a conductive microtip and a conductive surface, the tunneling current between the tip and the surface is exponentially and inversely proportional to the distance between the tip and the surface. Therefore, the small variation in distance between the tip and the surface can be amplified and thus easily measured. For example, microtips can detect the small motions of a proof mass of an accelerometer to determine acceleration. Similarly, the deformation of a micromembrane due to pressure can be measured by a microtip to determine pressure. A Goley-cell *infrared* (IR) detector employs a sealed flexible chamber that inflates when the air inside the chamber expands due to the absorption of an IR radiation by the IR absorber material in the chamber. The expansion of the chamber can be detected by a microtip so that the absorption of IR radiation can be measured.

- *Electronic Cooling* [15][64]: The Fowler-Nordheim equation describes electron tunneling as function of surface electric field distribution. We can modify the equation to account for the energy gained or lost due to the transportation of electrons. It is called the Nottingham heating/cooling effect. Under some specific circumstances, the Nottingham effect provides cooling. We can apply this cooling effect to fabricate micro cooling system based on emitter arrays. This cooling system can be used for electronic cooling.
- *Electronic Storage* [65]: Microtips can also be used as data storage devices based on the atomic force microscope, in particular thermomechanical recording. For data storage, a microfabricated cantilever with resistively heated tips can be used for writing data marks on a rotating polycarbonate disk. A microfabricated cantilever combining a sharp tip with an integrated piezoresistive sensor can be used for data readback from a rotating polycarbonate disk. The readback process is similar to the operation of atomic force microscopes. Since the size of marks written by the heated tip is close to the atomic scale, the storage density is estimated to be as high as about 500 Gbit/in^2.
- *Ion Propulsion*: A microtip array can also be used as an ionizer for the ion source of an ion propulsion system. Ion propulsion received attention in the space technology industry since the 1960s. It can be used for the propulsion system of satellites [66]. The microtip structures are similar to the emitter structures for field emission display. With a positive voltage difference applied between a microtip and a gate, noble gas

(xenon or krypton) molecules are ionized when flowing close to the microtip where a very strong electric field exists. The ions are repelled by the electric field around the microtip and thus create propulsion. The total propulsion can be increased substantially by an ionizer made of a large microtip array.

- Ion Beam Mass Spectrometer (IBMS): Similar to ion propulsion, microtip arrays can be used as an ionizer for the ion-beam mass spectrometer [67]. The structure and operation configuration of the microtip arrays for IBMS are similar to the those for ion propulsion. Gas molecules (samples) are ionized and repelled from the tips. A magnetic field can redirect the ion beams to hit on different locations based on their molecular weights, and thus the constitution of the sample can be determined.

CHAPTER 3

THE BOUNDARY ELEMENT METHOD IN ELECTROSTATICS ANALYSIS

For electrostatic simulations, there are two main types of solvers in the industry: the *finite-element method* (FEM) solvers and *the boundary-element method* (BEM) solvers. The approach of the finite-element method receives wide acceptance in the industry due to its availability, so most of the modeling work in field-emission device studies uses finite-element method. In this research work, the BEM solver FastLap is used for electrostatics simulations. The choice of this BEM solver provides significant advantages throughout the whole simulation process. In the first section, a brief discussion and comparison of these two approaches is presented. Two critical reasons why we use the BEM solver FastLap in this research work are described: easier meshing and better efficiency. In the second and third sections, the theory of the boundary integral equations for potential problems, and the formulation of the boundary element method based on those boundary integral equations are presented.

3.1 The Boundary Element Method (BEM) and The Finite Element Method (FEM)

Tremendous research about the finite-element method in numerical analysis of structures for engineering and science problems started in the 1950s due to the fast growth of computation capacity. Long-time consistent research contribution and testing have made the finite-element method the most widely available and reliable technique in the industry. The finite-element method has also been proved to be inadequate, inefficient, or even inaccurate in many applications. Since the 1970s, the boundary-element method began to receive attention because of its advantages over the finite-element method.

In general, the finite-element method approaches discretize the strictures into finite element meshes. In two-dimensional cases, the finite element is a planar surface consisting of straight lines that are connected by nodes on the corners of the surface. In three-dimensional cases, the finite element is a solid block or tetrahedron. The whole structure has to be meshed into finite elements even though for some cases most of the region is not of interest. The boundary-element method approach, on the other hand, only meshes the boundary of the structures, so for two-dimensional problems, the boundary elements are lines that are connected by nodes, and for three-dimensional problems, they are planar surfaces. Both methods need to solve inverse-matrix problems in order to obtain the unknowns on the nodes.

3.1.1 Advantages of The Boundary Element Method

The first advantage of boundary elements over finite elements is ease of meshing, and this advantage is one of the two critical reasons for using the boundary element solver in this research.

Since the governing equations of the boundary-element methods are integral equations containing only boundary integrals, only boundary meshes are needed for simulation. Furthermore, the boundary element method accepts discontinuous elements, which allow users to model three-dimensional structures without the requirement of matching up the nodes on the edge of two adjacent surfaces, as shown in Figure 3.1.1. These properties not only reduce the pre-process time and cost of modeling, but also give much better flexibility for problems with moving boundaries.

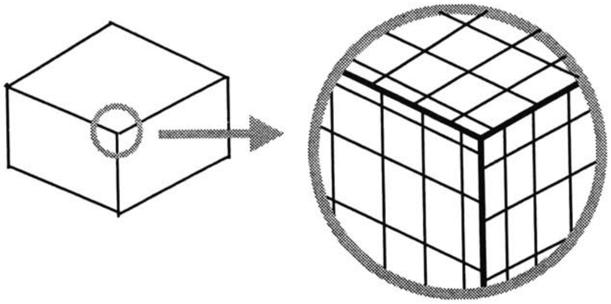


Figure 3.1.1 A boundary-element meshed block (meshes not shown) and its detailed meshes on a corner (shown in the circle). Note that the nodes on the edge of each surface do not need to be matched.

The second advantage is that adaptive error control techniques are much easier to apply to the boundary element method. The adaptive error control techniques detect the error during BEM and FEM simulations. If any unacceptable error is detected, a few schemes of mesh refinement can be performed. The boundary element method is much easier for refining meshes because of the simplicity of mesh and the acceptance of discontinuous meshes. Therefore, automatic adaptive error control techniques can be more easily integrated into the boundary element method than the finite element method.

3.1.2 Efficiency of the Boundary Element Method

A potential disadvantage of the boundary element method is poor analysis efficiency because BEM formulations generate dense matrices. In this sub-section, we will discuss why the boundary element method in general is not as efficient as the finite element method. Then we will point out that our BEM electrostatic solver FastLap, which employs the multipole acceleration technique, not only overcomes this disadvantage of the boundary element method, but also exceeds the performance of the finite-element method. That is the second critical reason for using the FastLap as our electrostatics solver in this research work.

As we will demonstrate in a later section, the boundary element method has to solve the inverse of fully populated (dense) matrices ($N_{EB} \times N_{EB}$, where N_{EB} is the number of elements). The typical cost for solving dense matrices is proportional to the square of the number of elements, i.e., N_{EB}^2. On the other hand, the finite element method only needs to solve sparse matrices ($N_{EF} \times N_{EF}$, where N_{EF}^2 is the number of total elements) that usually contain zeros in most entries of the matrices except for diagonal and off-diagonal elements, as shown in Figure 3.1.2. Because tremendous research work has been done in solving sparse matrices, the typical cost is proportional to N_{EF}. Taking a three-dimensional cubic block model as an example, the total number of elements for a boundary element model is $N_{EB} = 6 \cdot N_n^2$, where N_n is the number of elements on each edge of the cubic, and the total number of elements for a finite element model is $N_{EF} = N_n^3$. With simple substitutions, we can observe that the total cost of computation for the boundary element method is proportional to N_n^4, while the cost of the finite element method is proportional to N_n^3. Clearly, the efficiency of the boundary element method will be even worse for models with finer meshes.

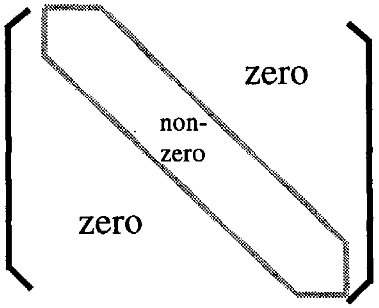


Figure 3.1.2 The sparse matrix generated by finite-element methods.

Fortunately, the boundary-element-method solver (FastLap) that is used in this study employs a technique called the multipole-accelerated method [25][68], which successfully makes the cost of computation proportional to N_{EB}, i.e., N_n^2. It significantly improves the efficiency and makes this boundary element code even better than commercial finite element solvers for the electrostatics analysis. This is the second critical reason why we use the BEM solver FastLap in our study.

The memory requirement of the boundary element method is almost the same as that of the finite element method. Although the matrices generated the boundary element methods are dense, their size is much smaller that that generated by the finite element method (i.e., $N_{EF} > N_{EB}$).

Table 3.1.1 lists the comparison of the finite element method and the boundary element method.

Category	Better Method
Preprocessing	Boundary Element Method
Computation Cost	In general: Finite Element Method
	In this work: Boundary Element Method with multi-pole acceleration technique is more efficient than Finite-Element-Method solvers
Memory requirement	Same

Table 3.1.1 Comparison of the finite element method and the boundary element method.

3.2 Boundary Integral Equations in Potential Problems

The boundary element method is a numerical technique to solve boundary integral equations (BIE). One of its important applications is to solve the potential problems, whose governing equation is the Laplace's or Poisson's equation. In this section we mainly focus on the boundary integral equation formulations for the Laplace's equation [69][70][71][72][73][74].

Equation (22) is the Laplace's equation. Figure 3.2.1 shows the schematic of a potential problem.

$$\nabla^2\varphi(\bar{x}) = 0 \qquad \bar{x} \text{ in } \Omega \tag{22}$$

where the potential φ that satisfies equation (22) in the entire domain Ω is subject to the boundary condition on the boundary S of Ω:

$$f\left(\varphi, \frac{\partial\varphi}{\partial n}, \bar{x}\right) = 0 \qquad \bar{x} \text{ on } S \tag{23}$$

where $\frac{\partial\varphi}{\partial n}$ is the normal derivative of φ on surface S.

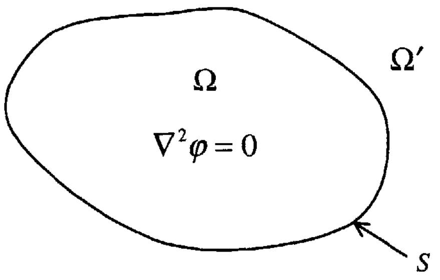


Figure 3.2.1 Schematic of a potential problem.

Then we apply Green's theorem (Green's second identity), which is the starting point for deriving boundary integral equations:

$$\int_{\Omega} u\nabla^2 w d\Omega - \int_{\Omega} w\nabla^2 u d\Omega = \int_{S} w\frac{\partial u}{\partial n}dS - \int_{S} u\frac{\partial w}{\partial n}dS \tag{24}$$

where u and w are two arbitrary scalar function that are continuous in the volume Ω. Note that this equation shows a reciprocal relationship between potential and normal flux density.

By assuming $w = \varphi$ and applying equation (22), the first term on the right hand side of equation (24) is eliminated and equation (24) can be written as:

$$\int_{\Omega} \varphi\nabla^2 u d\Omega = \int_{S} u\frac{\partial \varphi}{\partial n}dS - \int_{S} \varphi\frac{\partial u}{\partial n}dS \tag{25}$$

Now replace u with the Green's function $G(\overline{x},\overline{\xi})$, which is defined as:

$$\nabla^2 G(\overline{x},\overline{\xi}) = -\delta(\overline{x},\overline{\xi}) \qquad \overline{x}\in\Omega\,,\ \overline{\xi}\in S \tag{26}$$

where $\delta(\overline{x},\overline{\xi})$ is the Dirac delta function.

The Green's function makes the volume integral on the left-hand side of equation (25) equal to $\varphi(\overline{x})$, and thus:

$$\varphi(\overline{x}) = \int_{S} G(\overline{x},\overline{\xi})\frac{\partial \varphi}{\partial n}dS - \int_{S} \varphi\frac{\partial G(\overline{x},\overline{\xi})}{\partial n}dS \qquad \overline{x}\in\Omega\,,\ \overline{\xi}\in S \tag{27}$$

We have successfully eliminated the volume integral terms in equation (24) by applying Laplace's equation and Green's function. Equation (27) has a very important property that only one term depends on a point in the volume, which is the first term $\varphi(\overline{x})$ on the left-hand side of the equation. $G(\overline{x}, \overline{\xi})$ is also called the fundamental solution to Laplace's equation. For two-dimensional problems, $G(\overline{x}, \overline{\xi})$ is

$$G(\overline{x}, \overline{\xi}) = -\frac{1}{2\pi} \ln\left(\frac{1}{|\overline{x} - \overline{\xi}|}\right) \tag{28}$$

and for three dimentional problems, $G(\overline{x}, \overline{\xi})$ is

$$G(\overline{x}, \overline{\xi}) = -\frac{1}{4\pi|\overline{x} - \overline{\xi}|} \tag{29}$$

Since $\varphi(\overline{x})$ in equation (27) can either be in volume Ω or on boundary S, a multiplier $c(\overline{x})$ is introduced.

$$c(\overline{x})\varphi(\overline{x}) = \int_S G(\overline{x}, \overline{\xi}) \frac{\partial \varphi}{\partial n} dS - \int_S \varphi \frac{\partial G(\overline{x}, \overline{\xi})}{\partial n} dS \qquad \overline{x} \in \Omega \cup S, \ \overline{\xi} \in S \tag{30}$$

For the $\overline{x}$ inside Ω, $c(\overline{x})$ is equal to one; for $\overline{x}$ outside Ω, $c(\overline{x})$ is zero; for $\overline{x}$ on a smooth boundary, $c(\overline{x})$ is 0.5. If $\overline{x}$ is at a corner, then

$$c(\overline{x}) = \frac{\theta}{2\pi} \tag{31}$$

The definition of θ is illustrated in Figure 3.2.2.

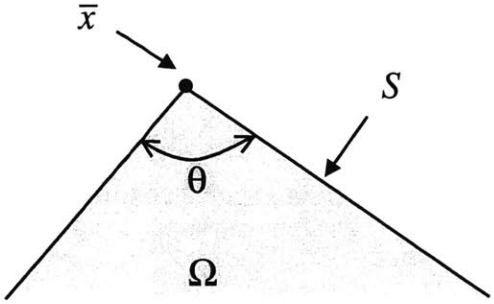


Figure 3.2.2 Definition of θ for the multiplier $c(\bar{x})$.

Equation (30) is completely in terms of boundary values, and is called a boundary integral equation for potential problems. Also, the derivation of equation (30) is generally referred to as the direct method of boundary integral equations, because the unknown functions (either $\varphi(\bar{x})$ or $\frac{\partial\varphi(\bar{x})}{\partial n}$ on S) are usually of direct interest to the person who solves Laplace's equation.

So far the interior volume bounded by the boundary has been focused upon, so equation (30) is for interior problems. For exterior problems ($\bar{x}\in\Omega'$, see Figure 3.2.1), we can follow the similar derivation described above and obtain:

$$-(1-c(\bar{x}))\varphi(\bar{x}) = \int_S G(\bar{x},\bar{\xi})\frac{\partial\varphi}{\partial n}dS - \int_S \varphi\frac{\partial G(\bar{x},\bar{\xi})}{\partial n}dS \qquad \bar{x}\in\Omega'\cup S,\ \bar{\xi}\in S \tag{32}$$

Note that the definition of $c(\bar{x})$ is the same as equation (31), so if $\bar{x}\in\Omega$, then $c(\bar{x})=0$.

Adding equations (27) and (32) results:

$$\begin{aligned}\varphi(\bar{x}) &= \int_S \sigma G(\bar{x},\bar{\xi})dS - \int_S \mu\frac{\partial G(\bar{x},\bar{\xi})}{\partial n}dS \qquad \bar{x}\in\Omega\cup\Omega' \\ c(\bar{x})\varphi^i(\bar{x}) + (1-c(\bar{x}))\varphi^e(\bar{x}) &= \int_S \sigma G(\bar{x},\bar{\xi})dS - \int_S \mu\frac{\partial G(\bar{x},\bar{\xi})}{\partial n}dS \qquad \bar{x}\in S\end{aligned} \tag{33}$$

where we define

$$\varphi^i(\bar{x}) = \lim_{\bar{x}^i\to\bar{x}}\varphi(\bar{x}^i), \qquad \bar{x}^i\in\Omega,\ \ \bar{x}\in S \tag{34}$$

$$\varphi^e(\bar{x}) = \lim_{\bar{x}^e \to \bar{x}} \varphi(\bar{x}^e), \qquad \bar{x}^e \in \Omega', \quad \bar{x} \in S \tag{35}$$

$$\mu = \varphi^i - \varphi^e \tag{36}$$

$$\sigma = \frac{\partial \varphi^i}{\partial n} - \frac{\partial \varphi^e}{\partial n} \tag{37}$$

For different types of boundary conditions, equation (33) can be simplified to different forms. Two examples are demonstrated below:

- $\varphi^i = \varphi^e = f$ on S (38)

$$\begin{aligned} \varphi(\bar{x}) &= \int_S \sigma G(\bar{x}, \bar{\xi}) dS \qquad \bar{x} \in \Omega \cup \Omega' \\ f(\bar{x}) &= \int_S \sigma G(\bar{x}, \bar{\xi}) dS \qquad \bar{x} \in S \end{aligned} \tag{39}$$

This formulation is also called the source-only formulation. The problem with this type of boundary is the single-layer potential problem.

- $\dfrac{\partial \varphi^i}{\partial n} = \dfrac{\partial \varphi^e}{\partial n}$ on S (40)

$$\begin{aligned} \varphi(\bar{x}) &= -\int_S \mu \frac{\partial G(\bar{x}, \bar{\xi})}{\partial n} dS \qquad \bar{x} \in \Omega \cup \Omega' \\ \varphi^i(\bar{x}) &= \frac{1}{2}\mu(\bar{x}) - \int_S \mu \frac{\partial G(\bar{x}, \bar{\xi})}{\partial n} dS \qquad \bar{x} \in S \end{aligned} \tag{41}$$

This formulation is also called the dipole-only formulation. The problem with this type of boundary is the double-layer potential problem.

Note that equation (39) and (41) are also referred to as indirect methods of boundary integral equations, because the unknown variables μ and σ are fictitious quantities that do not have physical meaning. However, those integral equations are easier to deal with numerically, and are favorable in some cases.

3.3 Boundary Element Method

In the previous section, the different kinds of boundary integral equations for potential problems were described, including the direct method (Green's theorem) and the indirect methods (single-layer and double-layer potential problems, or source-only and dipole-only formulations). The integral equations are derived from Green's second identity by applying the fundamental solutions of Laplace's equation [26][69].

Analytical solutions for the integral equations are only available for a few special cases. For most engineering applications, the boundary integral equations need to be integrated numerically by discretizing them into boundary elements. Equation (30) is taken as an example to demonstrate the procedure of the boundary element method formulations.

Figure 3.3.1 shows a two-dimensional boundary that is divided into boundary elements. Equation (30) can be written in discrete form:

$$c(\bar{x})\varphi(\bar{x}) = \sum_j \int_S G(\bar{x},\bar{\xi})\frac{\partial \varphi}{\partial n}dS_j - \sum_j \int_S \varphi \frac{\partial G(\bar{x},\bar{\xi})}{\partial n}dS_j \tag{42}$$

where j is the index of elements. Note that equation (42) is exactly the same as equation (30), where the whole boundary integration is just divided into a few small boundary integrals. However, it is the first step to formulate boundary element methods.

The black dots on the boundary are called nodes, which connect two adjacent boundary elements. On each node, there is a specific potential and flux density. Assuming there are N nodes on the boundary, the potential φ and flux density $\frac{\partial \varphi}{\partial n}$ can be approximated as:

$$\varphi = \gamma_1\varphi_1 + \gamma_2\varphi_2 + \ldots + \gamma_N\varphi_N \tag{43}$$

$$\frac{\partial \varphi}{\partial n} = \gamma_1\frac{\partial \varphi_1}{\partial n} + \gamma_2\frac{\partial \varphi_2}{\partial n} + \ldots + \gamma_N\frac{\partial \varphi_N}{\partial n} \tag{44}$$

where γ_i are the interpolation functions, which will be explained later. Equations (43) and (44) can be rewritten in the vector form as:

$$\varphi = \Gamma^T \Phi \tag{45}$$

$$\frac{\partial \varphi}{\partial n} = \Gamma^T \Phi_n \tag{46}$$

where $\Gamma^T = [\gamma_1, \gamma_2, \ldots, \gamma_N]$, $\Phi = [\varphi_1, \varphi_2, \ldots, \varphi_N]$ and $\Phi_n = \left[\frac{\partial \varphi_1}{\partial n}, \frac{\partial \varphi_2}{\partial n}, \ldots, \frac{\partial \varphi_N}{\partial n}\right]$.

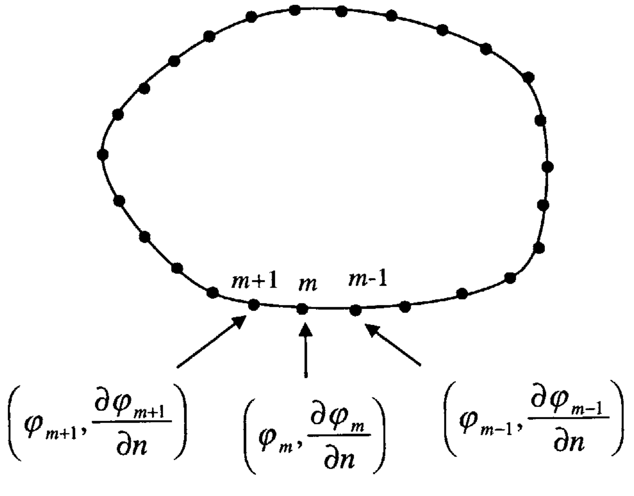


Figure 3.3.1 Two-dimensional boundary elements. m in the figure is the index of nodes.

The interpolation functions γ_i are in local coordinates. Figure 3.3.2 shows a two-dimensional (line) element with two nodes. On the element, the interpolation function of one node returns one for the point at the node, and returns zero for the point at the node on the other end of the element. For linear elements, the definitions of the interpolation functions in term of the local coordinate system (ς) are:

$$\begin{aligned} \gamma_k &= \frac{1}{2}(1+\varsigma) \\ \gamma_{k+1} &= \frac{1}{2}(1-\varsigma) \end{aligned} \tag{47}$$

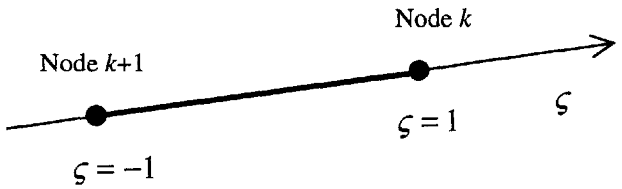


Figure 3.3.2 Local coordinate system for a two-dimensional (line) linear boundary element.

For quadratic elements (three nodes on the element), the definitions of the interpolation functions in term of the local coordinate system are:

$$\begin{aligned} \gamma_k &= \frac{1}{2}\varsigma(1+\varsigma) \\ \gamma_{k+1} &= \frac{1}{2}(1-\varsigma)(1+\varsigma) \\ \gamma_{k+2} &= \frac{1}{2}\varsigma(1-\varsigma) \end{aligned} \tag{48}$$

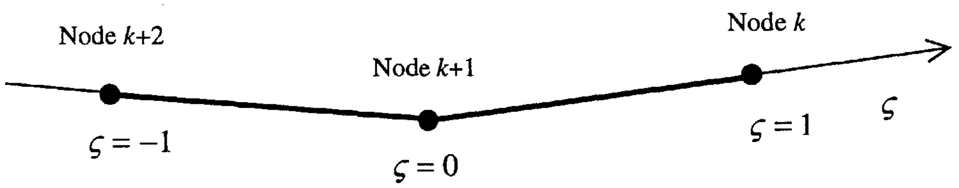


Figure 3.3.3 Local coordinate system for a two-dimensional (line) quadratic boundary element.

Then equation (42) can be rewritten as:

$$c(\bar{x})\Phi = \sum_{j}\int_{S} G(\bar{x},\bar{\xi})\Gamma^T dS_j \Phi_n - \sum_{j}\int_{S} \frac{\partial G(\bar{x},\bar{\xi})}{\partial n}\Gamma^T dS_j \Phi \tag{49}$$

where Φ and Φ_n are constant vectors that represent the values of potential and flux density at the nodes along the boundary. They are placed outside the integrals because they are independent of the integral. Also, the potential on the left hand side is written as the vector representation due to this discretization. Note that $G(\bar{x},\bar{\xi})$ and $\frac{\partial G(\bar{x},\bar{\xi})}{\partial n}$ are known functions, and Γ^T is a vector that contains local interpolation functions for all nodes. Therefore, the integrals for each boundary element can be evaluated numerically.

After evaluating the integrals, equation (49) can be rewritten as:

$$\mathbf{H}\mathbf{\Phi} = \mathbf{G}\mathbf{\Phi}_n \tag{50}$$

where $\mathbf{H}$ and $\mathbf{G}$ are $n \times n$ matrices and called the *influence matrices.*

As mentioned in the previous paragraph, $\mathbf{\Phi}$ and $\mathbf{\Phi}_n$ are constant vectors that represent the values of potential and flux density on each node respectively. However, the components in $\mathbf{\Phi}$ and $\mathbf{\Phi}_n$ are either given or unknown values depending on the boundary conditions imposed on the problem. For a problem that is solvable, it is necessary to have known values of either potential or flux density at each node.

For example, considering a Dirichlet boundary problem, $\mathbf{\Phi}$ is known on the whole boundary and equation (50) can be rearranged in a conventional way such that the known quantities are on the right-hand side and unknown quantities are on the left-hand side:

$$\mathbf{G}\mathbf{\Phi}_n = \mathbf{Z} \tag{51}$$

where $\mathbf{Z}$ is a known vector and evaluated from $\mathbf{H}\mathbf{\Phi}$. Thus $\mathbf{\Phi}_n$ is ready to be solved by linear algebra numerical techniques.

Similarly, for a Neumann boundary problem, equation (50) can be solved immediately.

For a mixed-type boundary problem that has Dirichlet and Neumann conditions imposed on the boundaries, showing in Figure 3.3.4, we can rearrange the columns in $\mathbf{H}$ and $\mathbf{G}$ as well as the components in $\mathbf{\Phi}$ and $\mathbf{\Phi}_n$ so that all the unknown variables are on the left-hand side, and solve for the unknown quantities. In Figure 3.3.4, the dashed line is the Neumann type boundary and the nodes 1,2,6 and 7 have given flux densities; the solid line is the Dirichlet type boundary and the nodes 3,4 and 5 have given potentials.

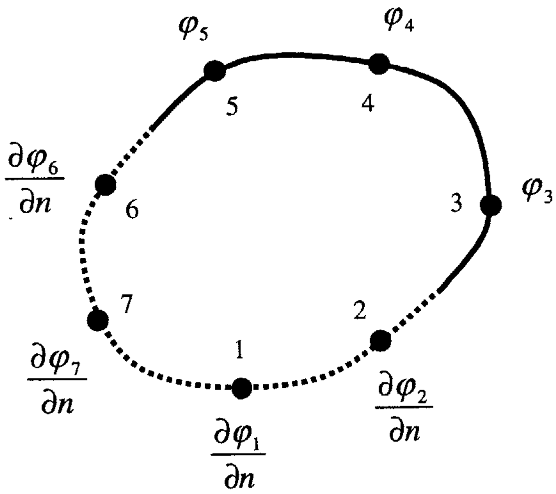


Figure 3.3.4 A mixed-type boundary problem. The solid line is Dirichlet type boundary, and the dashed line is the Neumann type boundary. The given quantities are shown beside the nodes.

Equation (52)is the equation of the system with detailed components.

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} & h_{16} & h_{17} \\ h_{21} & h_{22} & h_{23} & h_{24} & h_{25} & h_{26} & h_{27} \\ h_{31} & h_{32} & h_{33} & h_{34} & h_{35} & h_{36} & h_{37} \\ h_{41} & h_{42} & h_{43} & h_{44} & h_{45} & h_{46} & h_{47} \\ h_{51} & h_{52} & h_{53} & h_{54} & h_{55} & h_{56} & h_{57} \\ h_{61} & h_{62} & h_{63} & h_{64} & h_{65} & h_{66} & h_{67} \\ h_{71} & h_{72} & h_{73} & h_{74} & h_{75} & h_{76} & h_{77} \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \\ \varphi_5 \\ \varphi_6 \\ \varphi_7 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} & g_{15} & g_{16} & g_{17} \\ g_{21} & g_{22} & g_{23} & g_{24} & g_{25} & g_{26} & g_{27} \\ g_{31} & g_{32} & g_{33} & g_{34} & g_{35} & g_{36} & g_{37} \\ g_{41} & g_{42} & g_{43} & g_{44} & g_{45} & g_{46} & g_{47} \\ g_{51} & g_{52} & g_{53} & g_{54} & g_{55} & g_{56} & g_{57} \\ g_{61} & g_{62} & g_{63} & g_{64} & g_{65} & g_{66} & g_{67} \\ g_{71} & g_{72} & g_{73} & g_{74} & g_{75} & g_{76} & g_{77} \end{bmatrix} \begin{bmatrix} \partial \varphi_1/\partial n \\ \partial \varphi_2/\partial n \\ \partial \varphi_3/\partial n \\ \partial \varphi_4/\partial n \\ \partial \varphi_5/\partial n \\ \partial \varphi_6/\partial n \\ \partial \varphi_7/\partial n \end{bmatrix} \quad (52)$$

Since φ_3, φ_4 and φ_5 are given, we move those components as well as their corresponding columns in **H** to the right hand side, and move the unknown components $\frac{\partial \varphi_3}{\partial n}$, $\frac{\partial \varphi_4}{\partial n}$ and $\frac{\partial \varphi_5}{\partial n}$ as well as their corresponding columns in **G** to the left hand side. Thus, we obtain a new equation of system:

$$\begin{bmatrix} h_{11} & h_{12} & g_{13} & g_{14} & g_{15} & h_{16} & h_{17} \\ h_{21} & h_{22} & g_{23} & g_{24} & g_{25} & h_{26} & h_{27} \\ h_{31} & h_{32} & g_{33} & g_{34} & g_{35} & h_{36} & h_{37} \\ h_{41} & h_{42} & g_{43} & g_{44} & g_{45} & h_{46} & h_{47} \\ h_{51} & h_{52} & g_{53} & g_{54} & g_{55} & h_{56} & h_{57} \\ h_{61} & h_{62} & g_{63} & g_{64} & g_{65} & h_{66} & h_{67} \\ h_{71} & h_{72} & g_{73} & g_{74} & g_{75} & h_{76} & h_{77} \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \partial\varphi_3/\partial n \\ \partial\varphi_4/\partial n \\ \partial\varphi_5/\partial n \\ \varphi_6 \\ \varphi_7 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & h_{13} & h_{14} & h_{15} & g_{16} & g_{17} \\ g_{21} & g_{22} & h_{23} & h_{24} & h_{25} & g_{26} & g_{27} \\ g_{31} & g_{32} & h_{33} & h_{34} & h_{35} & g_{36} & g_{37} \\ g_{41} & g_{42} & h_{43} & h_{44} & h_{45} & g_{46} & g_{47} \\ g_{51} & g_{52} & h_{53} & h_{54} & h_{55} & g_{56} & g_{57} \\ g_{61} & g_{62} & h_{63} & h_{64} & h_{65} & g_{66} & g_{67} \\ g_{71} & g_{72} & h_{73} & h_{74} & h_{75} & g_{76} & g_{77} \end{bmatrix} \begin{bmatrix} \partial\varphi_1/\partial n \\ \partial\varphi_2/\partial n \\ \varphi_3 \\ \varphi_4 \\ \varphi_5 \\ \partial\varphi_6/\partial n \\ \partial\varphi_7/\partial n \end{bmatrix} \tag{53}$$

Obviously, the unknown quantities on the left-hand side of the equation (53) can be solved readily.

In the next chapter, the detailed procedure to solve those integral equations and obtain the electric field data essential for the trajectory simulation will be presented.

CHAPTER 4

MODELING PROCEDURE AND SIMULATOIN RESULTS FOR FEAs

In this chapter, the modeling procedure of field emission arrays is discussed based on the field-emission theory and the boundary element method described in previous chapters. Then the simulation results for various types of field emission devices will be presented, including simple FEAs, proximity-focused FEAs and *integrated-focus-electrode* (IFE) FEAs.

4.1 Simulation Procedure

Figure 4.1.1 shows a typical *computer-aided design* (CAD) system for the design of field emission devices [75]. The design parameters include geometry data for each component such as the cone, the gate, the focus, the gate aperture, the focus aperture, the tip radius of curvature, and the pitch distance between each emitters. Also, the operating parameters should be considered in the design phase. The operating parameters are the voltages applied on each component, such as the gate voltage, the anode voltage, and the IFE voltage.

With these design parameters, a three-dimensional solid model of the device is created. After meshing the device model and applying appropriate boundary conditions on all components of the device and boundaries, a BEM solver is used to simulate the electrostatics, and the electric field distribution in the space of interest is obtained. The current density as a function of position on the tip surface is calculated using the Fowler-Nordheim electron emission theory. The trajectories of emitted electrons can be calculated by the simulated electric field data. Then based on the electron trajectories, the distribution of the final positions of emitted electrons on the anode can be found and the spot size and emission current can be obtained.

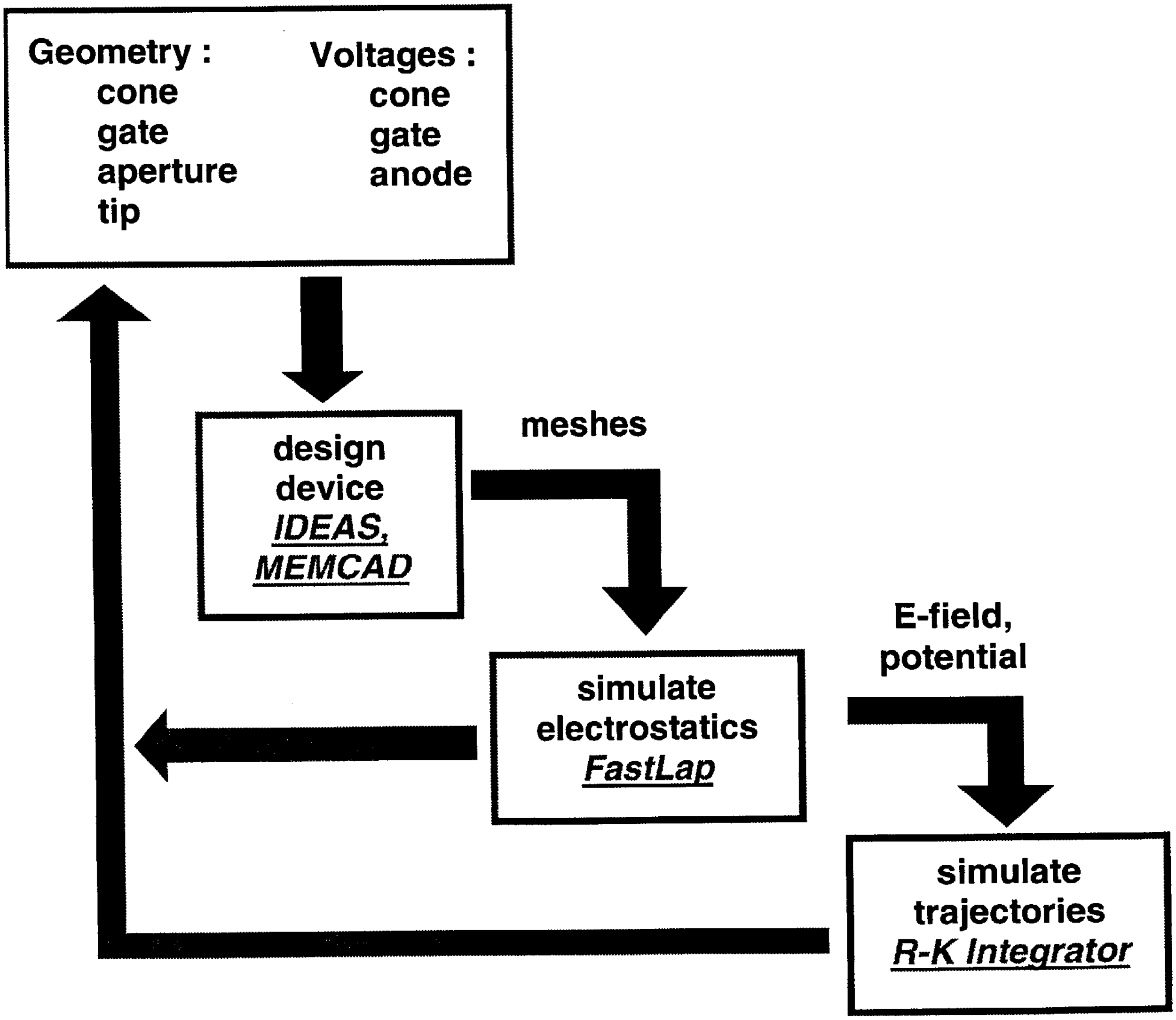


Figure 4.1.1 Procedure of a CAD system for design of field-emission devices.

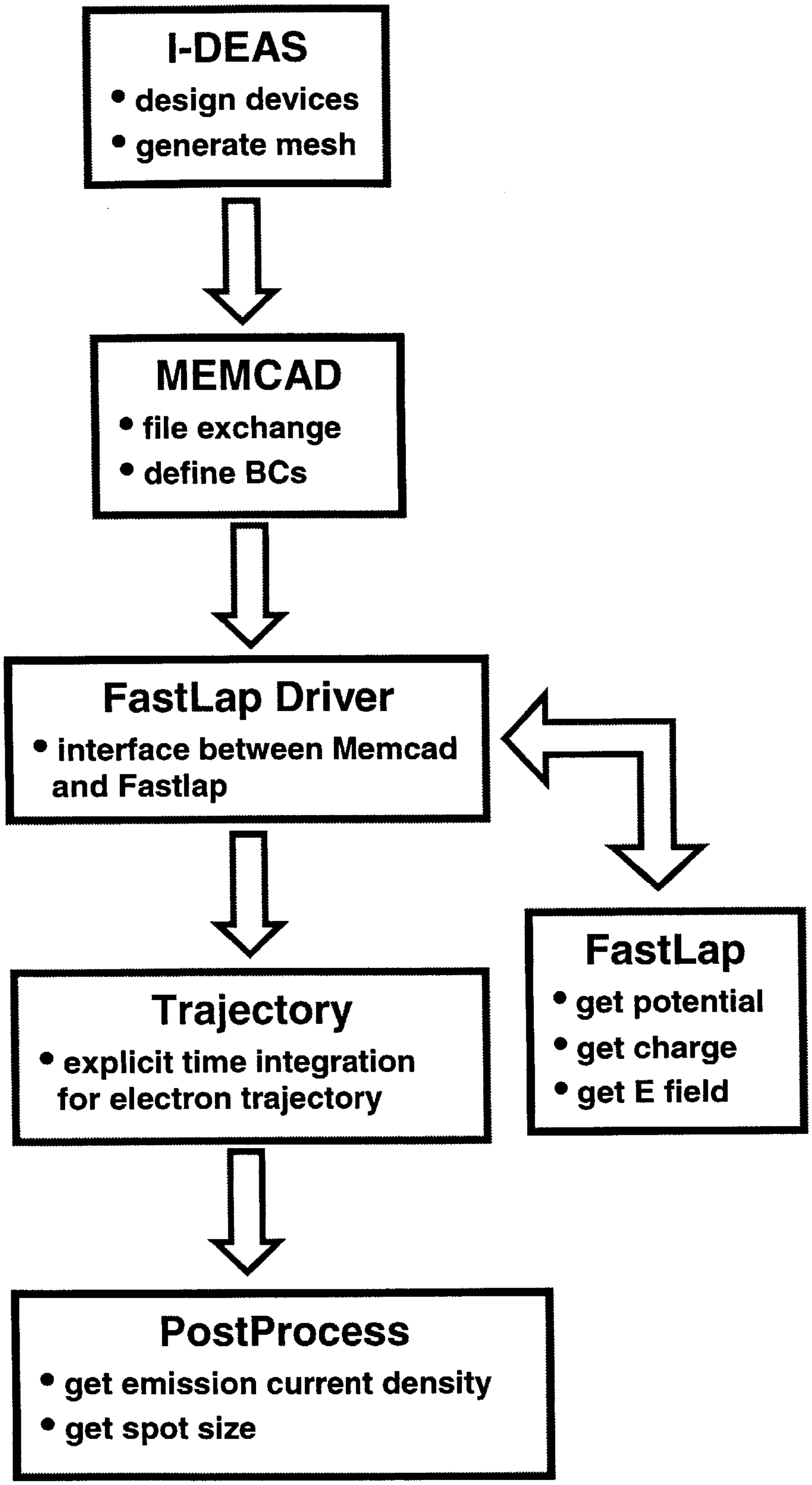


Figure 4.1.2 Block diagram of the process for modeling field emission devices in this study.

The block diagram of the CAD system used in this study is shown in Figure 4.1.2. A solid modeling package IDEAS [76] is used for creating solid models as well as boundary meshes for field emission devices. Note that since the ratio between the mesh length scale on the tip and that on other components of the device is about $\frac{1}{1000}$, well-controlled map meshes are recommended on the cone structures. The numerical model is then exported into MEMCAD, which is a MEMS modeling CAD package. The boundary conditions are applied to appropriate components and surfaces using MEMCAD.

FastLap is a group of C-language libraries that can be called by any other Fortran or C-language programs. The FastLap driver, which is written in C and C++, is the interface program between MEMCAD and FastLap. The FastLap driver reads the database files generated by MEMCAD, and then transfers the data into panel data that can be used by the FastLap. Also, the FastLap driver controls the electrostatics simulation sequences that are needed to obtain the electric field results for trajectory simulations.

The surface-electric-field distribution on the tip surface is used to calculate the emission current density distribution around the tip by the Fowler-Nordheim equation. A C-language-based trajectory simulator, which employs the Runge-Kutta explicit integration method, is used to calculate the trajectories of the electrons emitted from the panels on the emitter tip using the electric-field distribution in space. Finally, the spot size and emission current (density) are extracted from the trajectory simulation results by MATLAB scripts.

4.2 Electrostatics Simulations

4.2.1 Electrostatic Models

In CHAPTER 3, a concise description of the boundary integral equations in potential theory, and an example of formulation of the boundary element method are provided. In this section, the steps for solving a so-called mixed-type boundary problem using those boundary integral equation formulations are discussed. Figure 4.2.1 is a schematic of a field emission device.

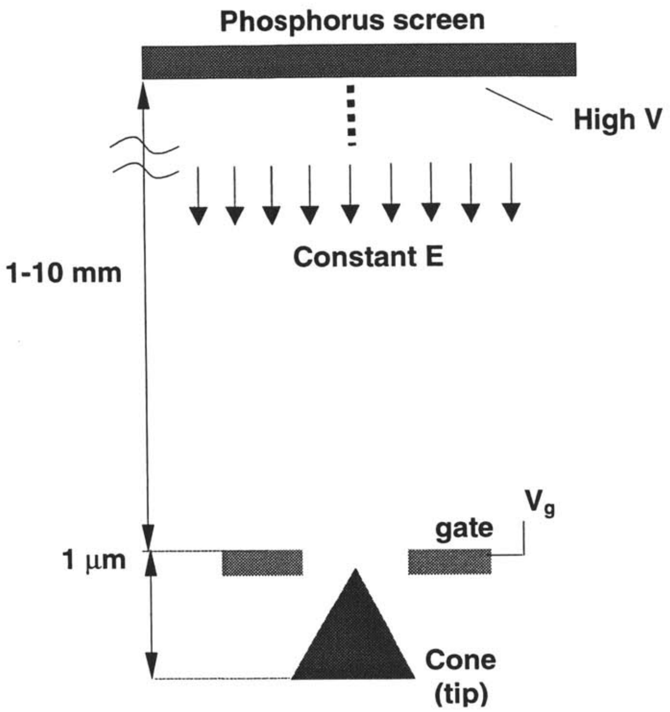


Figure 4.2.1 Schematic of a field emission device.

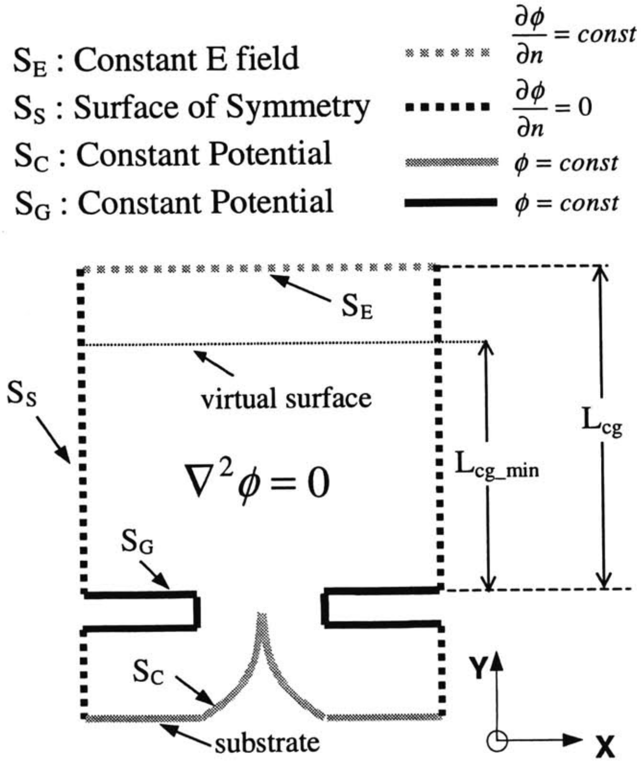


Figure 4.2.2 Boundary conditions of the model for the field emission device shown in Figure 4.2.1.

Note that in Figure 4.2.1 the size of the cone and gate structure is on the order of 1 micron, while the distance between the gate and the anode (phosphor screen) is on the order of 1mm (a difference of scale of more than 1000). The boundary element model of the device shown in Figure 4.2.1 is illustrated in Figure 4.2.2. The solid lines indicate the surfaces that have the Dirichlet-type boundary conditions, which contain constant potentials. The dotted lines indicate the surfaces that have Neumann-type boundary conditions, which contain constant values of normal derivatives of potentials.

Since this model represents an emitter located in an emitter array, the boundary conditions of symmetry are applied on the sides of the simulated domain. These symmetric boundary conditions force the electric field perpendicular to the surface ($\frac{\partial \varphi}{\partial n}$) to equal to zero. In other word, there are no electric field components that are normal to the side-walls of the simulated domain.

Due to the large scale difference between the cone-gate structure and the gate-anode distance, another Neumann-type boundary condition is applied on the top surface to represent the constant electric field caused by the voltage drop between the gate and the anode. As a result, only a model that has the same length scale of the cone-gate structure is simulated. Note that there is a minimum required distance between the top surface and the gate because the electric field distribution is not uniform due to the gate aperture and the cone structure. However, beyond a certain minimum distance above the gate, the electric field distribution becomes uniform and equals the ratio of the anode-gate voltage-difference to the anode-gate distance. The minimum distance can be found by requiring that the field distribution inside the computational domain be the same for the two cases of L_{cg_min} and L_{cg_min}+dL, where dL is a arbitrary positive number, and L_{cg_min} is the minimum required distance between the gate and the top wall.

FastLap is a boundary-element-method Laplace's equation solver. As described in the previous chapter, there are a few different variations of boundary integral equations derived from the Laplace equation. FastLap has the capability to simulate these boundary integral equations, depending on given boundary conditions and physical quantities of interest (i.e., charge distribution, potential or electric field).

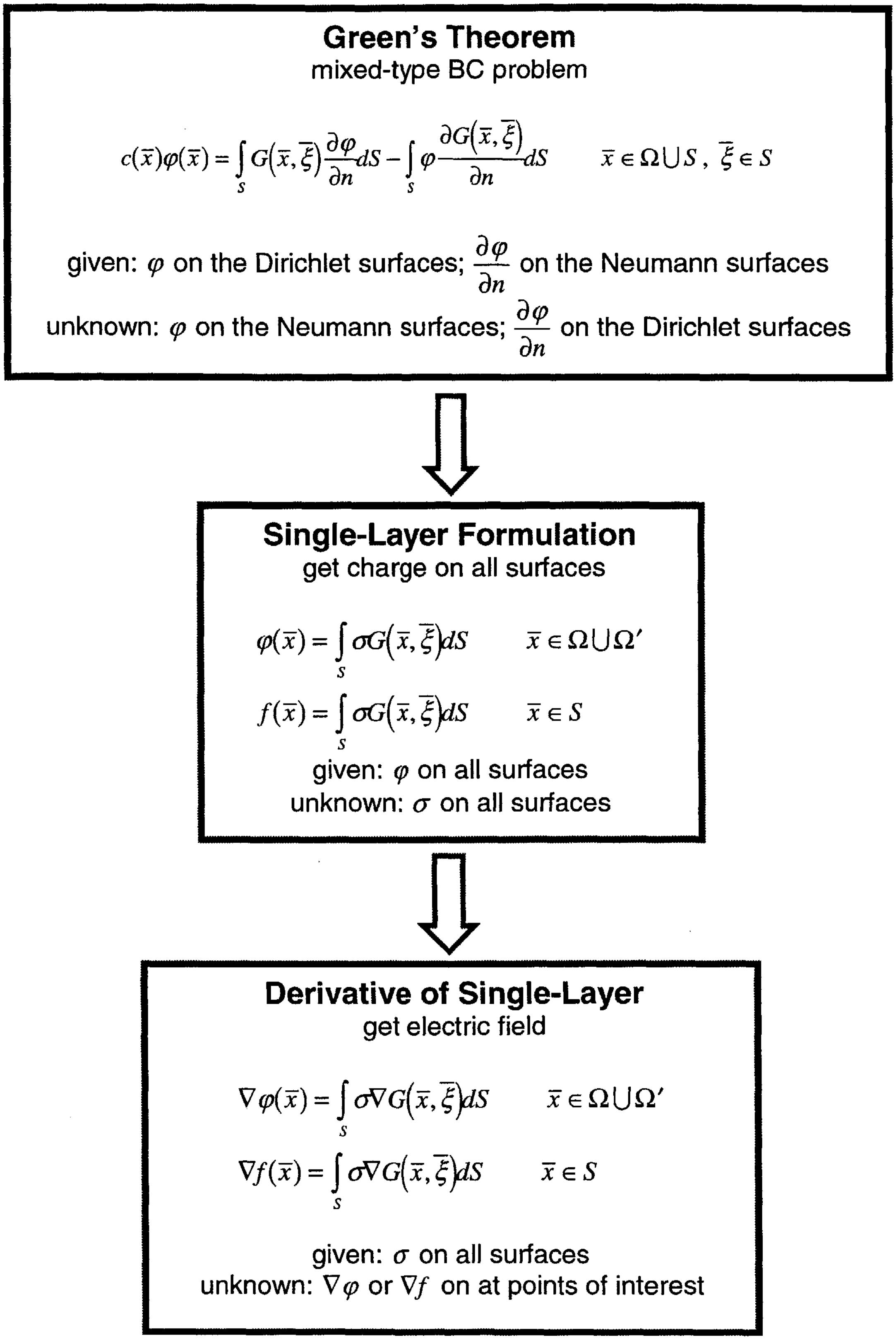
The numerical model shown in Figure 4.2.2 indicates that both the Dirichlet and the Neumann boundary conditions are applied to the model, so the boundary integral equation of the Green's Theorem formulation, equation (30), is needed for the first phase of the calculation.

The option of the Green's theorem formulation in FastLap solves the unknown potential on the Neumann surfaces, and solves the unknown flux density on the Dirichlet surfaces. After this step, complete information of the potential on all boundaries is obtained. Using this potential information, the indirect formulation of boundary integral equation is applied, which is the single layer formulation, equation (39). The indirect formulation calculates a tentative quantity called charge density. This quantity does not have any physical meaning; it is the difference of flux densities $\frac{\partial \varphi}{\partial n}$ on the boundary between interior and exterior problems, as described in CHAPTER 3. With the simulated charge density, the formulation of the derivative of the single layer is applied to calculate the electric field distribution on the boundary and in the space of interest.

Note that since FastLap can only calculate the electric field strength along the unit vector specified by users for each point of interest, FastLap is required to be called three times to obtain all components of electric field. FastLap calculates the three components of the electric field (F_x, F_y, F_z) at the point specified, by giving it the unit vector along the axes of the Cartesian coordinate system (i.e., $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\0\\1\end{bmatrix}$).

Under normal operating conditions of field-emission displays, the maximum emission current density around the tip is not high enough to cause the space charge effect [9][77]. Therefore, the electrostatic simulations in this study do not account for the space charge effect.

Figure 4.2.3 shows the steps which FastLap electrostatic simulation uses to obtain the electric field distribution. The given quantities and unknown quantities for each step are listed.



Green's Theorem

mixed-type BC problem

$$c(\bar{x})\varphi(\bar{x}) = \int_S G(\bar{x},\bar{\xi})\frac{\partial \varphi}{\partial n}dS - \int_S \varphi \frac{\partial G(\bar{x},\bar{\xi})}{\partial n}dS \qquad \bar{x} \in \Omega \cup S\,,\ \bar{\xi} \in S$$

given: φ on the Dirichlet surfaces; $\frac{\partial \varphi}{\partial n}$ on the Neumann surfaces

unknown: φ on the Neumann surfaces; $\frac{\partial \varphi}{\partial n}$ on the Dirichlet surfaces

Single-Layer Formulation

get charge on all surfaces

$$\varphi(\bar{x}) = \int_S \sigma G(\bar{x},\bar{\xi})dS \qquad \bar{x} \in \Omega \cup \Omega'$$

$$f(\bar{x}) = \int_S \sigma G(\bar{x},\bar{\xi})dS \qquad \bar{x} \in S$$

given: φ on all surfaces

unknown: σ on all surfaces

Derivative of Single-Layer

get electric field

$$\nabla\varphi(\bar{x}) = \int_S \sigma\nabla G(\bar{x},\bar{\xi})dS \qquad \bar{x} \in \Omega \cup \Omega'$$

$$\nabla f(\bar{x}) = \int_S \sigma\nabla G(\bar{x},\bar{\xi})dS \qquad \bar{x} \in S$$

given: σ on all surfaces

unknown: $\nabla\varphi$ or ∇f on at points of interest

Figure 4.2.3 Steps of electrostatics simulation to obtain electric field distribution using FastLap.

4.2.2 Accuracy Study of FastLap for FED simulations

In this sub-section, the results of the accuracy study of FastLap are presented, and the appropriate simulation parameters for accurate FED electrostatic simulations are proposed.

There are two important simulation parameters in FastLap: the order of multipole expansions (PN_{order}) and the depth (level) of decomposing computational domain (PN_{level}) [78][79][80]. PN_{order} controls the accuracy with which singularity influences are computed when they are approximated by multipole expansions. PN_{level} affects computational time because a smaller value of PN_{level} means more of the problem is done directly. In order to determine the appropriate values of these two parameters, a numerical model that has an analytical solution is constructed. This model is similar to the field-emission device model in terms of scale difference and boundary conditions.

Figure 4.2.4 shows the numerical model used for the accuracy study. The sphere is located at the center of the cube. The radius of the sphere is 50, and the side length of the cube is 2000. A constant potential of 100 is applied on the top surface of the cube, and zero potential is applied on the bottom surfaces (both of them are Dirichlet boundaries). Symmetric type boundary conditions (Neumann boundaries with $\frac{\partial \varphi}{\partial n} = 0$) are applied on the four side surfaces. The model represents an infinite two-dimensional array of spheres between two infinite plates. Since the radius of the sphere is much less than the size of the cube, the electrostatics solution in the cube is essentially equal to the solution of *a sphere in a uniform electric field*, whose analytical solutions of potential and electric field components in spherical coordinates are:

$$\varphi = \left(r - \frac{R^3}{r^2} \right) \cdot E_0 \cdot \cos(\theta) \tag{54}$$

$$E_r = -\left(1 + \frac{2 \cdot R^3}{r^3} \right) \cdot E_0 \cdot \cos(\theta) \tag{55}$$

$$E_\phi = 0 \tag{56}$$

$$E_\theta = \left(1 - \frac{R^3}{r^3}\right) \cdot E_0 \cdot \sin(\theta) \tag{57}$$

where R is the radius of the sphere, E_0 is the uniform field, and E_r, E_ϕ and E_θ are the electric field components in spherical coordinates.

The choice of PN_{level} is based on the experimentation. Small values of PN_{level} increase computational time because it is necessary to invert larger sub-matrices. However, small values of PN_{level} reduce the number of iterations. In this study, PN_{level} is fixed at 6 based on computation efficiency and convergence condition.

Figure 4.2.6, Figure 4.2.7, Figure 4.2.8 and Figure 4.2.9 are contour plots of relative error of the electric field component E_r on and outside the sphere for different values of PN_{order}. The blank region in the center of each plot is the sphere itself. The relative error at (r,ϕ,θ) is defined as:

$$Error(r,\phi,\theta) = \frac{\left|E_{simulated}(r,\phi,\theta) - E_{analytical}(r,\phi,\theta)\right|}{\left|E_{analytical}(r,\phi,\theta)\right|} \tag{58}$$

where $E_{analytical}(r,\phi,\theta)$ and $E_{simulated}(r,\phi,\theta)$ are the analytical solution and the simulated solution of a electric field component at (r,ϕ,θ), respectively.

The contours are plotted on a slice that is on a plane extended by the X and Y axes (the centroid of the sphere is at the origin of the coordinate system). The uniform field E_0 is in the Y direction. When PN_{order} is equal to 3 and 4, the error distributions are not well controlled. When PN_{order} is equal to 5 and 6, the maximum error is less than 3 %. This maximum error occurs around the region where the solution of the electric field is almost zero. Obviously, this maximum error is mainly due to the small value of $E_{analytical}$ in the denominator. On the other hand, the error is quite small (less than 1 %) at the region where E_r has a maximum value ($\theta = 0$, $r = R$). This phenomenon is encouraging for field emission simulation because the largest contribution to the emission current is from the region where maximum surface electric field occurs. Also, the error outside the sphere decays very fast to less than 0.2 %. For the electrostatic simulations in this work, PN_{order} is fixed at 5.

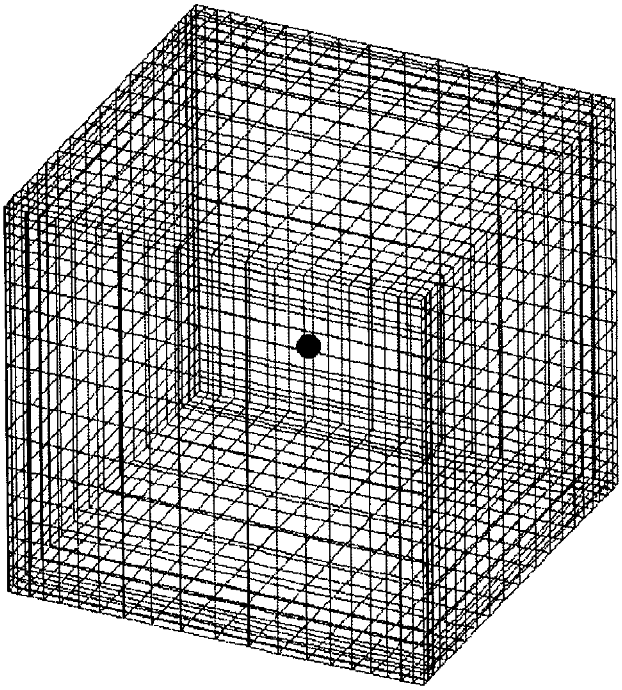


Figure 4.2.4 The numerical model of a sphere between two parallel plates.

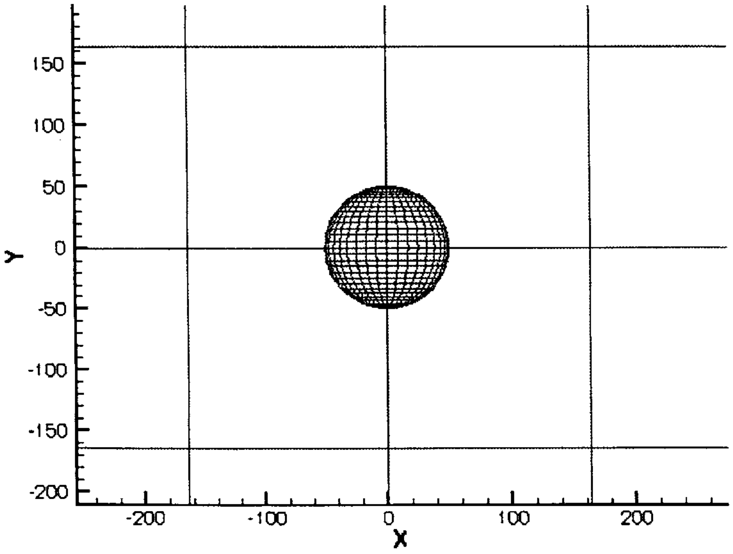


Figure 4.2.5 A closer view of the sphere shown in Figure 4.2.4.

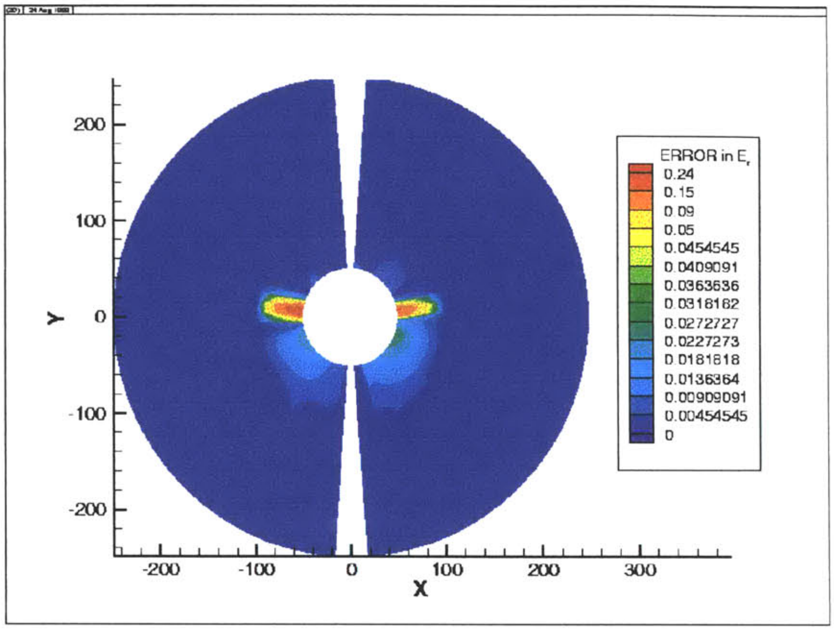


Figure 4.2.6 Contour plot of the relative error of the electric field component E_r around the sphere for PN_{level}=6 and PN_{order}=3.

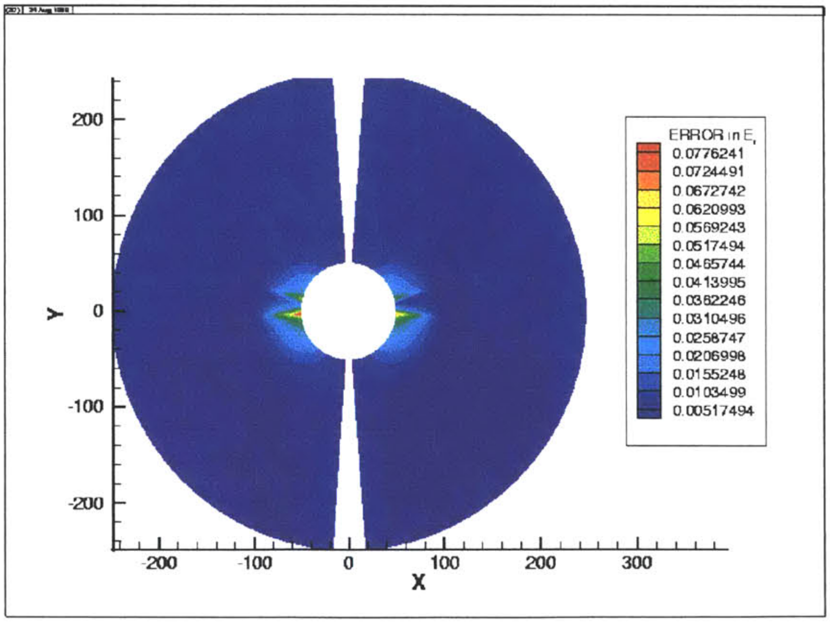


Figure 4.2.7 Contour plot of the relative error of the electric field component E_r around the sphere for PN_{level}=6 and PN_{order}=4.

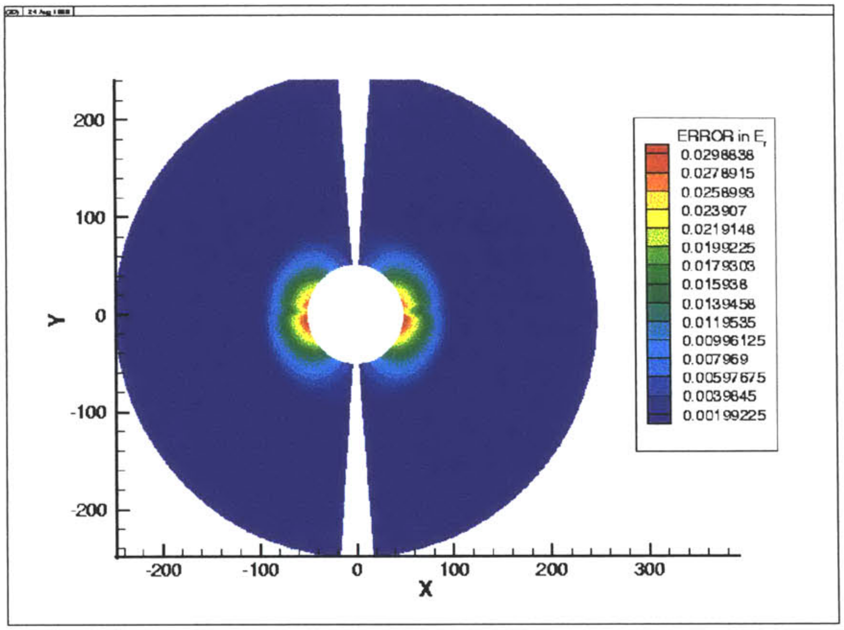


Figure 4.2.8 Contour plot of the relative error of the electric field component E_r around the sphere for PN_{level}=6 and PN_{order}=5.

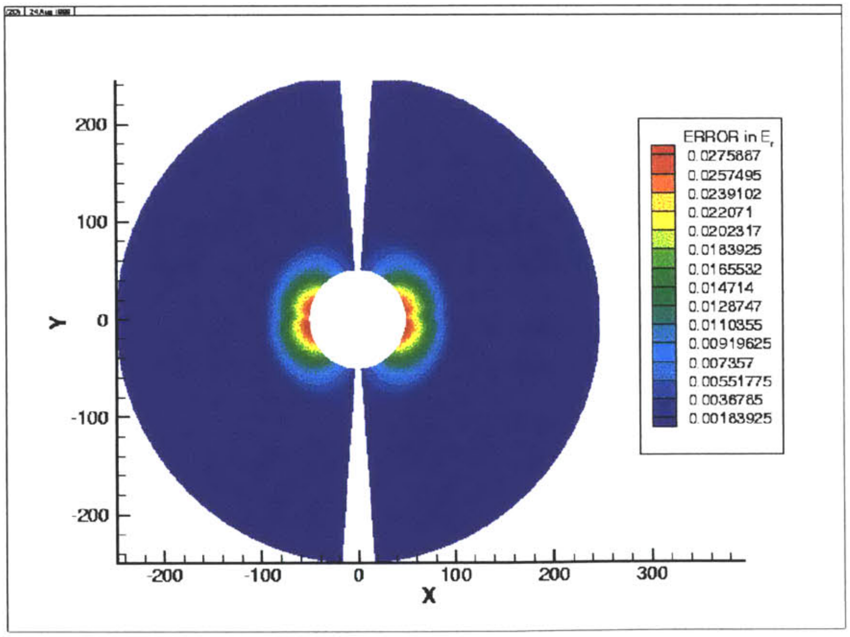


Figure 4.2.9 Contour plot of the relative error of the electric field component E_r around the sphere for PN_{level}=6 and PN_{order}=6.

4.3 Emission Current Calculation and Trajectory Simulation

Given the electric field distribution on the tip surface, the emission current density distribution on the tip is calculated by the Fowler-Nordheim equation, which has been described in CHAPTER 2. Since the tip is discretized into boundary element panels, the current contribution from each tip panel is the product of the current density on the panel and the area of the panel. The total emission current is the summation of the current contribution from each tip panel. Note that the boundary element model for the tip is a semi-sphere in this study. The maximum electric field occurs at the apex of the tip due to field enhancement. Since the Fowler-Nordheim equation indicates that emission current density is exponentially dependent on the electric field on the tip surface, the current density contributions from the panels that are close to the perimeter of the tip semi-sphere are very small compared with those from the panels that are close to the tip apex. Therefore, it is not necessary to include the current contribution from some panels on the cone structure even though they are close to the perimeter of the sphere of the tip.

The electron equation of motion is given by:

$$m\ddot{x}_i = -qF_i \qquad i = 1,2,3 \tag{59}$$

where m is the electron mass, and $\ddot{x}_i$ and F_i are the electron acceleration and the electric field at electron's current position. The i indicates the three directions (i= 1,2 and 3) in a Cartesian coordinate system.

The trajectory of an electron emitted from the centroid of a boundary element panel on the tip is calculated using a Runge-Kutta method (4th order) with an adaptive step-control scheme[81]. Since there are three equations of motion (one in each direction) that are second-order ordinary differential equations, each one can be reformulated as two first-order ordinary differential equations, and thus there are a total of six first-order ordinary differential equations to be integrated simultaneously during trajectory simulations.

In order to have better simulation efficiency, the electric fields are calculated at certain sets of points that span into the space of interest, before the trajectory simulation is performed. In the trajectory simulation, the electric field at the point of interest during integration is interpolated using the sets of electric field data.

The electric field data are grouped into two sets. The first one is called the local electric-field-point set and is used for local trajectory simulation. The definition of the local trajectory is the portion of a trajectory that is within a distance of about 10 times of the tip radius of curvature from the tip surface. For the electrons that are just emitted from the tip and are deflected back to the tip due to a strong focusing effect, the trajectory integration uses the local electric-field set. The reason for using the local set is because FastLap will give inaccurate electric field results if the electric-field points are:

- close to a panel (within a distance of about two times the diagonal of the panel), and
- NOT located on the line that is perpendicular to the centroid of *the* panel.

The proposed configurations for the tip boundary element mesh and the local electric-field-point set are shown in Figure 4.3.1, Figure 4.3.2 and Figure 4.3.3. The second set of points is the global electric-field-point set. The electric field data evaluated on the points in this set are used for the global trajectory simulation. The definition of the global trajectory is the portion of a trajectory that is outside the range defined for the local trajectory. The majority of the points in the computation domain are in the global electric-field set. Figure 4.3.4 is the two dimensional schematic that illustrates the local and global set of electric-field points.

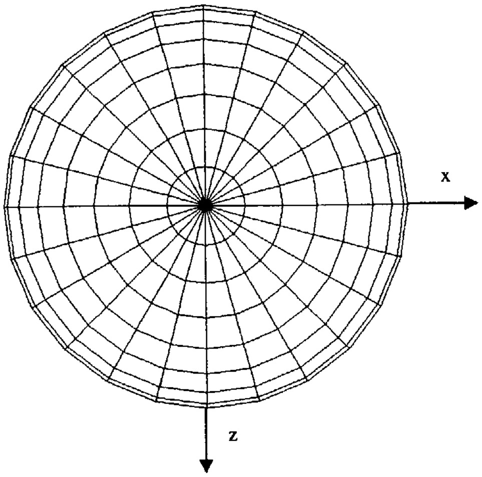


Figure 4.3.1 Top view of the boundary element mesh for the sold model of a tip that is a semi-sphere. Note that the hemisphere is symmetric to the Y-axis, which is pointing out of the paper.

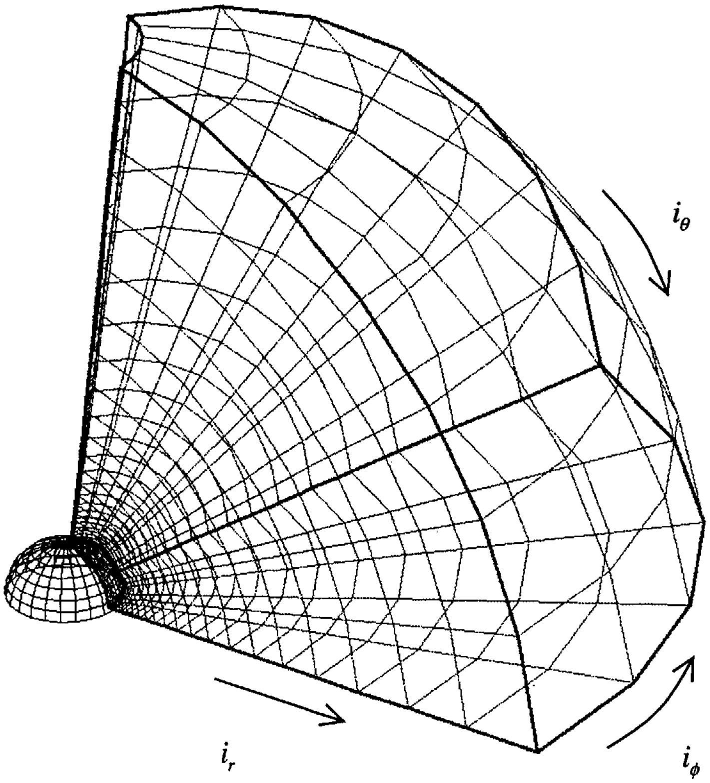


Figure 4.3.2 Configuration of local electric-field points. Note that only one quarter of the electric-field points are in the figure. Each electric-field point is located at the intersection of lines in the three-dimensional grid extending from the tip mesh.

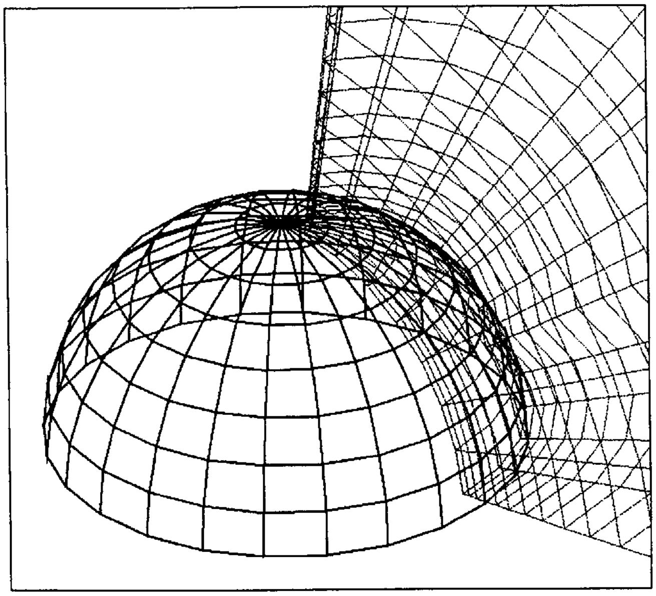


Figure 4.3.3 Close-up view of the tip mesh and part of the grid formed by local electric-field points.

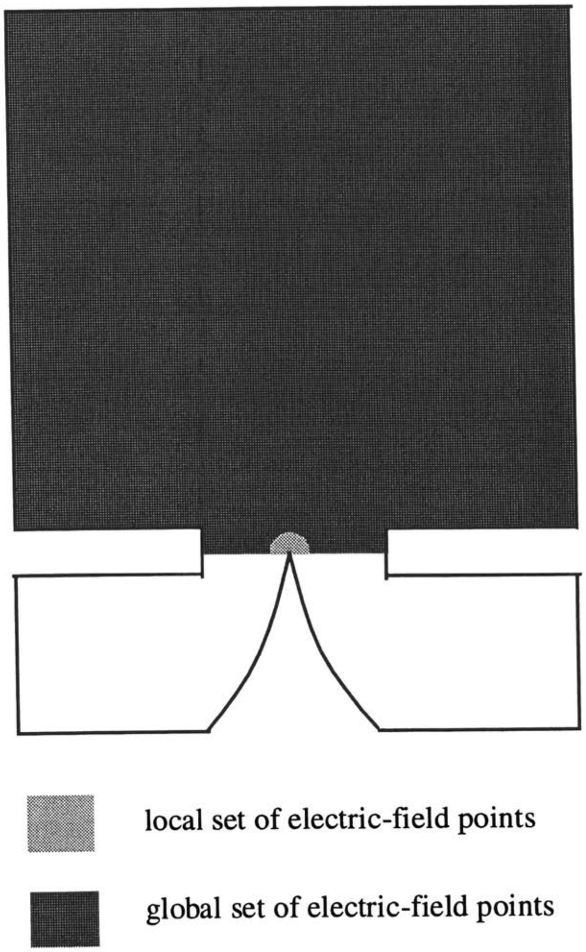


Figure 4.3.4 Schematic of the global and local sets of electric-field points in the simulation model.

The boundary element mesh for the tip shown in Figure 4.3.1 has the compatibility of indexing each panel in a spherical coordinate system. The local electric-field points are along the lines perpendicular to and projecting from the centroid of each tip panel, as shown in Figure 4.4.3. Since the electric-field points are located on the lines projected from the centroids of panels, FastLap will give accurate electric field results on the points that are close to the tip panels. The electric-field points can be addressed by the indices in a spherical coordinate system. With this property, the simulated electric-field results are stored in a three dimensional array $FARRAY_{LOCAL}\left(i_r, i_\phi, i_\theta\right)$ with three indices of $\left(i_r, i_\phi, i_\theta\right)$. The first step in the trajectory simulation is to transfer the electric field components from the Cartesian coordinate system into the Spherical coordinate system. During electron trajectory integration, the positions of electrons, which are emitted from the centroids of tip panels, are transferred into the spherical coordinate system

$(r,\phi,\theta,)$. The proper indices (i_r, i_ϕ, i_θ) in the spherical coordinate system are then evaluated based on the electron position in the spherical coordinate system, and the electric field components (F_r, F_ϕ and F_θ) at that position are calculated using linear interpolation in the spherical coordinate system. And finally, the electric field components are then transferred from the spherical coordinate system into the Cartesian coordinate system (F_x, F_y and F_z), and are used for the trajectory integration in the next time step. In short, the trajectory integration is done in the Cartesian coordinate system, while the interpolation of evaluating electric field components using the current positions of electrons is done in the spherical coordinate system. Note that the local electric-field points are much closer to each other around the tip than those far away from the tip. This is because the initial velocity of electrons is very slow, and a higher density of electric-field points gives better accuracy for the interpolation calculation.

Once electrons are outside the range defined by the local trajectory, the global set of electric-field data is used. The grid of the global electric-field points is shown in Figure 4.3.5. All points are located at the intersections of the straight lines that are parallel to x, y or z-axes. Similarly, the electric field components used in the trajectory integration are evaluated by interpolation based on the current positions of electrons. All the simulations for the global electric-field trajectories are done in the Cartesian coordinate system.

The integration of an electron trajectory stops right after the electron reaches the virtual surface (see Figure 4.2.2). This is because the electric field distribution on and above the virtual surface is uniform. The exact positions on the virtual surface where the electrons penetrate are calculated using interpolation. Since the electric field is uniform above this virtual surface, the final positions of electrons on the anode are calculated analytically based on the electrons' positions and velocities on this virtual surface, assuming that a constant vertical force is applied on electrons above the virtual surface.

Figure 4.3.6 and Figure 4.3.7 show the procedure for the local and global electric-field trajectory simulations, respectively.

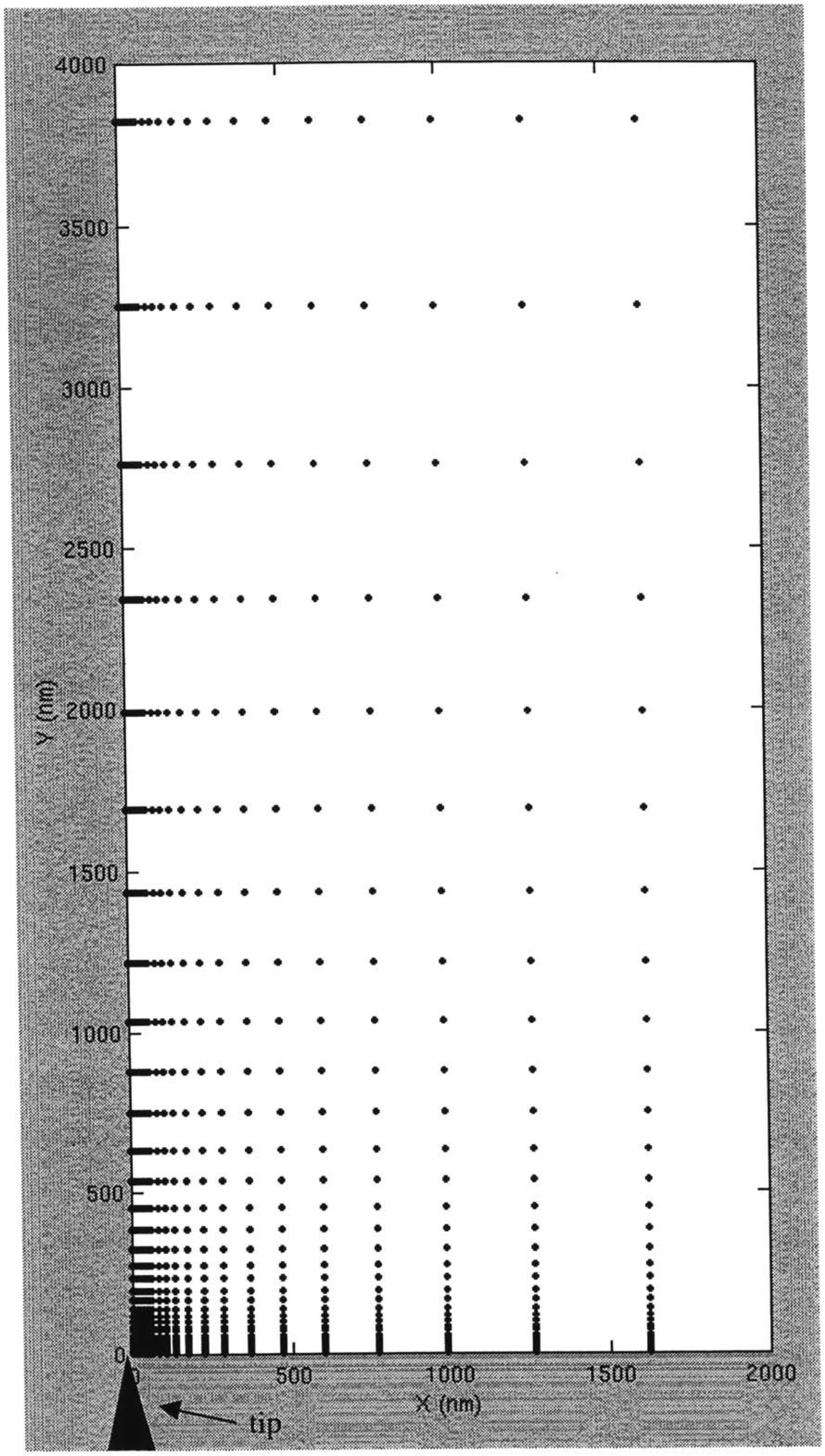


Figure 4.3.5 Side-view of the global electric-field points. Note that only one half of the points are shown. The tip is located at the origin (0,0,0).

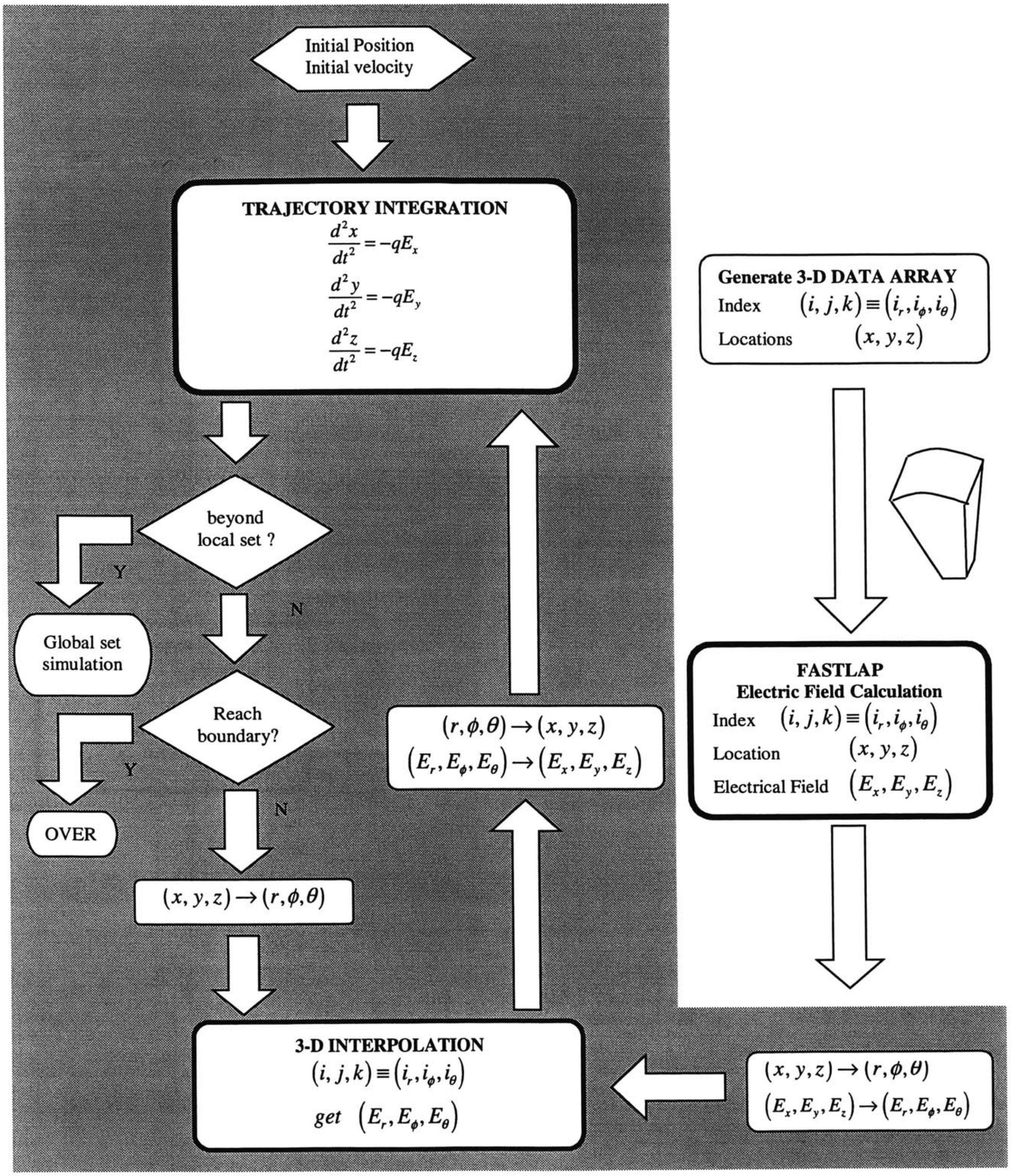


Figure 4.3.6 Flowchart of the procedure for the local electric-field trajectory simulation.

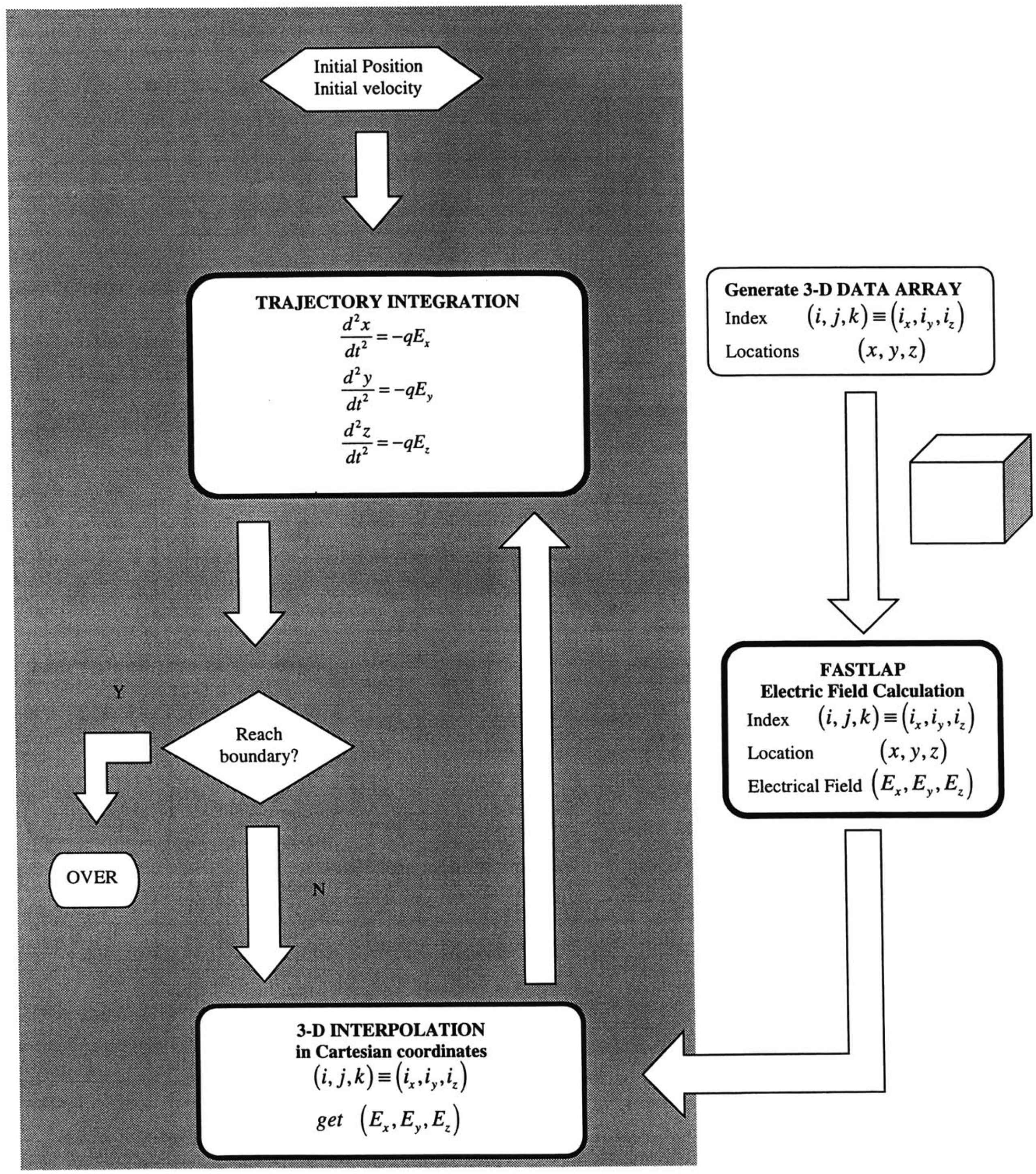


Figure 4.3.7 Flowchart of the procedure for the global electric-field trajectory simulation.

4.4 Simulated Results for Single Gated FEDs

Figure 4.4.1 shows the 3-D BEM surface mesh (4957 boundary elements) for a single emitter in a tip array with a pitch distance of 4 μm. The detailed dimensions can be found in Figure 4.4.2. A few closer views of the tip meshes are shown in Figure 4.4.3.

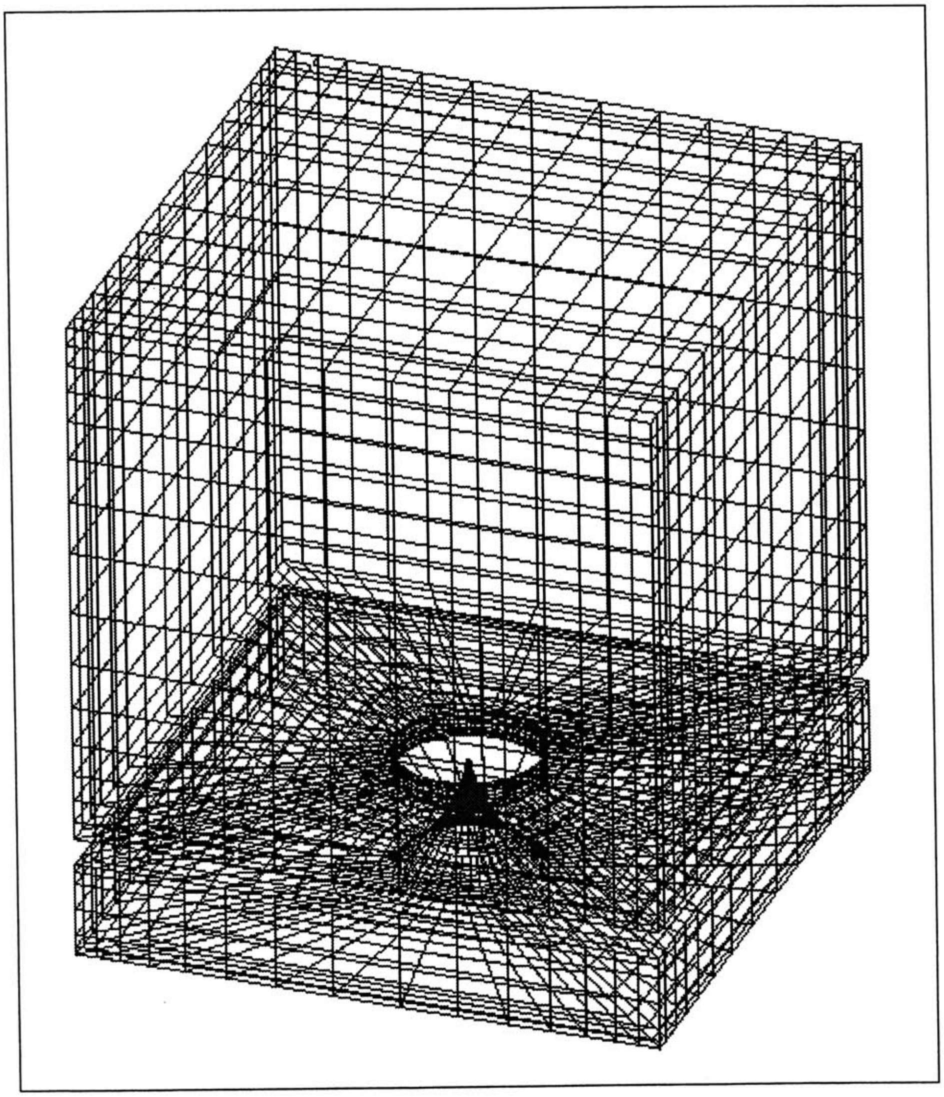
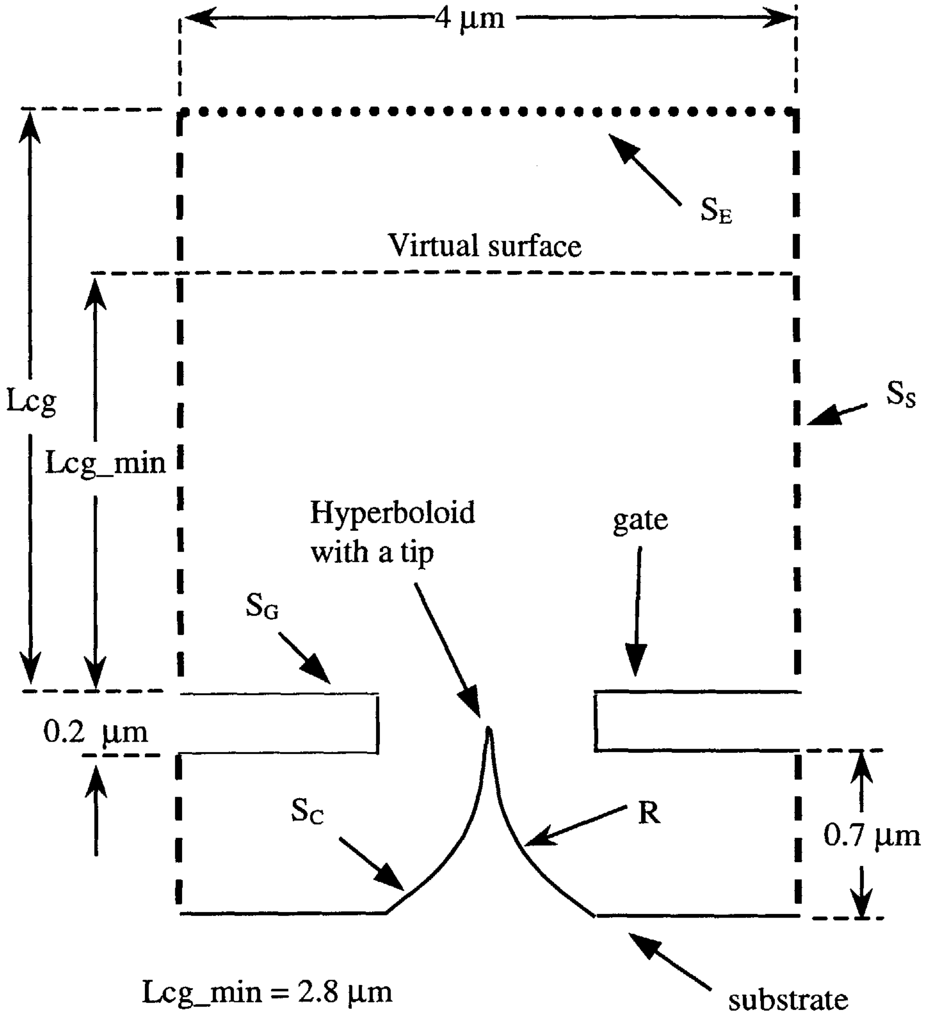


Figure 4.4.1 3-D boundary-element mesh plot of the emitter simulated. The total number of boundary elements in this model is 4957.



Lcg_min = 2.8 µm
Lcg = 4.0 µm
Radius of gate aperature = 0.5 µm
Raduis of hyperboloid base = 0.5 µm
Rad. of curvature of the hyperb. shape (R) = 1.2 µm

Figure 4.4.2 Configuration of the simulation model of the single gated FED.

In this model, the minimum required distance between the top surface of the simulation domain and the gate top surface (L_{cg_min}) is 2.8 µm. An SEM picture of a single emitter tip, which is fabricated by oxidation sharpening and CMP techniques, is shown in Figure 4.4.4. The radius of the gate aperture is 0.5 µm while the radius of curvature of the field emitter tip (with an added gold layer) is about 12.3 nm.

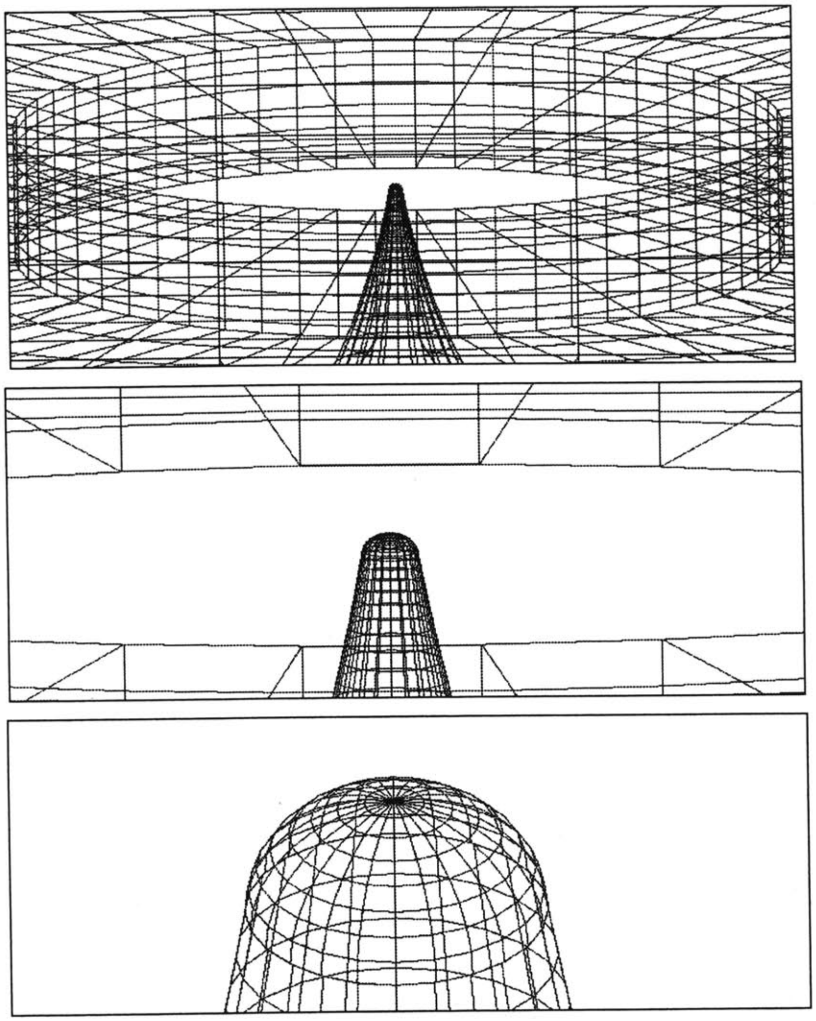


Figure 4.4.3 Close-up views of the tip with 8.2 nm radius of curvature.

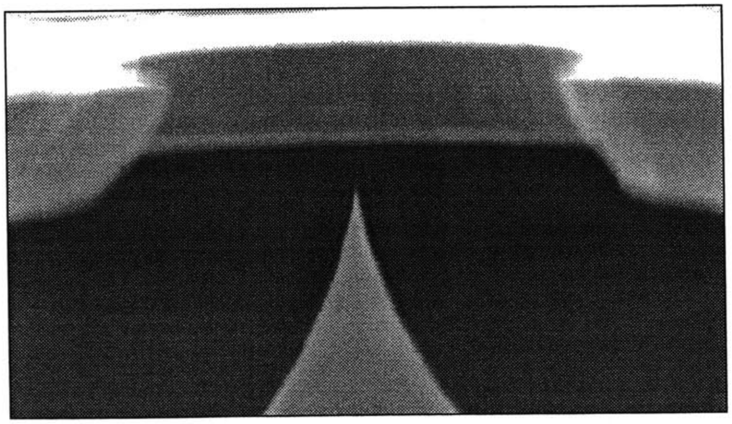


Figure 4.4.4 A SEM picture of the emitter simulated and tested.

The process of electrostatics simulation takes about 1 hour to obtain the electric field data that is needed for emission current and trajectory simulations. Note that since our model is symmetric in both the x and z-axes (see Figure 4.3.1), we only simulate the trajectories emitting from the panels on one quarter of the tip surface.

Figure 4.4.5 and Figure 4.4.6 show the I-V characteristics and the Fowler-Nordheim plots respectively for the single emitter, both measured and simulated. In the simulations, the tip radius of curvature (ROC) is selected to be 8.2 nm by matching the *slope* of the measured F-N curve. Note that this ROC is close to the estimated ROC from the SEM picture. With this ROC, the simulated and experimental data match very well *in both slope and magnitude* in the F-N plot up to the saturation region at high gate voltage. We conjecture that the saturation effect in the experimental F-N curve is due to electron velocity saturation in the hyperboloid's neck region where the cross-sectional area is very small. Note that the work function used in the simulation is 4.04 eV.

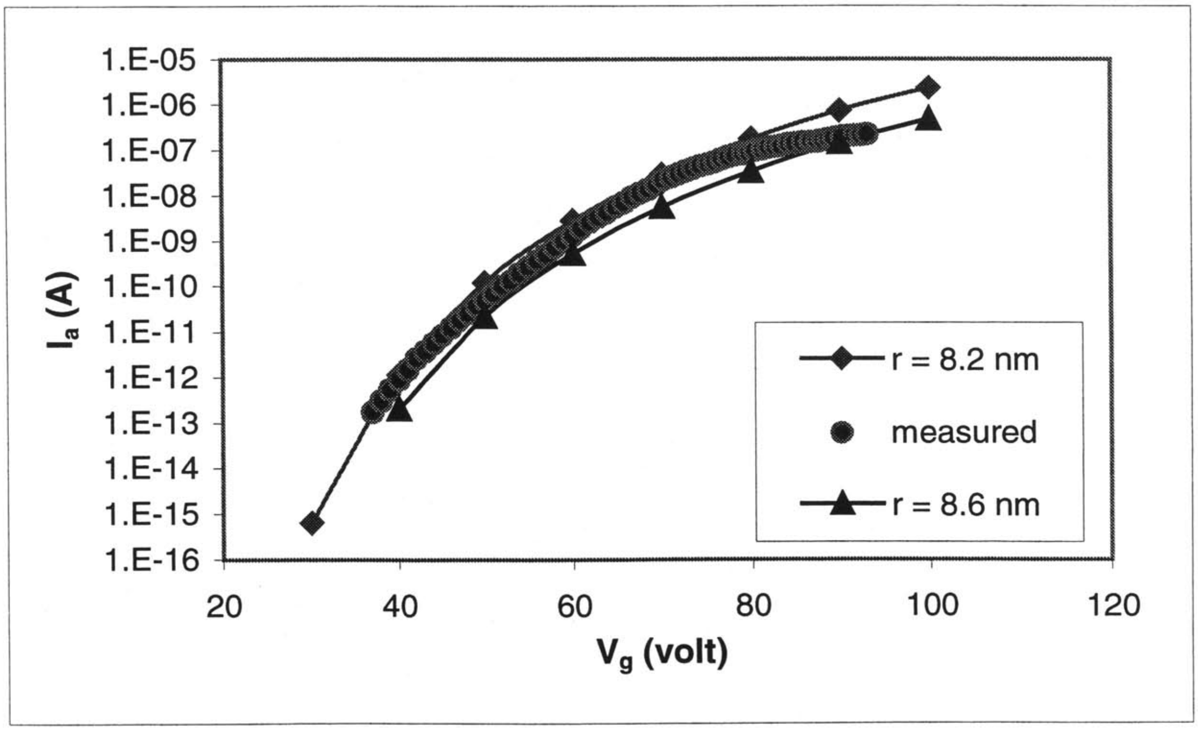


Figure 4.4.5 Simulated and measured I-V characteristics (anode current versus gate voltage) for a single emitter. Note that the experimental current data are the current per emitter measured from a 60x60 emitter array. The anode voltage is 1000 volts, and distance between gate and anode is 10 mm.

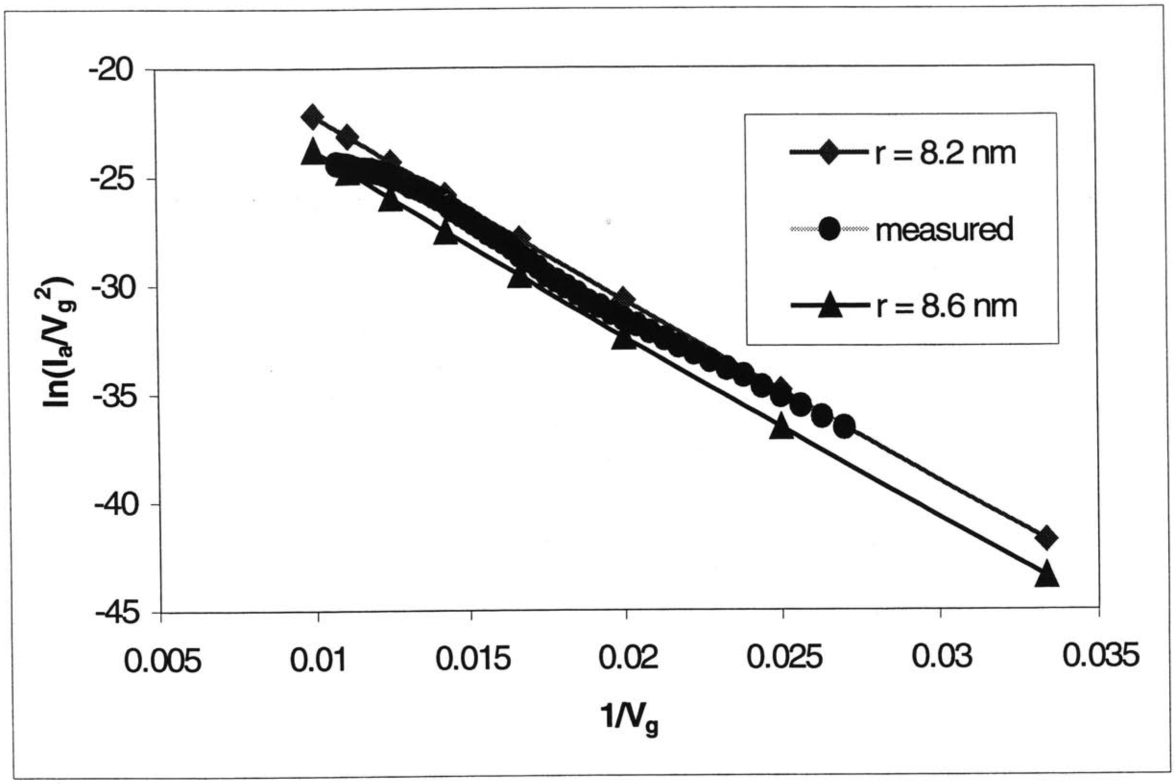


Figure 4.4.6 Fowler-Nordheim plot of the simulated and tested results. Note that the experimental current data are the current per emitter measured from a 60x60 emitter array. The anode voltage is 1000 volts, and distance between gate and anode is 10 mm.

Electron trajectories are calculated up to a *virtual surface* (see Figure 1) where the vertical field distribution is uniform. The final positions of electrons on the actual anode are calculated analytically based on the electrons' positions and velocities at this virtual surface. The emitter-array current distribution on the anode is then calculated by superimposing results from a single tip. The current density distribution for a 4-μm pitch 60x60 FEA at the phosphor screen is shown in Figure 4.4.7. The gate bias is 50V, the screen bias is 5kV and the screen to FEA distance is 10 mm. Figure 4.4.8 shows a phosphor screen image taken after 20 days of operation under the same bias conditions. The spot size is estimated to be 2.0 mm which is consistent with the estimated electron beam size of 1.8 mm if we account for electron beam scattering, carrier diffusion and photon scattering in the phosphor. This spot size is similar to the simulated data reported in [5].

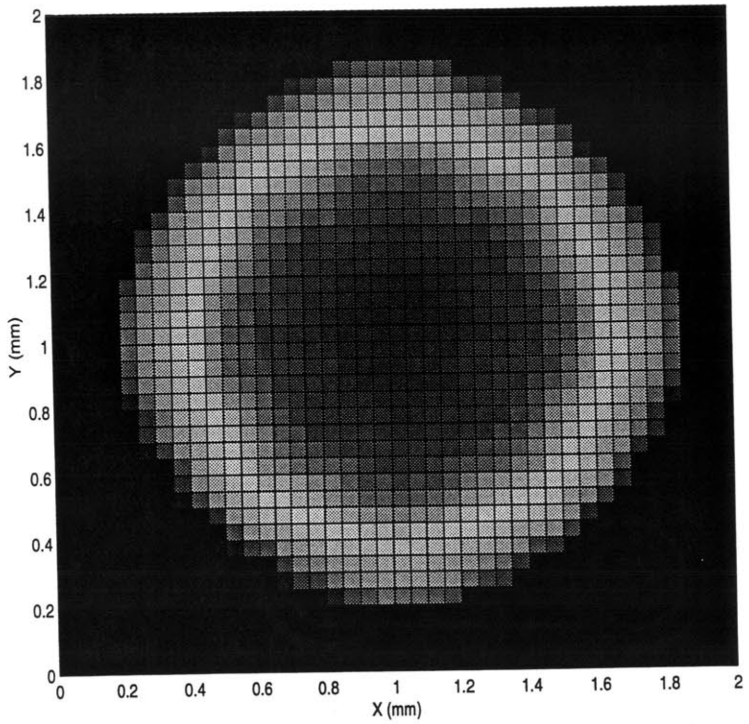


Figure 4.4.7 Anode current density distribution from a 60x60 emitter array. Note that (1) the gap between gate and anode is 10 mm, (2) the anode voltage is 5 kV, and (3) the gate voltage is 50 volts. With a constrast ratio of 1000, the spot diameter is about 1.8 mm. The total current of this array is $1.09\text{x}10^{-7}$ A.

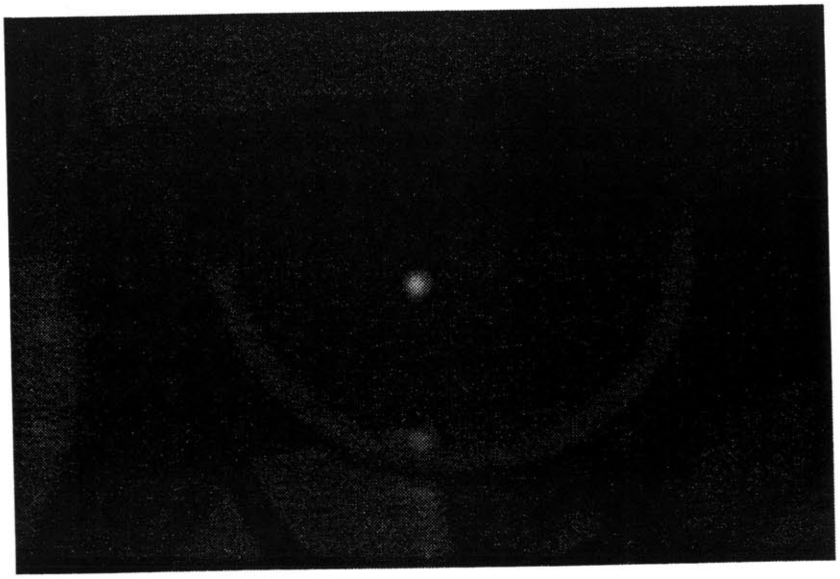


Figure 4.4.8 Picture of phosphor spot from a 60x60 emitter array.

4.5 Proximity Focusing Study

With the same BEM model shown in Figure 4.4.1 and Figure 4.4.2, the spot size and emission current density for the proximity focusing cases are calculated by changing the electric field boundary condition on the top Neumann boundary (S_E). In this study, the voltages on the anode and gate were held at 5,000 and 50 volts respectively, while the distance between the anode and gate was varied between 1 mm and 10 mm. The spot size and total current of a 60x60 FED array are shown in Figure 4.5.1.

As expected, the spot size increases with the cathode-anode distance while the emission current decreases due to the reduced anode field. The results at cathode-anode distance of 1 mm are consistent with previous work [5].

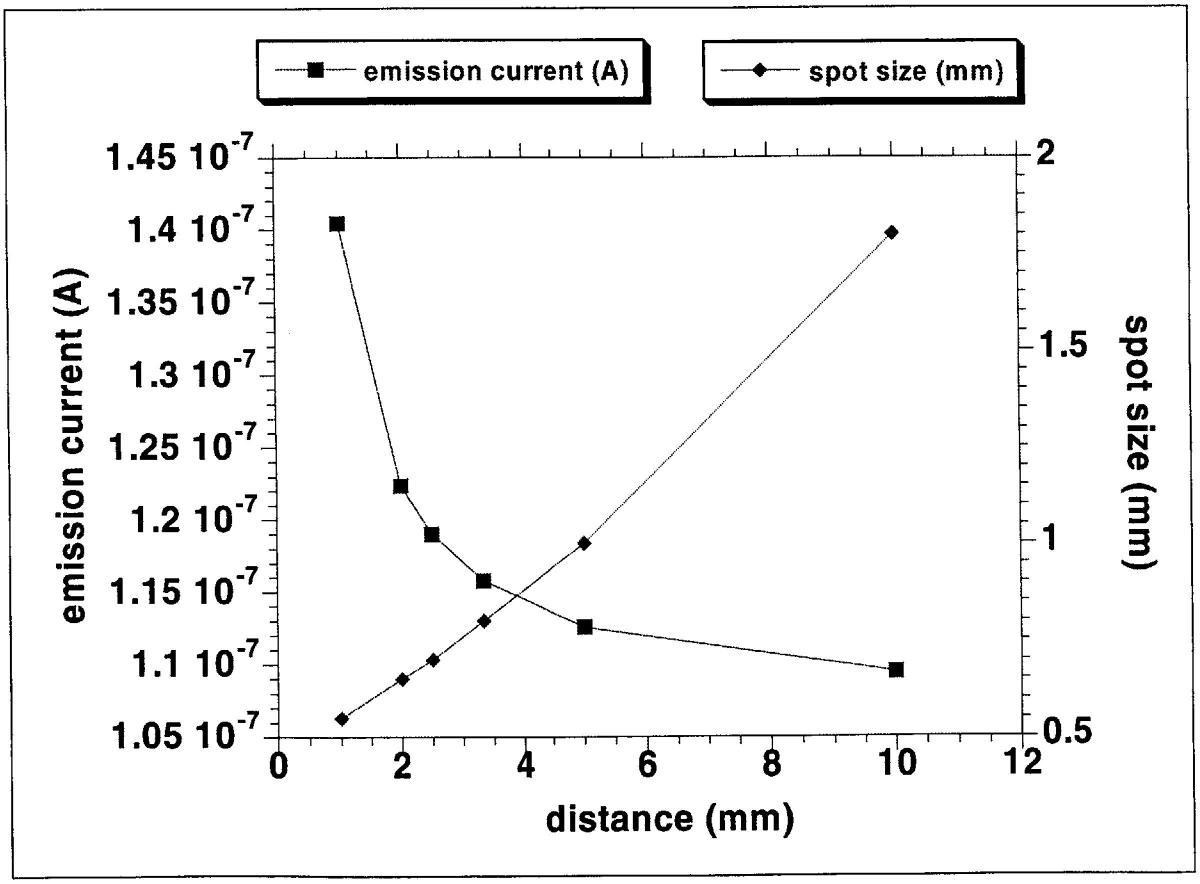


Figure 4.5.1 Simulated results of spot size and total emission curent vs. anode-gate distance for a 60x60 FED array shown in Figure 4.4.1.

4.6 Double Gated FEDs for Integrated Focusing

The integrated-focus-electrode (IFE) FEA structure shown in Figure 4.6.1 was also studied because of its technological importance. In this design, a second annular gate is used to reduce the beam size. The structure has been proposed as a solution to the problem of the trade-off between luminous efficiency and resolution in proximity focused FEDs. Figure 4.6.2 demonstrates the reduction in spot size as the focus voltage is decreased with a gate biased at 50V and a gate-anode distance of 10 mm. However, the emitted current decreases with focus voltage because of the reduction of the tip electric field, which is consistent with the results of Hosono [82]. The resolution of the IFE FEA increases as the focus voltage decreases; however, this also results in a decrease of emission current and brightness. It is possible to compensate for the decrease in emission current by increasing the gate voltage. A disadvantage of this approach is an increase in the power consumption of the display driver circuit.

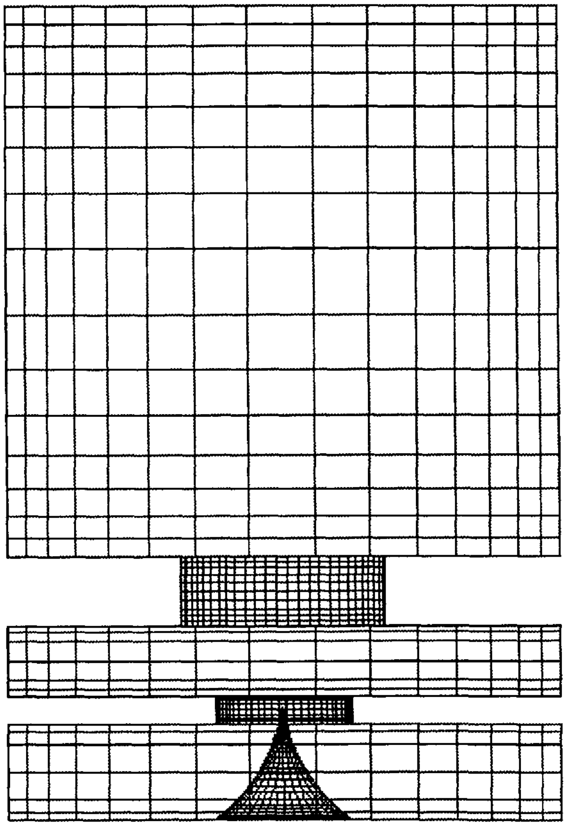


Figure 4.6.1 Side view of a 3-D mesh plot of an IFE FED. The focusing gate aperature is 1.5 μm in diameter, and thickness is 0.5 μm. The distance between gate and focusing gate is 0.5 μm.

With an adequate focus voltage, we can obtain even better (smaller) spot size than the proximity focusing case with a narrow gate-anode distance. However, higher gate voltage is needed in order to increase the emission current to the same emission current level used in the proximity focusing design.

Note that the minimum spot size occurs when 15 volts is applied on the focusing electrode. Below this voltage, the focus electrode over-compensates for the lateral velocity of the electrons after emission. The spot size increases because electrons cross the axis of symmetry as they travel from the gate aperture to the anode [45].

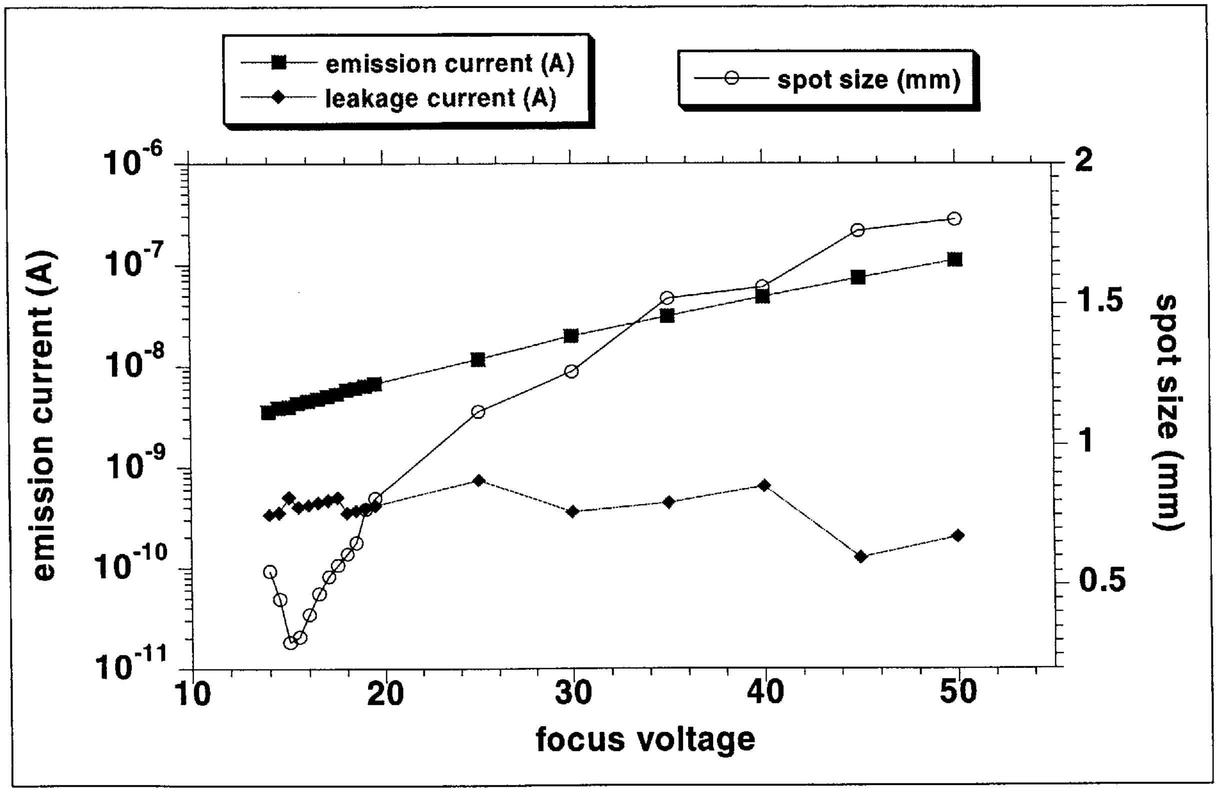


Figure 4.6.2 Simulated results for a 60x60 array of an IFE FED shown in Figure 10. Note that (1) the distance between gate and anode is 10 mm, (2) the anode voltage is 5 kV, and (3) the gate voltage is 50 volts.

CHAPTER 5

DESIGN OF EXPERIMENT

In this chapter, a series of methods of *design of experiment* (DOE) will be discussed. These methods are the foundation of design strategy for the FED design discussed in the next Chapter. We first introduce the two-level factorial design, which is used to estimate the main effects and interaction effects of input factors. A response surface can then be constructed after expanding the two-level factorial design into a central composite design. The regression analysis for generating a model that describes the relationship between input factors and responses is also presented.

5.1 Introduction to Design of Experiment

An experiment is "a test or series of tests in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes in the output response" [27][83]. An experimenter starts with a simple speculation regarding an experiment. Take the FED design as an example: "High voltage phosphor is preferable for our FEDs so that anode-gate distances of greater than 500 μm and IFE structures are required?" "How do we choose the IFE voltages for different gate voltage (i.e., different gray scale or colors) so that the spot size will keep the same?" or "Are there other factors that affect the spot size?" Speculations and questions of this kind form the statement of a problem: "How to find out the relationships between the spot size, emission current, gate voltage and IFE voltage?" or "I would like to know which method is the most effective way to keep the spot size the same while changing gray scale/colors of the FEDs."

After the formation of statement of problem, a listing of all the factors that might influence the response of the experiment is created. This step is crucial because it determines the choice of the type of experimental designs, and it also defines the group of factors with respect to which inferences can be drawn. Based on these factors, a suitable experimental design has to be chosen. At this point, the reevaluation and reformulation of the experimental design might become necessary due to the following reasons:

- the experimental design might become too big and too complicated to be carried out under existing conditions.
- the physical limitations imposed by the available experimental conditions may make it impossible to obtain the response of the experiment.

Furthermore, in any experimental design, in addition to understanding the influence of factors on the response, there is an objective that the experiment is designed to achieve, such as optimizing, minimizing, or maximizing the response. Also, the measurement of response contains an error that is unavoidable due to measurement environment, so it is necessary to develop an experimental design that will give appropriate results, even in the presence of experimental error.

Generally, there are two components of experimental design: (i) treatment design, (ii) and error-control (reduction) design. The treatment design determines the treatments to be included in the study. Usually, the treatments are chosen to have structural forms, in particular, level-factorial structure. The number of treatments is determined by the number of factors and the level combinations of each factor.

Aspects of treatment design are closely connected to aspects of error-control design. Error-control design determines the actual arrangement of treatments. Examples of such design are the completely randomized design, blocked design, Latin square design, etc. In most cases, the chosen treatment design is embedded into an appropriate error-control design using statistical analysis.

Note that in this study, the experiments are conducted using the 3-dimensional CAD tool described in the previous chapters. In other words, the system to be modeled is simulation based. The whole simulation process is in fact a series of *virtual experiments*, which employs the method of experimental design [27][84][85][86] on the numerical simulation to find out the optimal design of the devices without actually performing process experiments. Since the simulation intrinsically does not generate error, this study will mainly focus on the treatment design.

5.2 Two-Level Factorial Design

For experiments involving two or more factors, factorial design is the most efficient method to study the relationship between parameters and response. Especially, factorial designs can explore the interaction effect that one-at-a-time experiments cannot achieve [84].

The definition of the main effect is the effect of a factor on the response produced by a change in the level of the factor. In two level design, each factor has two levels, denoted by "-" and "+", and also called "low" and "high". Therefore, the main effect of a factor is the difference between the average response at the high level of the factor and the average response at the low level of the factor. For example, the main effect of a factor A is

$$Y_A = \bar{y}_{A+} - \bar{y}_{A-} \tag{60}$$

where $\bar{y}_{A+}$ and $\bar{y}_{A-}$ is the average response at the high and low level of factor A, respectively.

Note that all the responses from the complete combination of factor levels are being used to supply information on each of the main effects.

In most cases, however, the difference in response between the levels of one factor is not the same at different levels of other factors. These effects are called interaction effects, and can be calculated from factorial design response data. Note that sometimes the interactions between factors can overshadow the significance of the main effects, especially for two-factor interactions. In order to explain the importance of the interaction effects, a simple two-level factorial design with two factors will be discussed.

Figure 5.2.1 shows the experiment space for a two-factor two-level factorial design. The definition of interaction effect for this case is:

$$Y_{AB} = \frac{\left(\left(y_{A+,B+} - y_{A-,B+}\right) - \left(y_{A+,B-} - y_{A-,B-}\right)\right)}{2} = \frac{\left(\Delta y_{A,B+} - \Delta y_{A,B-}\right)}{2} \tag{61}$$

or

$$Y_{AB} = \frac{\left(\left(y_{A+,B+} - y_{A+,B-}\right) - \left(y_{A-,B+} - y_{A-,B-}\right)\right)}{2} = \frac{\left(\Delta y_{B,A+} - \Delta y_{B,A-}\right)}{2} \qquad (62)$$

The first equation calculates the difference between the difference of the two levels of factor A at upper level of factor B, and the difference of the two levels of factor A at the lower level of factor B. The second equation in fact is the same as the first one, but interprets the interaction effect from a different point of view.

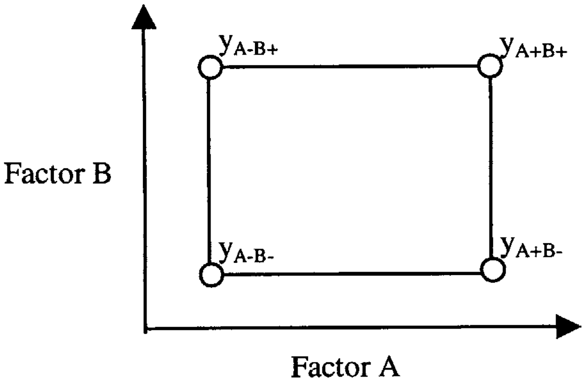


Figure 5.2.1 Experiment space of a two level factorial design with two factors.

Table 5.2.1 gives three sets of response values to demonstrate different cases of main effects and interaction effects. The calculation of these effects use equations (60) and (61). The first case shows that the difference in response between the levels of one factor is the same at the two levels of the other factor, so there is no interaction effect between the two factors. Figure 5.2.2 shows the concept of zero interaction. The responses vs. factor A at different levels of factor B are two parallel straight lines, which means that the change of level of factor B does not affect the change of effect due to varying the level of factor A.

Figure 5.2.3 and Figure 5.2.4 show the concept of mild and strong interaction between factors A and B. The advantage of the factorial design over one-factor-at-a-time method [27] is its ability to detect the interaction effects by using a nominal number of experiments.

RUN	Factor Level		**CASE I** Response	**CASE II** Response	**CASE III** Response
	A	B			
1	-	-	40	40	40
2	+	-	80	80	80
3	-	+	60	60	120
4	+	+	100	120	60

Main effect Y_A	40	50	-10
Main effect Y_B	20	30	30
Interaction effect Y_{AB}	0	10	-50

Table 5.2.1 Three cases of two-level factorial design with 2 factors.

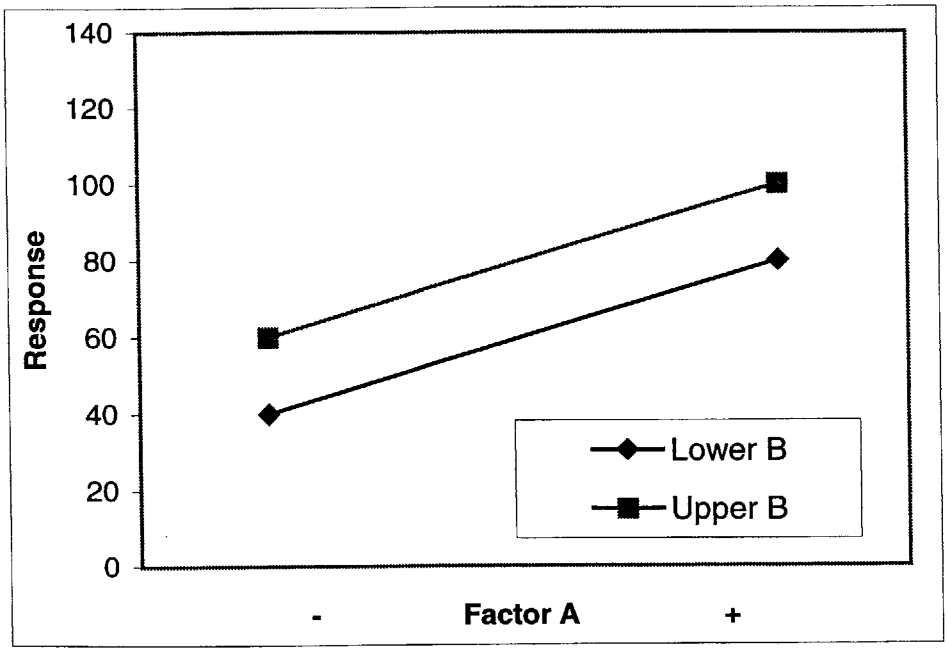


Figure 5.2.2 Case I : zero interaction between Factors A and B. The figure shows the response vs. factor A at different levels of Factor B.

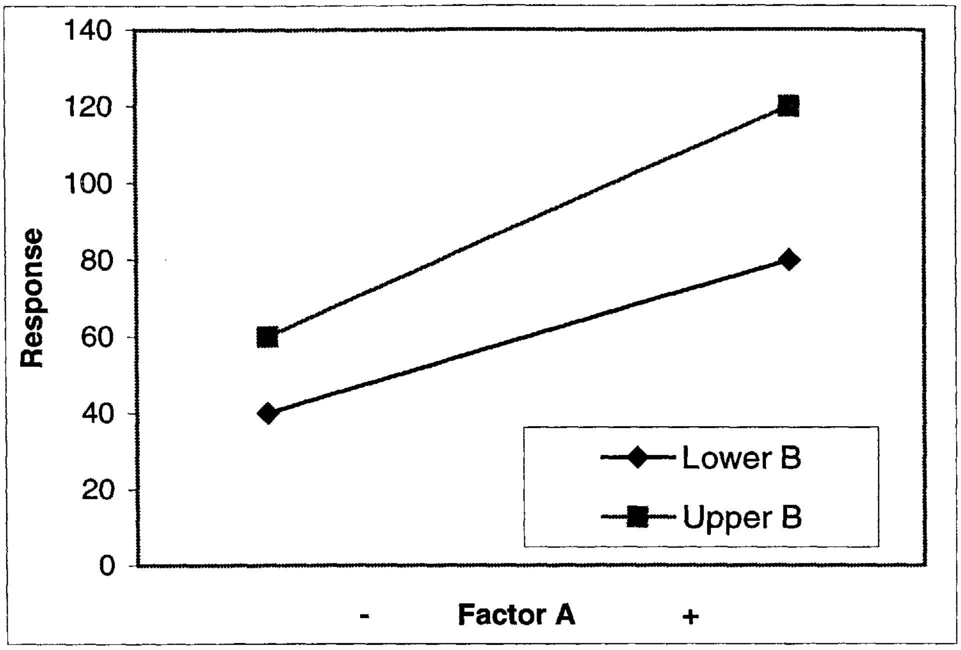


Figure 5.2.3 Case II: Mild interaction between Factors A and B. The figure shows the response vs. factor A at different levels of Factor B.

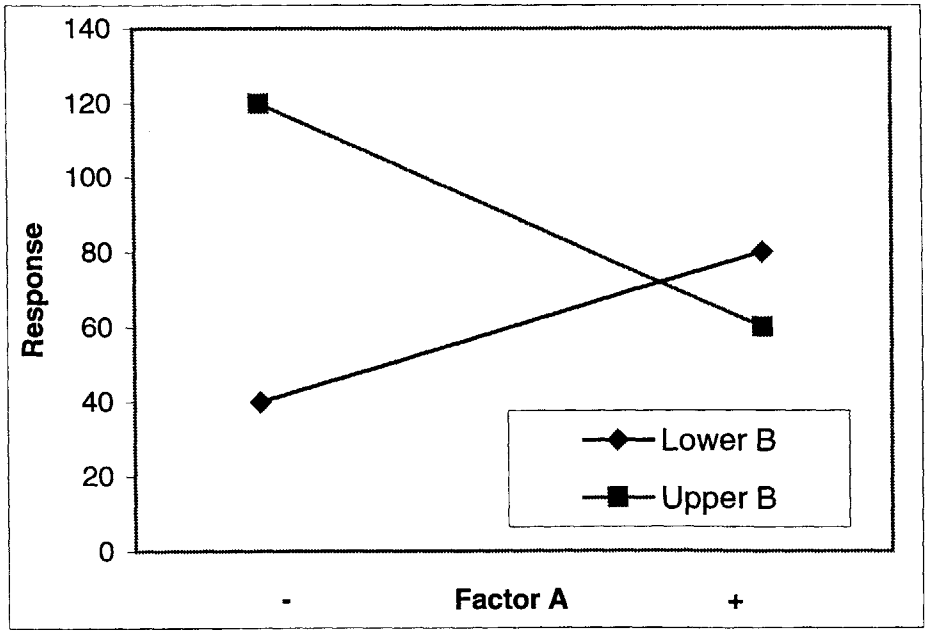


Figure 5.2.4 Case III: Strong interaction between Factors A and B. The figure shows the response vs. factor A at different levels of Factor B.

In two-level factorial design, the number of treatments is 2^k, where k is the number of factors. Table 5.2.2 shows two-level factorial design pattern up to four factors. For large k the number of treatments may become too large to accomplish. In this case, a fractional factorial design can be implemented.

Index of treatments	X1	X2	X3	X4
1	-	-	-	-
2	+	-	-	-
3	-	+	-	-
4	+	+	-	-
		Two factors (2^2=4)		
5	-	-	+	-
6	+	-	+	-
7	-	+	+	-
8	+	+	+	-
		Three factors (2^3=8)		
9	-	-	-	+
10	+	-	-	+
11	-	+	-	+
12	+	+	-	+
13	-	-	+	+
14	+	-	+	+
15	-	+	+	+
16	+	+	+	+
			Four factors (2^4=16)	

Table 5.2.2 two-level factorial design pattern up to four factors.

After the number of factors is determined, the next step is to identify the upper and lower limits of each parameter. Note that in factorial design analysis, the relationship between factors and responses are assumed to be linear. Unfortunately, in some cases there is an order-of-magnitude difference between upper and lower levels of a factor that are of interest, or there is an order-of-magnitude difference between the response of the upper level and the response of the lower level of a factor. For the former situation, it could be difficult to expand the results of the two-level factorial design to form a central composite design for response surface and optimization analysis (to be discussed later). In the latter situation, the linear characteristic of the two-level factorial design will give a poor estimate of the response between upper and lower levels. Some special treatments on these conditions will be proposed in the later sections.

5.3 Response Surface Methodology (RSM)

The *response surface methodology* (RSM) [87][88] is a group of mathematical techniques that can be used in the study and analysis of relationships between one or more measured responses and a number of input parameters (factors). It includes:

- Constructing a series of experiments (designing a set of experiments) which will give well-controlled responses in the range of factors that are of interest.
- Choosing a mathematical model that best fits the responses from the aforementioned design of experiments.
- Searching for the optimal combination of the design factors that gives the desired value of the responses.

The mathematical form of a response surface can be written as:

$$y = f(x_1, x_2, \ldots\ldots x_n) + \varepsilon \tag{63}$$

where n is the number of the input factors, and ε is the error observed in a response.

Contour plots are frequently used to help visualize the shape of the response surfaces. In most cases, the functional form of the relationship between the response and the input factors is unknown, so usually a low-order polynomial that is a function of the input factors is employed. The first-order model of a response surface is

$$y = c_0 + c_1 x_1 + c_2 x_2 + \ldots\ldots + c_n x_n + \varepsilon \tag{64}$$

where n is the number of input factors, and $c_1, c_2, \ldots\ldots c_n$ are unknown coefficients giving the linear relationship between $x_1, x_2 \ldots\ldots x_n$ and y. Note that each factor is related to the response in a linear fashion. Also, all factors independently affect the response.

If interaction effects of a model are significant, the first-order model can be expanded to handle interaction by adding terms consisting of product of factors. For example, if interactions between any two factors are of interest, the response surface model is:

$$\begin{aligned} y = c_0 + c_1 x_1 + c_2 x_2 + \ldots\ldots + c_n x_n + \\ c_{1,2} x_1 x_2 + c_{1,3} x_1 x_3 + \ldots\ldots + c_{n-1,n} x_{n-1} x_n + \varepsilon \end{aligned} \tag{65}$$

Finally, a reasonable model to consider when the knowledge of the response curve is unclear is the quadratic response model:

$$\begin{aligned} y = c_0 + c_1 x_1 + c_2 x_2 + \ldots\ldots + c_n x_n + \\ c_{1,1} x_1^2 + c_{2,2} x_2^2 + \ldots\ldots + c_{n,n} x_n^2 + \\ c_{1,2} x_1 x_2 + c_{1,3} x_1 x_3 + \ldots\ldots + c_{n-1,n} x_{n-1} x_n + \varepsilon \end{aligned} \tag{66}$$

The quadratic form includes linear terms, interaction terms and squared terms. This form of response surface can represent the curvature of a response in the system, and is called a second-order model. Obviously, the models shown in equations (64), (65), and (66) are unlikely to be a good approximation of a functional relationship of a system for wide ranges of all factors. However, for relatively small ranges of input factors, the polynomial models usually work well.

The evaluation of the coefficients in the polynomial models can be done by the method of least-square regression analysis. A brief description of the method of least squares is as follows.

Assuming that equation (64) is the response function whose coefficients are unknown. The matrix notation is:

$$\mathbf{y} = \mathbf{X}\mathbf{c} + \mathbf{e} \tag{67}$$

where $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_k \end{bmatrix}$, $\mathbf{e} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$,

$k+1$ is the number of coefficients (number of unknowns), and n is number of treatments (given responses and corresponding values of input factors).

In order to minimize the error due to fitting, we can define the "badness of fit" as:

$$BF = (\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc}) \tag{68}$$

By differentiating equation (68) with respect to the unknown $c_1, c_2, \ldots.. c_n$, we can obtain

$$\mathbf{X'Xc} = \mathbf{X'y}$$

which is also known as the matrix form of the least square *normal equations.* The unknown coefficients $c_1, c_2, \ldots.. c_n$ can be easily calculated by multiplying $(\mathbf{X'X})^{-1}$ on both sides of the normal equation.

$$\mathbf{c} = (\mathbf{X'X})^{-1}\mathbf{X'y} \tag{69}$$

Note that whatever solution we take, the vector $\mathbf{Xc}$ is unique given a particular vector $\mathbf{y}$, and the minimum value of $BF\left(= (\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc})\right)$ is obtained. This procedure gives a set of coefficients $(\mathbf{c})$ that ensures a minimum value of the accumulated error of responses for all the treatments (combinations of given input factors). The quality of the response prediction depends on the choice of the forms of the polynomial functions and the combination of the treatments $\mathbf{X}$ (i.e., the experiment design). In other words, the selections of types of experimental designs and forms of polynomial are critical to construct the best response surface. The method of least squares is only a technique to extract the parameters of the polynomial function we choose, and it does not affect the accuracy of the response prediction in this study.

The method of least squares can also be applied to the polynomials as equations (65) and (66). Since all the factor levels of each treatment are given, the interaction and pure quadratic terms can be evaluated first, then the same procedure can be used to estimate the parameters for each term.

5.4 Central Composite Design : A Combination of Factorial Design and RSM

In this study, we choose the *central composite design* (CCD) [87][88] because it not only possesses the advantage of analyzing the effects of individual factors and interactions, but it also allows experimenter to create response surfaces of the model. The central composite design was proposed by Box and Wilson in 1951, which is an alternative class of design to the three-level factorial design, but in fact is a combination of a two-level factorial design with few more treatments appended to ensure better capability of prediction of second-order responses. Each factor is used at five different levels. The central composite design consists of three parts:

- A two level factorial design of k factors. The upper and lower levels of each factor are denoted as -1 and $+1$. There are a total of 2^k treatments.
- Axial points. Each axial point is at a distance of α from the center of the design. The easiest choice of axial points is to set a factor at a level either $+\alpha$ or $-\alpha$ $(\alpha > 1)$, then set the levels of all the other factors to 0. There are a total of $2 \cdot k$ treatments in this part.
- A center point. The level of each factor at the center point is zero.

Table 5.4.1 shows the configuration of the central composite design. Note that since in this study numerical simulations (virtual experiments) do not give any random error, unlike the typical central composite design used in physical experiments, there are no replicates for each treatment, and the total number of treatments in our central composite design is $2^k + 2 \cdot k + 1$, where k is the number of factors.

Figure 5.4.1 and Figure 5.4.2 illustrate each treatment of the central composite design for 2 and 3 factors.

TREATMENTS				DESCRIPTION
X_1	X_2	....	X_k	
±1	±1	....	±1	two level factorial points (2^k points)
$-\alpha$	0	....	0	axial points ($2 \cdot k$ points)
$+\alpha$	0	....	0	⋮
0	$-\alpha$	....	0	⋮
0	$+\alpha$	....	0	⋮
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
0	0	....	$-\alpha$	⋮
0	0	....	$+\alpha$	⋮
0	0	0	0	center point

Table 5.4.1 Configuration of a central composite design.

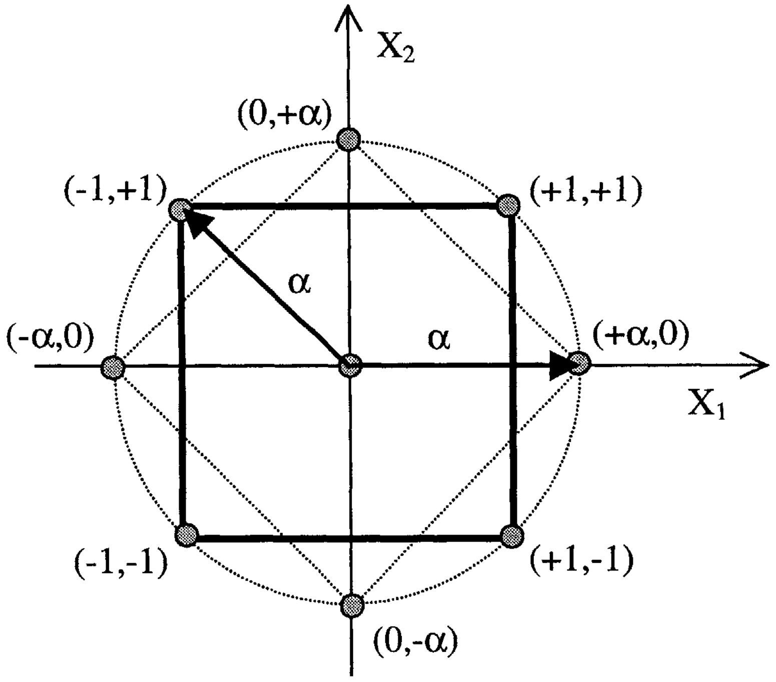


Figure 5.4.1 Representation of a central composite design with two factors.

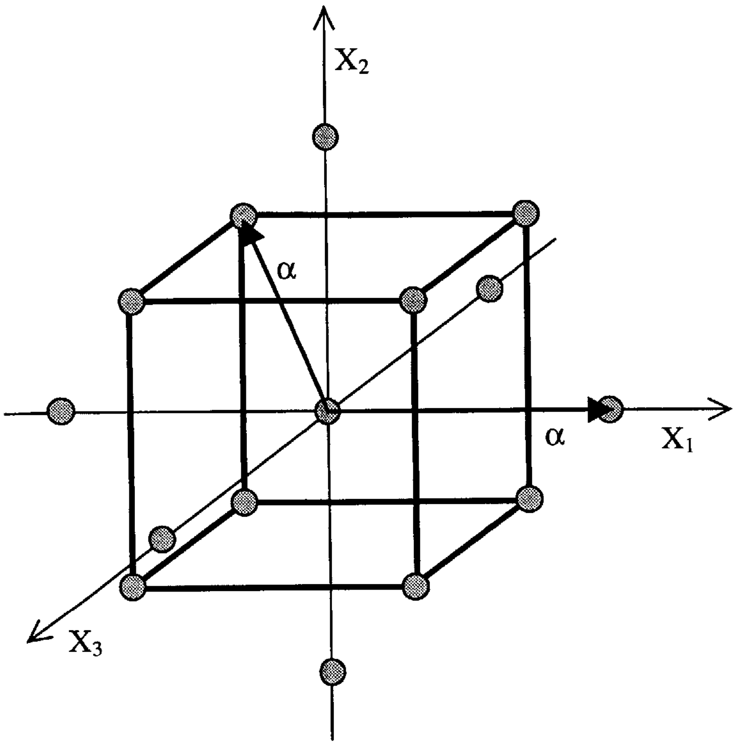


Figure 5.4.2 Representation of a central composite design with three factors.

There are a few ways to determine the value of α. Most of them are dependent on the constraints due to the analysis of variance. Since our virtual experiments do not have random error and the analysis of variance is not of interest in this study, the choice of α is based on the spherical property, which puts all the factorial and axial points of treatments on the surface of a sphere of radius $\sqrt{k}$. Because all the treatments have the same distance from their locations to the center point, a quadratic polynomial may give best prediction of responses. Figure 5.4.1 demonstrates geometrically that all the treatments are located on the same circle of radius $\sqrt{2}$, and Figure 5.4.2 demonstrates geometrically that all the treatments are located on the same sphere of radius $\sqrt{3}$. More discussion about the choice of α can be found in [87].

CHAPTER 6

DESIGN STRATEGY OF FEDS

In this chapter, a FED design procedure and results are presented. The chapter describes the integration of the simulation tools, which is essential for efficient experimental design for FEDs, followed by design examples for two types of FEDs. For proximity focused FEDs, a preliminary experimental design using six parameters is used. After the screening process, regression models for generating response surfaces are built using four out of the original six parameters. For integrated-focus-electrode FEDs, the preliminary experimental design has six parameters, followed by a four-parameter design and a three-parameter design to determine the optimal performance. The choice of parameters and their ranges of values are described in detail.

6.1 Introduction

A typical FED structure has several design parameters (input factors in the design of experiment), such as the applied voltages on various sub-components of the structure (gate, focus, anode and cone), and the dimensions of the structure (the radius of curvature of the tip, the apertures of the gate and focus, thickness of the gate and focus, the distance between gate and focus, etc.). Finding a set of design parameters which gives optimal performance of the device requires a wise strategy in order to minimize the required computation. The design strategy for the FED design consists of the following steps:

- Characterization: identify the significant design parameters (factors) and study the joint effects of the design parameters of FEDs by screening experiments (full/fractional two-level factorial designs)
- Optimization: use the response surface method to find the set of parameters which optimizes the device performance
- Macromodel: find analytical expressions for device performance as functions of device geometry and applied voltages

The main purpose of the first step is to determine the parameters that have significant influences on the response of the system. This step is necessary if the number of preliminarily chosen parameters is greater than or equal to six [85]. A large number of parameters is undesirable in constructing response surfaces. Also, it is rare for a practical system to have more than three significant parameters. Our approach in this step is to employ a two-level factorial design and determine the effects of the preliminarily chosen parameters. Also, by the factorial design, the interaction effects of design parameters on a response can be easily revealed.

After the significant parameters are determined, a new two-level factorial design process is performed with these significant parameters. The ranges of values of these parameters applied on this new two-level factorial design are also revised based on the previous results. A few more treatments are appended to the treatments of the new two-level factorial design to form the central composite design, which is primarily used for constructing response surfaces. The coefficients of a quadratic polynomial of the response surface are calculated by using the linear regression. The optimization can then be performed on the polynomial to determine the set of values of the parameters that give the desired system response.

The final step is to determine the device performance and response as a function of design parameters. Quadratic polynomials are originally used to represent the relationship between them. However, the quadratic form may not be a desired analytical form for the macromodel of a system. An ideal approach is to use a functional form of an existing analytical solution of the response of a simple FED that can be found in literature, then extract the coefficients of the functional form by the response results from the virtual experiments [89].

6.2 Mesh Generation and Integration of Simulation

As described in the previous section, a large number of treatments are performed to determine the effects from each possible significant parameter during the step of characterization. Although our virtual experiments are actually performed through computer simulation, for situation requiring a large number of simulations with different configurations, the most time consuming process may not be the computational time. The generation of solid models for structures with different dimensions requires tremendous effort if a standard solid modeling package (such as IDEAS) is used *manually*.

Also, if the execution of each step in a complete simulation for a single treatment is not integrated, users have to monitor the whole simulation process and provide appropriate commands between two consecutive steps to keep the whole procedure moving forward, which significantly reduces the efficiency of the simulation. As a result, in order to effectively perform a large number of simulations, automation of mesh generation and integration of simulation are essential [90].

Figure 6.2.1 is the block diagram of the integrated system for the design of experiment of FEDs. Listed below is the description of main codes (scripts) that perform each critical step during a complete design of experiment:

➤***DOE_CCD.m** : matlab code*

(2-3 hours for 77 configurations of a six-parameter central composite design)

- ➤*generate parameters of different runs (configurations) of CCD*
  - ✧ *create data structure for all simulation and mesh generation (about 100 variables)*
  - ✧ *create sub-directory*
  - ✧ *generate panel data for Fastlap 2.0*
  - ✧ *generate mesh data for TECPLOT (visualization)*
  - ✧ *create soft-links for binary executable files*
    - o *yjyP_quad, yjyQ_quad, yjyQE_quad, area_DOE, organizeQE, RUN_fastlap, tyjylg, maparray*

        - ✧ *create data : boundary condition input file (surface_BC.dat)*
        - ✧ *create data : centroid coordinates and panel area for tip panels (tip_area.dat; tip_centroid.dat)*
        - ✧ *create data : local and global E-field evaluation points in space (local_Ept.dat; global_Ept.dat)*
        - ✧ *create data : surf_pts.dat and surf_area.dat for analyzing emission current*
        - ✧ *create data : trajectory calculation input data (paralg.dat)*
        - ✧ *create data : emission current analyzer input data (para_map.dat)*
        - ✧ *create data : spot size analyzer input data (para_maparray.dat)*
    - ➤ *generate RUN (UNIX script)*
- ➤ ***RUN:*** *UNIX script*
  (3-5 days for 77 runs of a central composite design)
    - ➤ *call all electrostatics and trajectory simulation C codes for all runs*
- ➤ ***PostProcess.m*** *: matlab code (2-3 hours for 77 responses)*
    - ➤ *extract all current components and spot size for each run*
- ➤ ***LinearRegression.m*** *: matlab code*
    - ➤ *Full quadratic polynomial response surface of central composite design*
    - ➤ *analyzed responses*
        - ✧ *maximum current density on anode*
        - ✧ *spot size (diameter)*
        - ✧ *total emission current*
        - ✧ *Figure of merit*
- ➤ ***FindOptim.m*** *: matlab code*
    - ➤ *constraint for each parameters*
    - ➤ *optimize selected response*

Figure 6.2.2 shows the steps of a complete experimental design (listed by the main code (script) names). DOE.m is a Matlab script that generates solid models of the device and creates all the necessary data and executable files into specific sub-directories for each treatment. The computation time for DOE.m is approximately equal to $2 \cdot N_T$ (in minutes), where N_T is the number of treatments. The BEM models generated by DOE.m can be visualized by a package TECPLOT [91]. RUN is a UNIX C-shell script generated by DOE.m. This script executes all the electrostatic and trajectory simulations for each treatment in an experimental design. This step is the most time consuming step in terms of computation. The approximate computation time (the number of panels for each treatment is less than 8000) is equal to N_T hours. PostProcess.m is a Matlab script that extracts the emission currents, leakage currents and spot sizes from the electrostatics and trajectory simulation results (responses). LinearRegression.m creates quadratic polynomials that fit the responses, and FindOptim.m analyzes the response and uses the constrained optimization schemes provided by Matlab to optimize the responses.

The dependence of the internal files created and used by each code (script) is tabulated in Table 6.2.1. Note that in addition to the script LinearRegression.m, we use a statistics package JMP [92] to analyze the factorial design and response surface results. This package provides efficient ways to visualize the effect of each parameter and to estimate the coefficients of the quadratic curve fitting for each response surface.

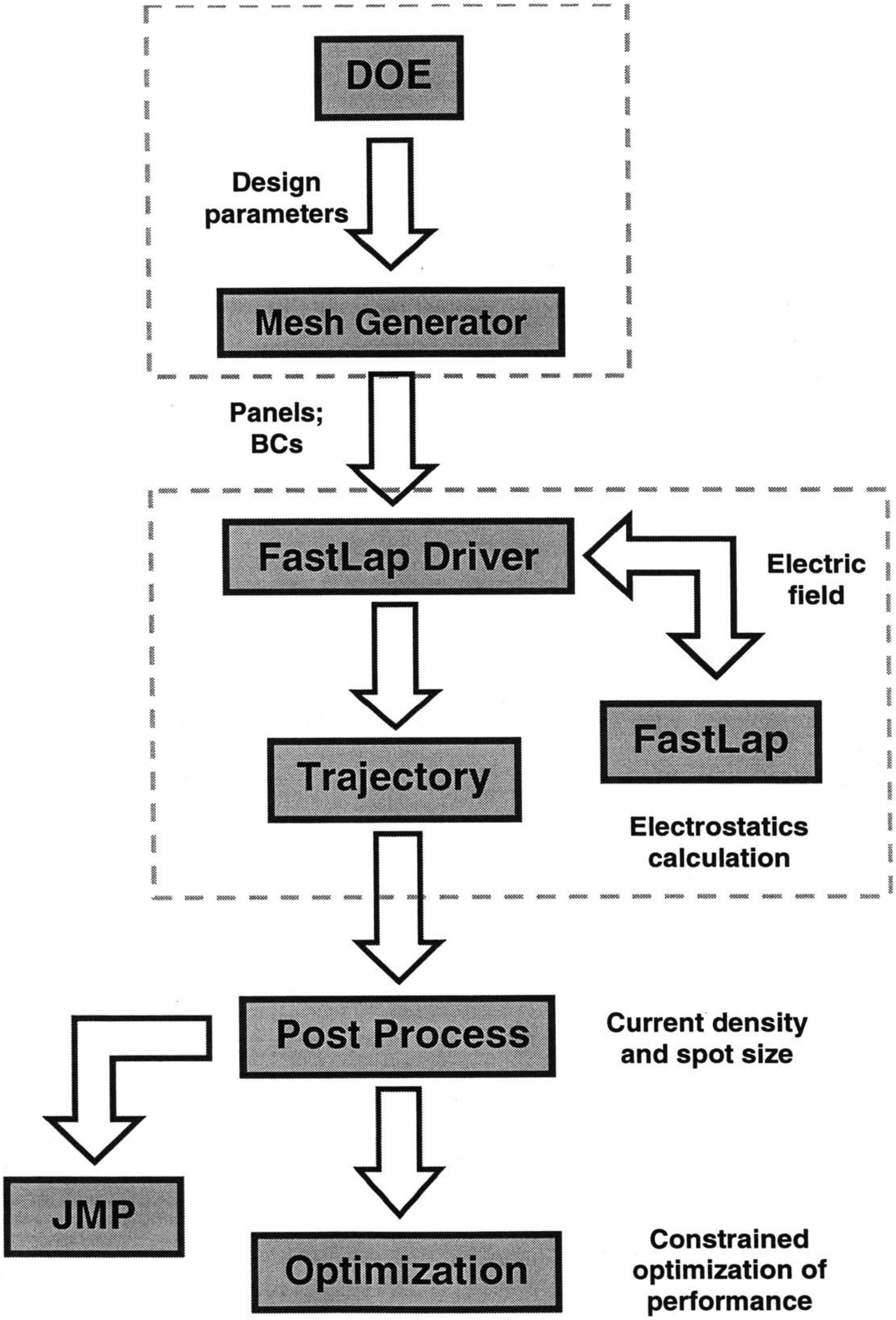


Figure 6.2.1 Block diagram of the system for experimental design of FEDs.

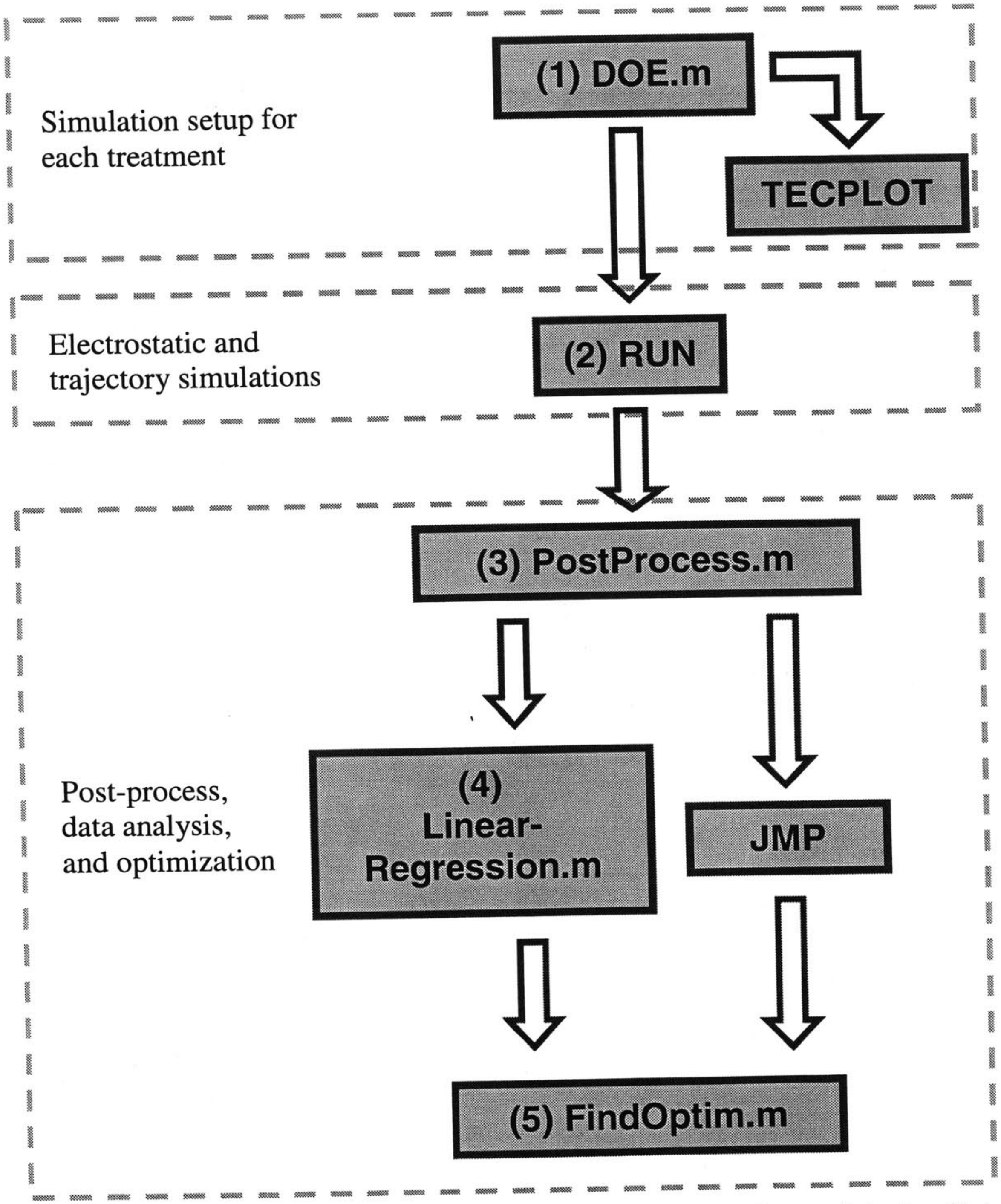


Figure 6.2.2 Steps of simulation for a complete cycle of a experimental design (listed by code (script) names).

Files	File type	Created by	Used by
yjy_quad.dat	Data	1	2
FEM.dat	Data	1	TECPLOT
surface_BC.dat	Data	1	2
yjyP_quad	Executable	1	2
yjyQ_quad	Executable	1	2
yjyQE_quad	Executable	1	2
area_DOE	Executable	1	2
organizeQE	Executable	1	2
RUN_fastlap	Executable	1	2
tyjylg	Executable	1	2
maparray	Executable	1	3
tip_area.dat	Data	1	2
tip_centroid.dat	Data	1	2
local_Ept.dat	Data	1	2
global_Ept.dat	Data	1	2
surf_pts.dat	Data	1	3
surf_area.dat	Data	1	3
paralg.dat	Data	1	3
para_map.dat	Data	1	3
para_maparray.dat	Data	1	3
Final.dat	Data	3	JMP

Table 6.2.1 Dependence of files of an integrated system for simulation of an experimental design system. The numbers listed on the table are the code index shown in Figure 6.2.2.

6.3 Proximity Focused FED Design Example

6.3.1 Preliminary Analysis: Six Parameters

The first example of the experimental design is for proximity focused FEDs. The basic structure of a proximity focused FED is shown in Figure 6.3.1. In this design strategy, six parameters that are possibly influential on the responses (emission current, spot size, etc.) are chosen for a preliminary characterization simulation. The central composite design is used as the first design strategy to filter out the significant parameters. These parameters are:

- the tip radius of curvature (ROC),
- the gate voltage (V_{gate})
- the radius of gate aperture (R_{gate})
- the uniform electric field (E_{top}) on the top surface of the simulation model
- the gate thickness (L_{gate})
- the vertical position of tip relative to the center of the gate aperture (L_{tip}).

Note that these parameters are chosen based on intuition and past experience.

In order to further expand the factorial design into a central composite design that is efficient for building a full quadratic device response surface, the range of values for each parameter must be carefully designed. Since the Fowler-Nordheim equation indicates an exponential relationship between tip surface electric field and emission current density, a small change of the tip radius of curvature gives rise to a significant change on the emission current density. This is because the surface electric field is approximately inversely-proportional to the tip radius of curvature. Consequently, the tip *ROC* is transferred into a logarithmic scale so that the five levels of the tip radius of curvature (in a central composite design) will not give a large variation in the responses (current densities and spot sizes) that would be difficult to be captured by a quadratic polynomial.

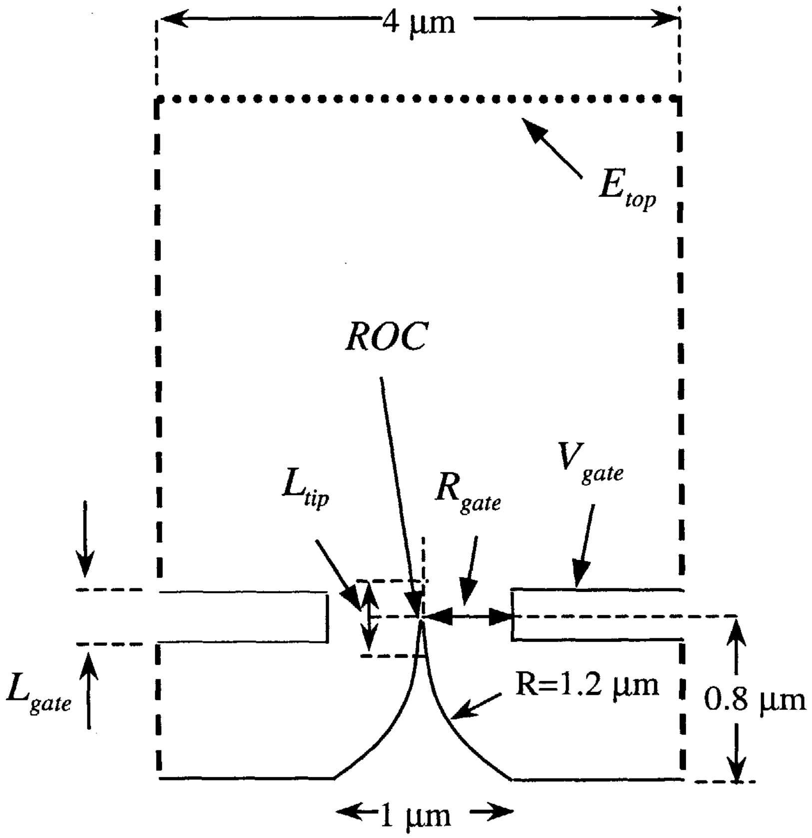


Figure 6.3.1 Basic structure of a proximity focused FED.

Note that we use the uniform electric field (E_{top}) on the top surface of the simulation model as a design parameter to account for the anode-gate gap distance (L_{gate_anode}). Assuming a constant voltage (V_{anode}) applied to the anode, L_{gate_anode} can be calculated directly from this electric field value ($E_{top} = \frac{V_{anode}}{L_{gate\_anode}}$). In order to thoroughly investigate the effect of proximity focusing, the difference between the maximum and minimum levels of the anode-gate gap distance is more than one order of magnitude (e.g., 2 cm vs. 0.5 mm). Consequently, the parameter E_{top} is also transformed to a logarithmic scale so that the five levels of the central composite design can be chosen easily.

In this experimental design, the following parameters are fixed:

- Pitch distance of each cone structure in an emitter array: 4 μm

- Anode voltage: 5000 volts
- Cone base radius: 0.5 μm
- Cone height: 0.8 μm (with zero variation in the vertical position of the tip)
- Cone shape: hyperboloid with 1.2 μm radius of curvature on the side shape

With six design parameters, a full two-level factorial design has 64 (2^6) treatments. Because adding another 13 treatments is not a big overhead for 64 treatments, we started with the parameter space for a central composite design (total treatments: $2^6 + 2 \cdot 6 + 1 = 77$). The axial-point level α is equal to $\sqrt{6} = 2.4495$. The high and low levels of parameter values for the two-level factorial design are chosen based on past operating experience (for V_{gate} and E_{top}) and reasonable fabrication technology consideration (for the rest of the parameters). The value of the center point is the average of the high and low levels. Similarly, the values for axial points can be calculated by extrapolation. Note that the center-point and axial-point values for the tip ROC and E_{top} are calculated after transferring the values of low and high levels into a logarithmic scale.

Table 6.3.1 lists the design parameters and the levels used in the central composite design. The levels in this table will be used for the polynomial fitting of the response surfaces. Except for $\ln(ROC)$ and $\ln(E_{top})$, these values are also directly used for the electrostatics and trajectory simulations. Table 6.3.2 shows the values of the tip ROC and E_{top} used for the electrostatics and trajectory calculations. The values of high and low levels of the ROC and E_{top} are first determined in this table, then transferred into a logarithmic scale and put into Table 6.3.1. After the values of the center-point and axial-point levels are calculated in Table 6.3.1, these values are transferred back into a linear scale and put into Table 6.3.2.

Parameters	-	0	+	$-\alpha$	$+\alpha$	unit
$\ln(ROC)$	1.7917	2.0471	2.3025	1.4215	2.6728	ln(nm)
V_{gate}	45	60	75	23.257	96.742	Volt
R_{gate}	0.35	0.45	0.55	0.20505	0.69495	μm
$\ln(E_{top})$	13.1	14.2736	15.4	11.5	17.1	ln(V/m)
L_{tip}	-0.1	0	0.1	-0.245	0.245	μm
L_{gate}	0.2	0.25	0.3	0.1275	0.3725	μm

Table 6.3.1 The range of FED design parameters for a central composite design ($\alpha = 2.4495$).

Parameters	-	0	+	$-\alpha$	$+\alpha$	unit
ROC	6	7.7460	10	4.14349	14.481	nm
E_{top}	500000	1581138	5000000	94237	26528946	V/m

Table 6.3.2 The values of the five levels of a central composite design for tip ROC and E_{top} in linear scale ($\alpha = 2.4495$). Note that the values of the center and axial points are calculated in the logarithmic scale then transferred back to the linear scale.

The simulated responses are the total anode current (I_a), spot size (D) in diameter from a 60x60 emitter array, and the figure of merit (FOM). The figure of merit is defined as:

$$FOM = \frac{\log(I_a)}{D} \tag{70}$$

where I_a and D have the units of pA and mm respectively. Note that during the calculation of the FOM, the anode currents that have values less than 1 pA are set equal to 1 pA to avoid a negative value of FOM.

The simulated responses are first analyzed using the linear regression technique. Full quadratic polynomial models are created for the response-surface analysis. The linear regression analysis can be done using Matlab scripts and/or JMP. The errors in the quadratic models are also studied in order to determine which coefficients in the models should be disregarded. Note that the anode current (I_a) is transferred into a logarithmic scale in the linear regression analysis because of the exponential relationship between the emission current and the gate voltage. Figure 6.3.2, Figure 6.3.3 and Figure 6.3.4 show the responses vs. the significant parameters.

For a full quadratic model with six parameters, there are a total of 28 unknown coefficients to be determined by linear regression techniques. However, some of the terms are not *statistically significant* on the response and have to be excluded to reduce the error. This can be done by studying the t-ratio of each coefficient. The definition of the t-ratio is the ratio of the coefficient estimate to its standard error, and it lists the test statistics for the hypotheses that each coefficient is zero. If the hypothesis is true, then this statistic has a Student's t distribution. If the hypothesis is rejected, it is inferred that the variation accounted for by the coefficient is significantly greater than the random variance. Looking for a t ratio greater than two in absolute value is a common rule-of-thumb for judging significance because it approximates the 0.05 significance level. The exact values for the significance level depend on the number of treatments (degrees of freedom).

Note that the parameters that are *statistically significant* on the response may not be *physically significant* on the response. The effects on responses given by statistically significant parameters are significantly higher than random or unexplained variations. In other words. statistical significance means that there is a *real* relationship between the input and the response. Among statistically significant parameters, the parameters that affect the responses substantially by small variations of the parameters are the *physically significant* parameters. Physically significant parameters are the short-cut parameters to achieve optimal performance, if they can be easy to be tuned by users.

The JMP package provides detailed reports about the t-ratio that give users information to determine which terms should be ignored (i.e., their coefficients are zero). Table 6.3.3, Table 6.3.4, Table 6.3.5, Table 6.3.6, Table 6.3.7 and Table 6.3.8 list the summaries of fit as well as the coefficient estimates for the three responses. In this work, deterministic simulations rather than physical experiments are performed, so there is no measure of random error using replicates. Instead, the lack of fit (reported as the standard error) from the experimental design responses is used in calculating the t-ratio and identifying significant parameters.

In the summary of fit table, R^2 is defined as the proportion of the variation in the response that can be attributed to the terms in the model rather than to random error or lack of fit. In other words, it is a measure of degree of fit, from 0 (no fit) to 1 (exact fit). The definition of R^2 is:

$$R^2 = \frac{SS_T - SS_E}{SS_T} = 1 - \frac{SS_E}{SS_T} \tag{71}$$

where SS_T is the total sum of squares:

$$SS_T = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \tag{72}$$

and SS_E is the residual sum of squares:

$$SS_E = \mathbf{y}'\mathbf{y} - \mathbf{c}'\mathbf{X}'\mathbf{y} \tag{73}$$

However, a large value of R^2 does not mean that the regression model is a good one because R^2 can always be increased by adding more terms to the model. The adjusted R^2 accounts for the degrees of freedom that depends on the number of treatments and number of coefficients. Since adding unnecessary terms to the regression model often decreases the value of the adjusted R^2, some users prefer using the adjusted R^2. The mathematical definition of adjusted R^2 is:

$$R^2_{adjusted} = 1 - \frac{SE_E/(n-p)}{SS_T/(n-1)} = 1 - \frac{n-1}{n-p}\left(1 - R^2\right) \tag{74}$$

where n is the number of treatments, and p is the number of coefficients.

The Prob>|t| in the coefficient estimate table lists the observed significance probability calculated from each t ratio. It is the probability of getting by chance a t ratio greater (in absolute value) than the computed value, given that the hypothesis is true. Often, a value below .05 (or sometimes .01) is interpreted as evidence that the parameter is significantly different from zero.

The response of the models are optimized using the optimization functions provided by Matlab Toolboxes. Constraints of the input parameters are needed during the optimization because the minimum/maximum values of the responses usually occur at the boundary of the design space (minimum/maximum values of the design parameters). The results of this preliminary optimization can be summarized as:

- ROC, V_{gate}, R_{gate}, E_{top}, and L_{tip} have significant effects on the responses.
- The minimum value of the tip ROC gives the maximum emission current density and FOM. This is because a smaller ROC results in higher surface electric field on the tip. However, the ROC has little effect on the spot size.
- The maximum value of E_{top} gives the minimum spot size. This is because a large value of E_{top} gives a small value of anode-gate separation (the anode voltage is fixed at 5000 volts in this design), and thus the electrons are collected by the anode with small spread. E_{top} has little effect on the emission current density compared with other statistically significant parameters. Also, this parameter has the most significant effect on the spot size and FOM among other parameters.
- The minimum value of R_{gate} gives the maximum value of current density because small R_{gate} results in large electric field around the tip. However, the smaller the R_{gate}, the larger the spot size due to higher distorsion of electric field around the gate aperture. Also, the effects of R_{gate} on the current density and the spot size are relatively small compared with other significant parameters, which can be shown in Figure 6.3.2 and Figure 6.3.3. Therefore, FOM is not sensitive to R_{gate}. Note that reducing R_{gate} is a typical way to decrease the turn-on voltage of the device. Traditionally, the turn-on voltage is defined as the gate voltage that gives rise to an emission current of 1 nA for a single emitter.
- Although a larger value of L_{tip} gives a higher value of current density, it results in a larger spot size. Since L_{tip} has a greater effect on current density than on spot size, the FOM increases as L_{tip} increases. Note that the effect on FOM is relatively small compared to the effects of other parameters.

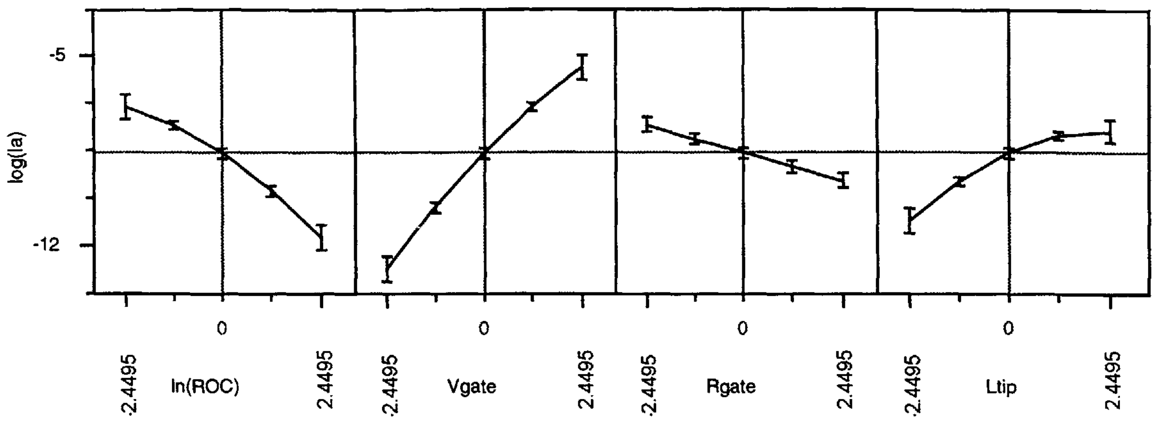


Figure 6.3.2 $\log(I_a)$ vs. its four statistically significant parameters ($\ln(ROC)$, V_{gate}, R_{gate} and L_{tip}) at their zero levels.

R^2	0.978303
Adjusted R^2	0.975389
Root Mean Square Error	0.316406
Mean of Response	-8.94258
Treatments	77

Table 6.3.3 Summary of fit for the response $\log(I_a)$.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	-8.599832	0.093861	-91.62	<.0001
$\ln(ROC)$	-0.992012	0.036294	-27.33	<.0001
V_{gate}	1.5202826	0.036294	41.89	<.0001
R_{gate}	-0.423592	0.036294	-11.67	<.0001
L_{tip}	0.6789014	0.036294	18.71	<.0001
$\ln(ROC)*\ln(ROC)$	-0.117572	0.042231	-2.78	0.0070
$V_{gate}*\ln(ROC)$	0.1143873	0.039551	2.89	0.0052
$V_{gate}*V_{gate}$	-0.089565	0.042231	-2.12	0.0376
$L_{tip}*\ln(ROC)$	-0.114059	0.039551	-2.88	0.0053
$L_{tip}*L_{tip}$	-0.140123	0.042231	-3.32	0.0015

Table 6.3.4 Coefficient estimates for the response $\log(I_a)$.

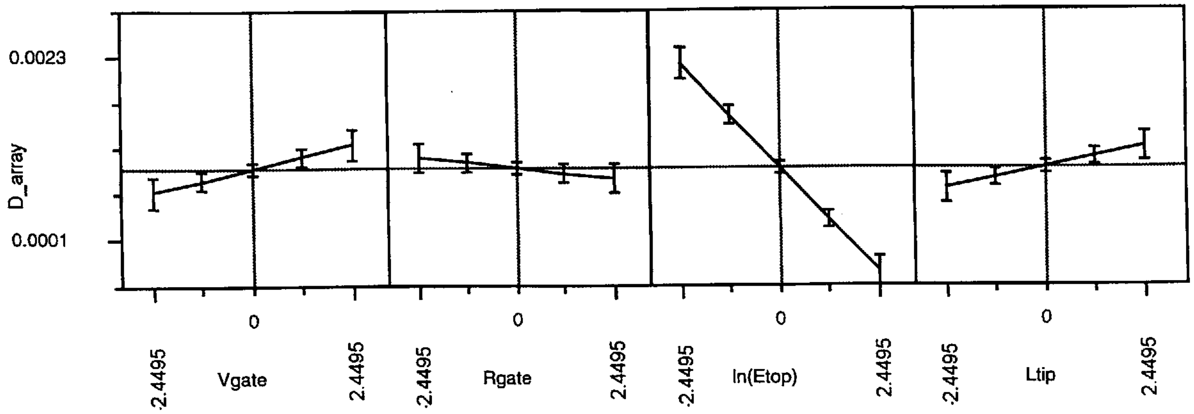


Figure 6.3.3 Spot size from a 60x60 emitter array (D) vs. its four statistically significant parameters (V_{gate}, R_{gate}, $\ln\left(E_{top}\right)$ and L_{tip}) at their zero levels.

R^2	0.785612
Adjusted R^2	0.767236
Root Mean Square Error	0.000299
Mean of Response	0.000953
Treatments	77

Table 6.3.5 Summary of fit for the response D.

Term	**Coefficients**	**Std Error**	**t Ratio**	**Prob>|t|**
Intercept	0.0009526	0.000034	27.96	<.0001
V_{gate}	0.0001164	0.000034	3.39	0.0011
R_{gate}	-0.000052	0.000034	-1.50	0.1376
$\ln\left(E_{top}\right)$	-0.00051	0.000034	-14.86	<.0001
L_{tip}	0.0001009	0.000034	2.94	0.0044
$\ln\left(E_{top}\right) * V_{gate}$	-0.000103	0.000037	-2.75	0.0076
$L_{tip} * \ln\left(E_{top}\right)$	-0.000089	0.000037	-2.39	0.0194

Table 6.3.6 Coefficient estimates for the response D.

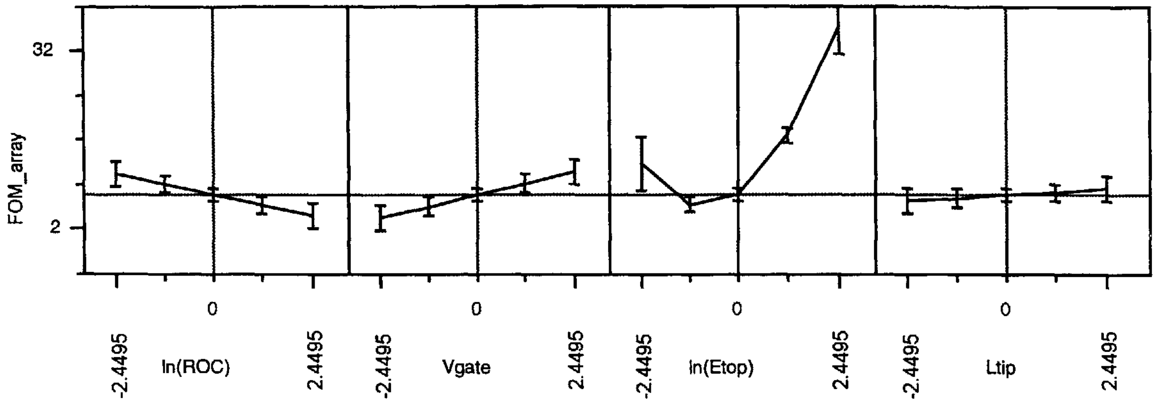


Figure 6.3.4 *FOM* vs. its four statistically significant parameters (R_{gate}, V_{gate}, $\ln(E_{top})$ and L_{tip}) at their zero levels.

R^2	0.784334
Adjusted R^2	0.762454
Root Mean Square Error	3.260556
Mean of Response	10.31255
Treatments	77

Table 6.3.7 Summary of fit for the response *FOM* .

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	7.5456453	0.554875	13.60	<.0001
$\ln(ROC)$	-1.424987	0.374011	-3.81	0.0003
V_{gate}	1.5648807	0.374011	4.18	<.0001
$\ln(E_{top})$	4.7175768	0.374011	12.61	<.0001
L_{tip}	0.4017305	0.374011	1.07	0.2865
$\ln(E_{top})*\ln(ROC))$	-0.77137	0.407569	-1.89	0.0626
$\ln(E_{top})*V_{gate}$	1.2887801	0.407569	3.16	0.0023
$\ln(E_{top})*\ln(E_{top})$	2.8033095	0.417512	6.71	<.0001

Table 6.3.8 Coefficient estimates for the response *FOM* .

6.3.2 The Second-Phase Analysis : Four Parameters

The preliminary study (six parameters) of response surfaces and optimization for a proximity-focused FED design did not give very precise prediction for the responses, which is indicated by:

- large confidence intervals in some plots of responses vs. parameters
- values of R^2 and adjusted R^2 that are not satisfactory (relatively small)

This inaccuracy is predictable because of the following reasons:

- a quadratic polynomial model might not be sufficient or adequate to represent the responses
- the wide ranges of the parameter values usually give worse fitting than narrow ranges
- parameters that do not have a statistically significant effect on the responses nevertheless contribute small variations that appear as apparent errors

As a result, a second-phase design of experiment is implemented and performed based on the experience of this preliminary study. In the new design, tip-ROC, V_{gate}, E_{top} and L_{tip} are chosen as the design parameters. In the following paragraphs, the reasons why these four parameters are chosen and why the other two parameters are excluded will be discussed. In addition to reducing the number of parameters, the second design is centered near an expected optimal point, based on the initial design.

The common characteristic of these four parameters is that all of them have significant effects on at least one of the responses. Although the gate aperture R_{gate} also has significant impact on the emission current, we exclude it because:

- We always need the smallest R_{gate} to obtain the lowest turn-on voltage for FEDs [93][94]. However, the value of R_{gate} depends on process technology/facility available. Once the process is determined, the smallest R_{gate} is determined.
- In our fabrication process, R_{gate} is easy to control so that its process variation is small.

On the other hand, although we know the smallest tip-ROC always provides best performance, it is one of the most difficult parameters to control in the fabrication process. We include it in the second-phase experimental design to study the response caused by its variation. The new range of values of this parameter is narrowed down to the best sharpness that can be achieved by the current process facility.

Also, the tip-ROC is kept in the linear scale in this second-phase analysis due to the following reasons:

- Narrower range of values. It is not necessary to transform ROC to a logarithmic scale to avoid very small value for the negative axial point in the central composite design.
- Keep the compatibility of the functional form of the Fowler-Nordheim equation for building a macromodel.

Not only does L_{tip} have significant effect on the response, but it is also difficult to control in the fabrication process. It is certainly chosen for the second-phase experimental design. The range of parameters is narrowed to the practical variation caused by fabrication process.

E_{top} depends on the gap between the anode and gate, and is the most significant parameter for beam focusing. With fixed anode voltage, higher E_{top} represents smaller anode-gate distance, and consequently smaller spot size. However, the minimum anode-gate distance depends on packaging technology, such as spacer strength and vacuum level. E_{top} is chosen so that the relationship between various packaging technologies and spot sizes can be studied.

V_{gate} is the only *operational parameter* in the proximity focused device design, which not only switches on/off the emitter, but also controls the gray-scale/color of the image on the phosphor anode. Its characteristic has to be studied for the operation of the device. Its new range of values does not differ much from its previous range, but is just narrowed to a range that covers the turn-on voltage and a reasonable operational range of voltage. A range of the operational voltage is chosen so that:

- the range of the anode current always gives reasonable gray scale

- an inexpensive driving circuit will be sufficient to generate required switching gate voltages [95].

The first-phase design also shows that the gate thickness L_{gate} does not effect the responses. In the second-phase design, L_{gate} is fixed at 200 nm.

Table 6.3.9 and Table 6.3.10 list the four design parameters chosen for the second-phase experimental design and their five levels. The range of the parameters are narrower and around the range where the optimal performance occurs. Since there are four design parameters, α is equal to 2 ($\sqrt{4}$). There are a total of 25 treatments in the central composite design ($2^4 + 2 \cdot 4 + 1$).

Parameters	**-**	**0**	**+**	**$-\alpha$**	**$+\alpha$**	**unit**
ROC	7	8	9	6	10	nm
V_{gate}	50	60	70	40	80	Volt
$\ln(E_{top})$	13.8	14.6	15.4	13.0	16.2	ln(V/m)
L_{tip}	-0.05	0	0.05	-0.1	0.1	μm

Table 6.3.9 The range of FED design parameters for a central composite design ($\alpha = 2$).

Parameters	**-**	**0**	**+**	**$-\alpha$**	**$+\alpha$**	**unit**
E_{top}	1000000	2236068	5000000	447213	11180339	V/m

Table 6.3.10 The values of the five levels of a central composite design for E_{top} in a linear scale ($\alpha = 2$). Note that the values of the center and axial points are calculated in the logarithm scale then transferred back to the linear scale.

Figure 6.3.5, Figure 6.3.6 and Figure 6.3.7 show the response vs. parameters of the second-phase experimental design. Table 6.3.11, Table 6.3.12, Table 6.3.13, Table 6.3.14, Table 6.3.15 and Table 6.3.16 show the statistical results and model estimates for the response surfaces. With fewer parameters and narrower ranges of parameters, the regression models are much better than those from the previous experimental design. Note that the regression model for the spot size D includes a cubic term of $\ln\left(E_{top}\right)$. Without this term, the adjusted R^2 is less than 0.96.

The summary of this refined second-phase experimental design is:

- The regression models provide accurate estimate of the responses.
- E_{top} (anode-gate distance) is the most critical parameter (physically significant parameter) that affects the performance of the device.
- Although L_{tip} is statistically significant in the model, its effect on the responses are relatively insignificant compared to other parameters in the response of the figure of merit.
- A cubic term of $\ln\left(E_{top}\right)$ is used for a better regression model.

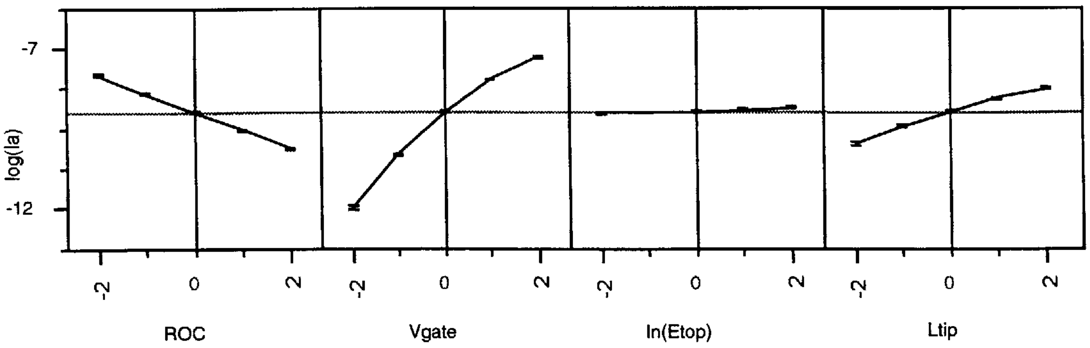


Figure 6.3.5 $\log(I_a)$ vs. its four statistically significant parameters (ROC, V_{gate}, $\ln(E_{top})$ and L_{tip}) at their zero levels.

R^2	0.999826
Adjusted R^2	0.999678
Root Mean Square Error	0.024832
Mean of Response	-9.17019
Treatments	25

Table 6.3.11 Summary of fit for the response $\log(I_a)$.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	-8.983569	0.013467	-667.1	<.0001
ROC	-0.568293	0.005069	-112.1	<.0001
V_{gate}	1.1719373	0.005069	231.20	<.0001
$\ln(E_{top})$	0.0393035	0.005069	7.75	<.0001
L_{tip}	0.4283518	0.005069	84.51	<.0001
$V_{gate} * ROC$	0.095145	0.006208	15.33	<.0001
$V_{gate} * V_{gate}$	-0.163852	0.005634	-29.08	<.0001
$\ln(E_{top}) * V_{gate}$	-0.010931	0.006208	-1.76	0.1018
$\ln(E_{top}) * \ln(E_{top})$	0.0075931	0.005634	1.35	0.2007
$L_{tip} * ROC$	0.0158366	0.006208	2.55	0.0241
$L_{tip} * V_{gate}$	-0.057985	0.006208	-9.34	<.0001
$L_{tip} * L_{tip}$	-0.038136	0.005634	-6.77	<.0001

Table 6.3.12 Coefficient estimates for the response $\log(I_a)$.

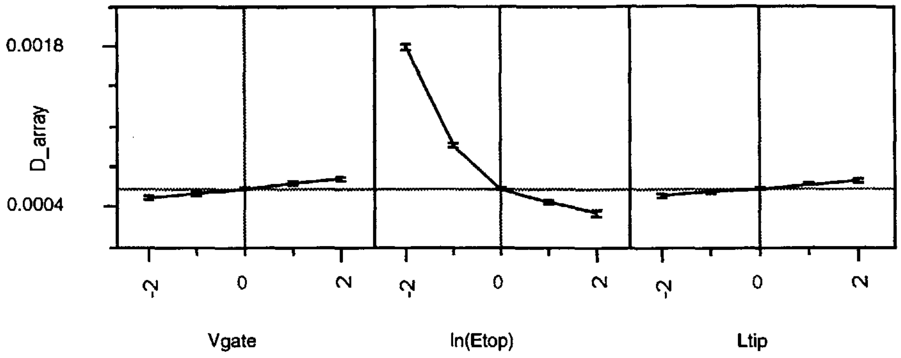


Figure 6.3.6 Spot size from a 60x60 emitter array (D) vs. its statistically significant parameters (V_{gate}, $\ln(E_{top})$ and L_{tip}) at their zero levels.

R^2	0.998192
Adjusted R^2	0.997448
Root Mean Square Error	0.000016
Mean of Response	0.000679
Treatments	25

Table 6.3.13 Summary of fit for the response D.

Term	Coefficients		Std Error	T Ratio	Prob>|t|
Intercept	0.0005567	c_0	0.000005	122.82	<.0001
V_{gate}	0.0000427	c_1	0.000003	12.80	<.0001
$\ln(E_{top})$	-0.000206	c_2	0.000006	-35.69	<.0001
L_{tip}	0.0000302	c_3	0.000003	9.05	<.0001
$\ln(E_{top}) * V_{gate}$	-0.00003	c_4	0.000004	-7.27	<.0001
$\ln(E_{top})*\ln(E_{top})$	0.0001274	c_5	0.000003	38.93	<.0001
$L_{tip} * \ln(E_{top})$	-0.000023	c_6	0.000004	-5.74	<.0001
$\ln(E_{top})*\ln(E_{top})*\ln(E_{top})$	-0.000039	c_7	0.000002	-16.56	<.0001

Table 6.3.14 Coefficient estimates for the response D.

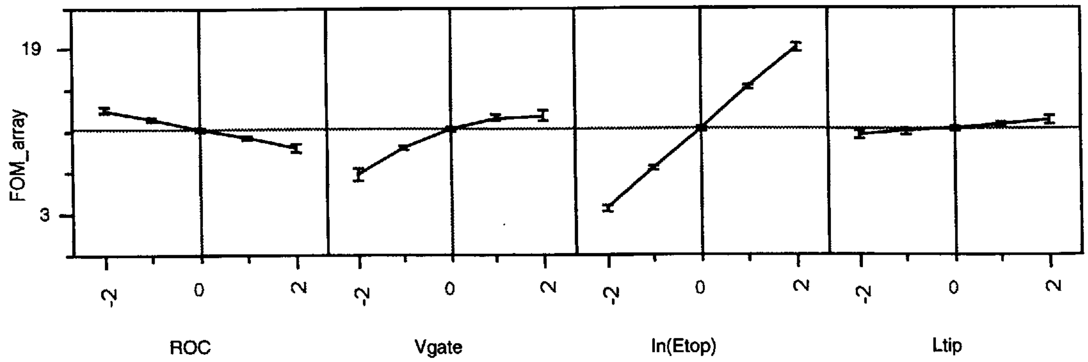


Figure 6.3.7 *FOM* vs. its four statistically significant parameters (ROC, V_{gate}, $\ln(E_{top})$, and L_{tip}) at their zero levels.

R^2	0.996025
Adjusted R^2	0.993185
Root Mean Square Error	0.360637
Mean of Response	10.89188
Treatments	25

Table 6.3.15 Summary of fit for the response *FOM*.

Term	**Coefficients**	**Std Error**	**t Ratio**	**Prob>|t|**
Intercept	11.288143	0.100023	112.86	<.0001
ROC	-0.912151	0.073615	-12.39	<.0001
V_{gate}	1.4318121	0.073615	19.45	<.0001
$\ln(E_{top})$	3.9165695	0.073615	53.20	<.0001
L_{tip}	0.3375702	0.073615	4.59	0.0004
$V_{gate}*ROC$	0.1629026	0.090159	1.81	0.0923
$V_{gate}*V_{gate}$	-0.412773	0.072185	-5.72	<.0001
$\ln(E_{top})*ROC$	-0.293976	0.090159	-3.26	0.0057
$\ln(E_{top})*V_{gate}$	0.70921	0.090159	7.87	<.0001
$L_{tip}*ROC$	0.1348215	0.090159	1.50	0.1570
$L_{tip}*\ln(E_{top})$	0.3314346	0.090159	3.68	0.0025

Table 6.3.16 Coefficient estimates for the response *FOM*.

6.3.3 Macromodels

The next step of modeling is to create the macromodels of the device. The regression models by themselves are macromodels of the device. For example, the macromodel of the spot size from a 60x60 emitter array can be created using the results in Table 6.3.14.

$$D = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_1 x_2 + c_5 x_2^2 + c_6 x_2 x_3 + c_7 x_2^3 \quad (75)$$

where $c_0 \ldots\ldots c_7$ are listed in Table 6.3.14, and

$$x_1 = \frac{\ln\left(V_{gate}\right) - \left(\text{'0' level of } \ln\left(V_{gate}\right)\right)}{\left(\text{'+' level of } \ln\left(V_{gate}\right)\right) - \left(\text{'0' level of } \ln\left(V_{gate}\right)\right)} \quad (76)$$

$$x_2 = \frac{\ln\left(E_{top}\right) - \left(\text{'0' level of } \ln\left(E_{top}\right)\right)}{\left(\text{'+' level of } \ln\left(E_{top}\right)\right) - \left(\text{'0' level of } \ln\left(E_{top}\right)\right)} \quad (77)$$

$$x_3 = \frac{\ln\left(L_{tip}\right) - \left(\text{'0' level of } \ln\left(L_{tip}\right)\right)}{\left(\text{'+' level of } \ln\left(L_{tip}\right)\right) - \left(\text{'0' level of } \ln\left(L_{tip}\right)\right)} \quad (78)$$

6.4 Integrated-Focus-Electrode FED Design Example

6.4.1 Preliminary Analysis : Six Parameters

In this section, examples of experimental design are demonstrated *for integrated-focus-electrode* (IFE) FEDs. The IFE FEDs has a second electrode that converges electron beams and reduces spot sizes. The main advantage of the IFE FEDs over proximity FEDs is that the IFE FEDs have the flexibility of using high voltage phosphors while achieving reasonably small spot sizes. We use a central composite design as the first design strategy to filter out the significant parameters. The preliminary design parameters for the IFE FED are:

- the tip radius of curvature (ROC)
- the gate voltage (V_{gate})
- the focus gate (electrode) voltage (V_{focus})
- the gate aperture radius (R_{gate})
- the difference of the aperture radii between the gate and the focus electrode (dR_{focus})
- the distance between the upper surface of the gate and the lower surface of the focus electrode (L_{gf})

The preliminary design parameters are listed in Table 6.4.1 and Table 6.4.2. The schematic of an IFE FED and its design parameters are shown in Figure 6.4.1. The design parameters and the ranges of their values listed in Table 6.4.1 are chosen by intuition, past experimental experience, and the experience from the previous examples (design of proximity focused FEDs). Similar to the proximity focusing cases, the tip radius of curvature (ROC) is transformed into a logarithmic scale for better fitting in the regression analysis. Note that in this experimental design L_{tip} is not selected as one of the design parameters. Although L_{tip} has statistical significance on responses, its physical significance on FOM is relatively small.

In this experimental design, the fixed parameters are:

- Pitch distance of each cone structure in an emitter array: 4 μm

- Anode voltage: 5000 volts
- Gate-anode distance: 1 cm
- Cone base radius: 0.5 μm
- Cone height: 0.8 μm (with zero variation in the vertical position of the tip)
- Cone shape: hyperboloid with 1.2 μm radius of curvature on the side shape
- Gate thickness: 0.2 μm
- Focus electrode thickness: 0.5 μm

A relatively large value of the gate-anode distance (1 cm) is chosen so that the minimum spot size predicted by this design will be the upper bound value, since the minimum achievable spot size can be further reduced by decreasing the gate-anode distance. Note that in recent designs of FEDs using high voltage phosphor, the anode-gate distance has been successfully reduced to the 1 to 1.25 mm range.

There are a total of 77 treatments in this experimental design, and the level of the axial point is equal to 2.4495 ($\sqrt{6}$).

Parameters	-	0	+	-α	+α	**Unit**
$\ln(ROC)$	1.79176	2.04717	2.30259	1.42154	2.67280	ln(nm)
V_{gate}	45	60	75	23.2575	96.7425	Volt
V_{focus}	25	32.5	40	14.129	50.871	Volt
R_{gate}	0.35	0.45	0.55	0.20505	0.69495	μm
dR_{focus}	0.15	0.25	0.35	0.0051	0.495	μm
L_{gf}	0.4	0.6	0.8	0.1101	1.0899	μm

Table 6.4.1 The range of FED design parameters for a central composite design ($\alpha = 2.4495$).

Parameters	-	0	+	-α	+α	**Unit**
ROC	6	7.7460	10	4.1435	14.4805	nm

Table 6.4.2 The values of the five levels of a central composite design for tip *ROC* in linear scale ($\alpha = 2.4495$). Note that the values of the center and axial points are calculated in the logarithm scale then transferred back to the linear scale.

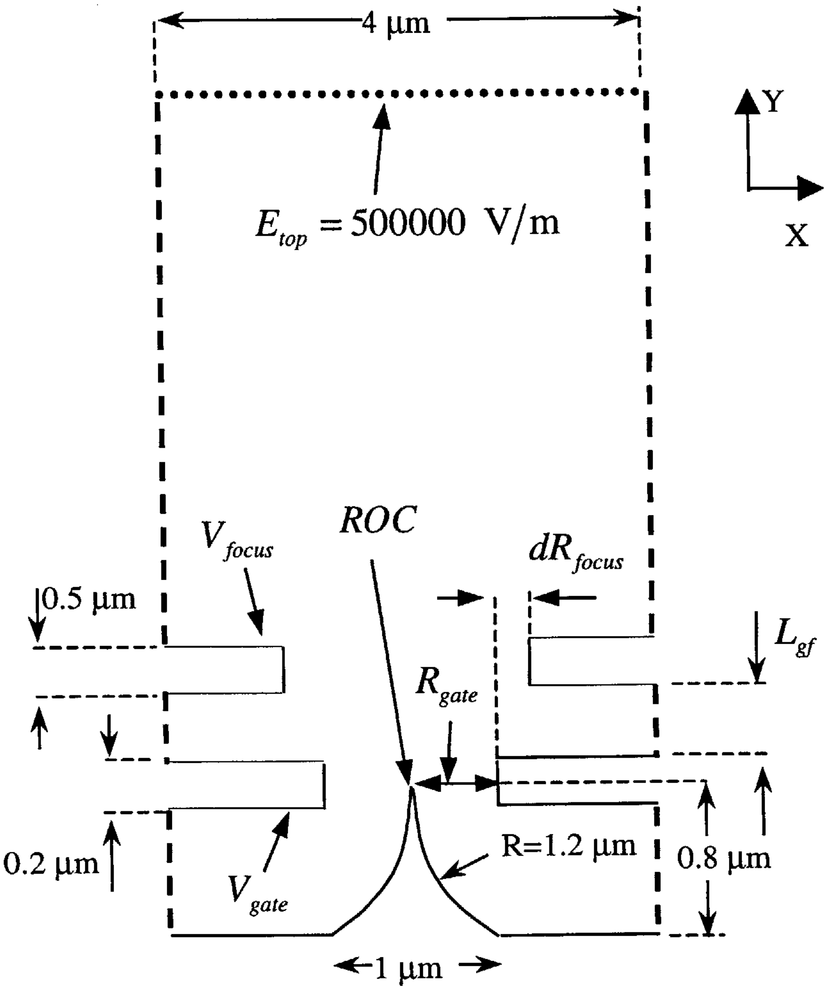


Figure 6.4.1 Basic structure of an integrated-focus-electrode (IFE) FED.

Figure 6.4.2, Figure 6.4.3, and Figure 6.4.4 are the response plots of emission current, spot size and figure of merit vs. their significant parameters. The statistical data and the coefficient estimates of these plots are listed in Table 6.4.3, Table 6.4.4, Table 6.4.5, Table 6.4.6, Table 6.4.7, and Table 6.4.8. In some treatments, the potential drop between the gate and the focusing gate is so large that all the emitted electrons either are deflected back to the cathode or are intercepted by those two gates. Since there is no anode current in these cases, the figure of merit and the spot size are not defined. We exclude these treatments from the regression analysis.

All of the six design parameters have statistically significant effects on the responses of total anode current (I_a), spot size of a 60x60 array (D), and figure of merit (FOM). The regression model of the anode current gives excellent fitting, but the models for the spot size and the figure of merit are relatively inaccurate. Nevertheless, the constrained optimization of the regression models performed by the Matlab functions gives information about the parameter ranges for optimal performance, which will be used for the second phase experimental design.

The IFE FEDs have two operational parameters: the gate voltage (V_{gate}) and the focus voltage (V_{focus}). The optimization also explores the range of voltages that give the desired anode current and spot size. The constraints applied to the optimization are no longer the boundary of the parameters. Instead, the constraints are the minimum required anode current and the maximum tolerable spot size.

The optimization process for the IFE FEDs not only gives information about the desired dimensions of the device, but also provides the information for choosing the operating ranges of the operational parameters.

Note that Figure 6.4.3 shows that the spot size decreases as V_{gate} increases. However, in the proximity focusing example, Figure 6.3.3 and Figure 6.3.6 indicate that the spot size increases with the V_{gate}. The contradictory behavior in Figure 6.4.3 is because an increase in V_{gate} increases the voltage difference between V_{gate} and V_{focus}, and thus increases the focusing effect.

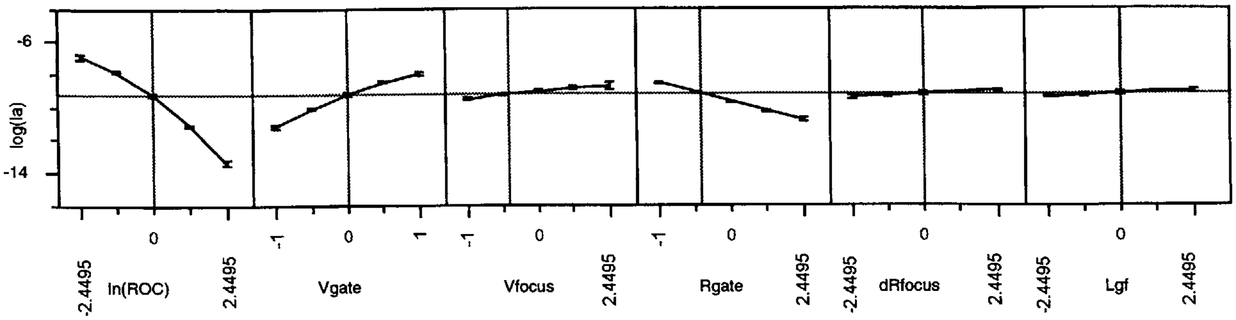


Figure 6.4.2 $\log(I_a)$ vs. its six significant parameters ($\ln(ROC)$, V_{gate}, V_{focus}, R_{gate}, dR_{focus} and L_{gf}) at their zero levels.

R^2	0.998851
Adjusted R^2	0.998277
Root Mean Square Error	0.091444
Mean of Response	-9.97717
Treatments	61

Table 6.4.3 Summary of fit for the response $\log(I_a)$.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	-9.282474	0.037833	-245.4	<.0001
$\ln(ROC)$	-1.322477	0.013786	-95.93	<.0001
V_{gate}	1.5903174	0.020221	78.65	<.0001
V_{focus}	0.2497762	0.020221	12.35	<.0001
R_{gate}	-0.65171	0.015033	-43.35	<.0001
dR_{focus}	0.0797127	0.01222	6.52	<.0001
L_{gf}	0.0639624	0.011731	5.45	<.0001
$\ln(ROC)*\ln(ROC)$	-0.148197	0.012486	-11.87	<.0001
$V_{gate}*\ln(ROC)$	0.2281521	0.015604	14.62	<.0001
$V_{gate}*V_{gate}$	-0.31655	0.034613	-9.15	<.0001
$V_{focus}*\ln(ROC)$	0.0791646	0.015604	5.07	<.0001
$V_{focus}*V_{gate}$	-0.04949	0.020221	-2.45	0.0189
$V_{focus}*V_{focus}$	-0.024182	0.018273	-1.32	0.1932
$R_{gate}*\ln(ROC)$	-0.109024	0.013199	-8.26	<.0001
$R_{gate}*V_{gate}$	0.0989056	0.01589	6.22	<.0001
$R_{gate}*V_{focus}$	0.0706751	0.01589	4.45	<.0001
$dR_{focus}*V_{gate}$	0.0234877	0.013641	1.72	0.0928
$dR_{focus}*V_{focus}$	-0.025775	0.013641	-1.89	0.0661
$dR_{focus}*R_{gate}$	-0.017726	0.012871	-1.38	0.1761
$L_{gf}*V_{focus}$	-0.035969	0.013022	-2.76	0.0086
$L_{gf}*R_{gate}$	0.046573	0.012768	3.65	0.0008

Table 6.4.4 Coefficient estimates for the response $\log(I_a)$.

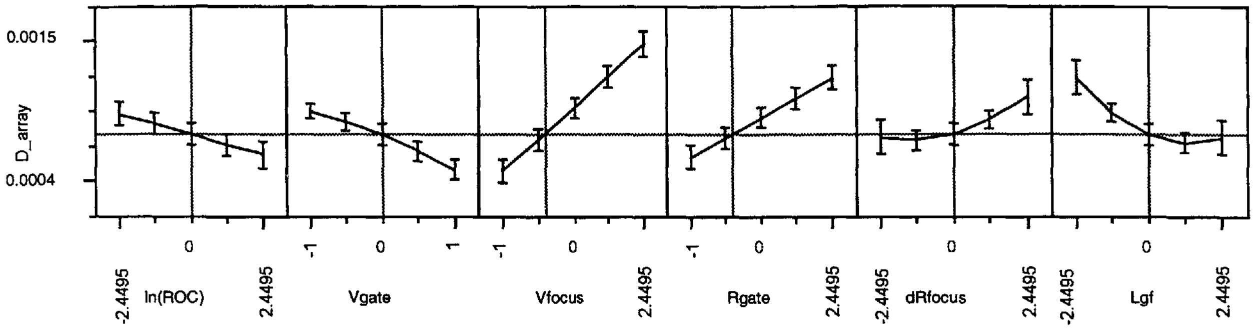


Figure 6.4.3 Spot size from a 60x60 emitter array (D) vs. its six statistically significant parameters ($\ln(ROC)$, V_{gate}, V_{focus}, R_{gate}, dR_{focus} and L_{gf}) at their zero levels.

R^2	0.939612
Adjusted R^2	0.915737
Root Mean Square Error	0.000087
Mean of Response	0.000945
Treatments	61

Table 6.4.5 Summary of fit for the response D.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	0.0007665	0.000041	18.80	<.0001
$\ln(ROC)$	-0.000066	0.000011	-5.90	<.0001
V_{gate}	-0.000231	0.000017	-13.55	<.0001
V_{focus}	0.0002865	0.000016	17.70	<.0001
R_{gate}	0.0001798	0.000014	12.57	<.0001
dR_{focus}	0.0000653	0.000011	5.87	<.0001
L_{gf}	-0.000095	0.000011	-8.58	<.0001
$V_{gate}*V_{gate}$	-0.000048	0.000033	-1.47	0.1487
$V_{focus}*V_{gate}$	0.0000847	0.000017	4.97	<.0001
$R_{gate}*\ln(ROC)$	0.0000176	0.000012	1.41	0.1654
$R_{gate}*V_{gate}$	0.0000779	0.000015	5.21	<.0001
$R_{gate}*V_{focus}$	-0.000044	0.000015	-2.96	0.0050
$dR_{focus}*V_{focus}$	0.0000186	0.000012	1.51	0.1393
$dR_{focus}*R_{gate}$	-0.000025	0.000012	-2.04	0.0472
$dR_{focus}*dR_{focus}$	0.0000222	0.000012	1.81	0.0770
$L_{gf}*V_{focus}$	-0.000028	0.000012	-2.29	0.0273
$L_{gf}*R_{gate}$	0.0000215	0.000012	1.78	0.0827
$L_{gf}*L_{gf}$	0.0000347	0.000012	2.83	0.0070

Table 6.4.6 Coefficient estimates for the response D.

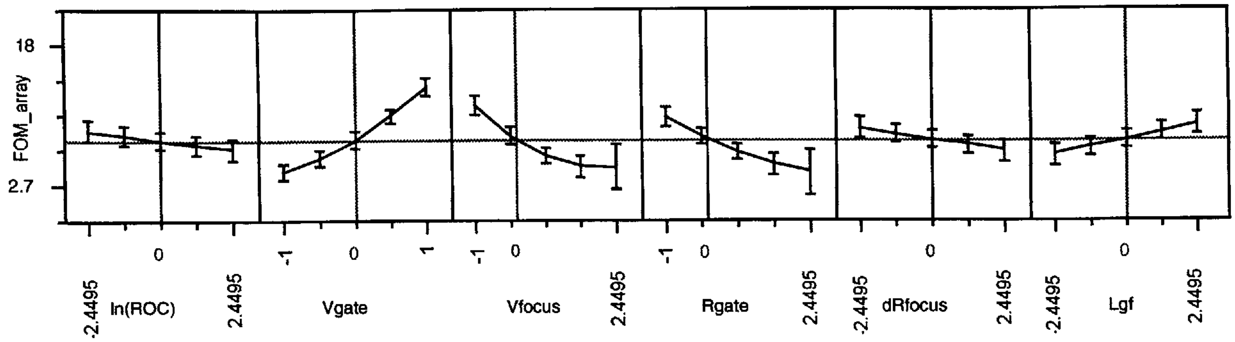


Figure 6.4.4 FOM vs. its six statistically significant parameters ($\ln(ROC)$, V_{gate}, V_{focus}, R_{gate}, dR_{focus} and L_{gf}) at their zero levels.

R^2	0.929173
Adjusted R^2	0.901172
Root Mean Square Error	1.196288
Mean of Response	7.076858
Treatments	61

Table 6.4.7 Summary of fit for the response FOM.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	7.5872147	0.452154	16.78	<.0001
$\ln(ROC)$	-0.428051	0.161798	-2.65	0.0113
V_{gate}	4.6148515	0.246322	18.74	<.0001
V_{focus}	-2.962265	0.246322	-12.03	<.0001
R_{gate}	-2.176964	0.211476	-10.29	<.0001
dR_{focus}	-0.491792	0.149897	-3.28	0.0021
L_{gf}	0.6932587	0.153984	4.50	<.0001
$V_{gate}*\ln(ROC)$	0.3823991	0.180896	2.11	0.0404
$V_{gate}*V_{gate}$	1.0278382	0.454448	2.26	0.0288
$V_{focus}*V_{gate}$	-2.013324	0.246322	-8.17	<.0001
$V_{focus}*V_{focus}$	0.7064485	0.235677	3.00	0.0045
$R_{gate}*V_{gate}$	-1.137996	0.211476	-5.38	<.0001
$R_{gate}*V_{focus}$	0.5054388	0.211476	2.39	0.0213
$R_{gate}*R_{gate}$	0.2968865	0.229967	1.29	0.2036
$dR_{focus}*R_{gate}$	0.3042587	0.166295	1.83	0.0742
$L_{gf}*\ln(ROC)$	0.2029518	0.167375	1.21	0.2319
$L_{gf}*V_{focus}$	0.3243146	0.171008	1.90	0.0646
$L_{gf}*R_{gate}$	-0.524122	0.167375	-3.13	0.0031

Table 6.4.8 Coefficient estimates for the response FOM.

6.4.2 The Second-Phase Analysis : Four Parameters

The six-parameter experimental design is not able to provide a good data set that builds good regression models for responses. Therefore, a second-phase design with fewer parameters and narrower ranges of parameters is implemented. There are four parameters in the second-phase experimental design: ROC, V_{gate}, *ratio* (the ratio of V_{focus} to V_{gate}) and L_{gf}. Some of these parameters are selected from the previous parameter set, and one of them is a new parameter. In the following paragraphs, we will discuss the criteria of picking these parameters from the previous set, as well as the definition of the new parameter.

As discussed in the previous section, the gate aperture radius R_{gate} is strongly dependent on process facility/technology, so the smallest value of aperture radius that is compatible with our process is chosen (0.5 μm) for the second-phase experimental design.

The regression models show that the dR_{focus} (the difference between gate aperture radius and focus aperture radius) has the following properties:

- As dR_{focus} decreases, D decreases and FOM increases.
- When dR_{focus} is in the range between the center point and the negative axial point, its effect on the response becomes flat.
- As dR_{focus} is close to its minimum value, the leakage current increases.

Consequently, for the second-phase experimental design, the value of dR_{focus} is set at the center point (i.e., R_{focus} is 0.75 μm) in the first-phase design. This value not only gives the smallest spot size and the maximum FOM, but also prevents the device from generating too much leakage current.

According to the previous results of six design parameters, if the potential drop between the gate and the focus is too large, all the emitted electrons are deflected toward the emitters and the anode current becomes zero. In order to avoid this condition, the ratio of the focusing voltage to the gate voltage, denoted as $ratio = \frac{V_{focus}}{V_{gate}}$, is used as a new design parameter. Based on the optimization results of the previous six-parameter design, 0.40 is chosen as the value for the negative axial point of the ratio in order to achieve a small spot size without deflecting all the emitted electrons towards emitters (zero anode current).

A larger L_{gf} gives better focusing effect and less negative effect on the emission current. Also, its effect on the spot size becomes flat when L_{gf} increases. It is included as one of the design parameters to study its effect on focusing. A value of L_{gf} larger than 500 nm is preferred because a high-voltage difference between the gate and the focus may cause a breakdown through a thinner oxide layer.

Finally, since the tip *ROC* is difficult to control, it is included as one of the design parameters in order to investigate its effect. Its range of values is around the practical range that can be achieved by the process facility. Also, similar to the reasons described in the previous section, this parameter is used without transforming it into a logarithmic scale.

Table 6.4.9 lists the new set of design parameters as well as their ranges of values. There are a total of 25 treatments, and α is equal to 2.

Parameters	-	0	+	-α	+α	**unit**
ROC	7	8	9	6	10	nm
V_{gate}	50	60	70	40	80	Volt
V_{focus}/V_{gate}	0.55	0.70	0.85	0.40	1.00	Volt/Volt
L_{gf}	0.7	0.8	0.9	0.6	1	μm

Table 6.4.9 The range of FED design parameters for a central composite design ($\alpha = 2$).

The simulation results build excellent regression models of responses. Figure 6.4.5, Figure 6.4.6, and Figure 6.4.7 are the response plots vs. their statistically significant parameters, and Table 6.4.10, Table 6.4.11, Table 6.4.12, Table 6.4.13, Table 6.4.14, and Table 6.4.15 are the summaries of fit and coefficient estimates for each response. The summary is:

- Excellent fit of regression models
  - The ratio of V_{focus} to V_{gate} is the most physically significant parameter on the spot size, while all other parameters are almost insensitive. A decrease in the ratio is equivalent to an increase in the voltage difference between the gate and the focus, and hence the spot size decreases.
  - The ratio of V_{focus} to V_{gate} is the most physically significant parameter on the figure of merit.
  - Higher order terms (third and/or fourth order) L_{gf} and *ratio* are needed to build good regression models for D and FOM.

Note that the spot size is almost insensitive to all parameters except the ratio of V_{focus} to V_{gate}. This is a very important observation for IFE FEDs. For example, as long as the ratio of V_{focus} to V_{gate} is fixed, V_{gate} can be varied freely to adjust the grayscale/color of the spot on a screen, without changing spot size significantly. This characteristic can be used as a guideline for the operation of IFE FEDs.

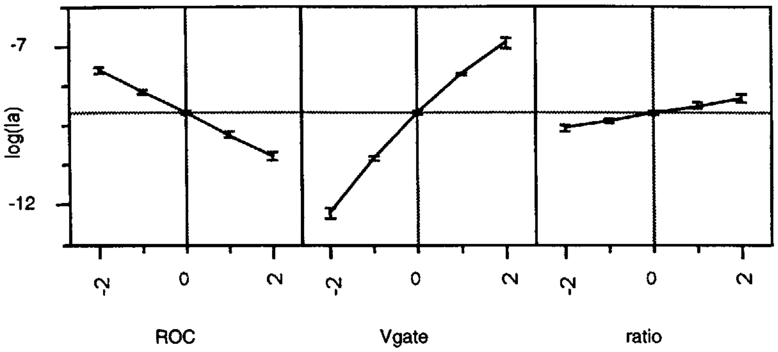


Figure 6.4.5 $\log(I_a)$ vs. its three statistically significant parameters (ROC, V_{gate}, and $V_{focus}/V_{gate}(= ratio)$) at their zero levels.

R^2	0.996546
Adjusted R^2	0.995638
Root Mean Square Error	0.08912
Mean of Response	-9.55022
Treatments	25

Table 6.4.10 Summary of fit for the response $\log(I_a)$.

Term	**Coefficients**	**Std Error**	**t Ratio**	**Prob>|t|**
Intercept	-9.452862	0.024717	-382.4	<.0001
ROC	-0.602908	0.018192	-33.14	<.0001
V_{gate}	1.1795592	0.018192	64.84	<.0001
$ratio$	0.2050547	0.018192	11.27	<.0001
$V_{gate}*ROC$	0.1007799	0.02228	4.52	0.0002
$V_{gate}*V_{gate}$	-0.101417	0.017838	-5.69	<.0001
$V_{gate}*\ln(ROC)$	0.1079529	0.032677	3.30	0.0045
$V_{gate}*V_{gate}$	-0.086997	0.027269	-3.19	0.0057
$dV_{focus}*V_{gate}$	-0.040309	0.032677	-1.23	0.2352

Table 6.4.11 Coefficient estimates for the response $\log(I_a)$.

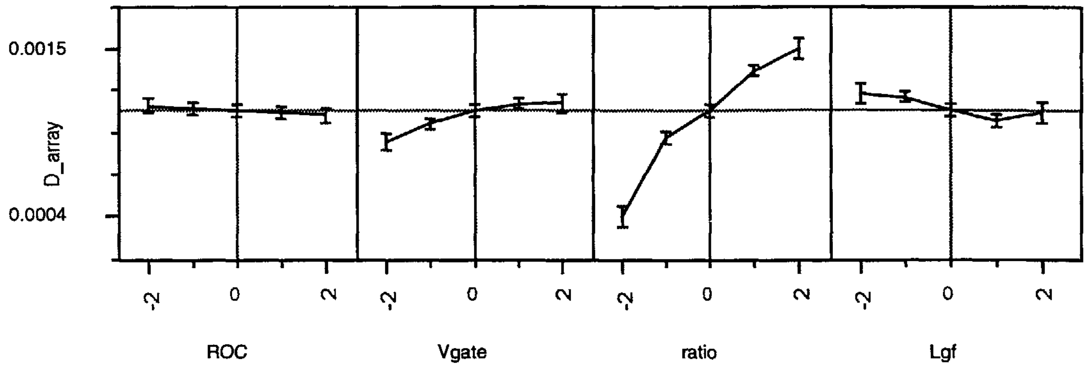


Figure 6.4.6 Spot size from a 60x60 emitter array (D) vs. its four statistically significant parameters (ROC, V_{gate}, $V_{focus}/V_{gate}(= ratio)$ and L_{gf}) at their zero levels.

R^2	0.995122
Adjusted R^2	0.986991
Root Mean Square Error	0.00003
Mean of Response	0.0011
Treatments	25

Table 6.4.12 Summary of fit for the response D.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	0.0010917	0.000018	62.33	<.0001
ROC	-0.000016	0.000006	-2.52	0.0326
V_{gate}	0.0000635	0.000006	10.26	<.0001
$ratio$	0.0002021	0.000011	18.84	<.0001
L_{gf}	-0.000092	0.000011	-8.55	<.0001
$V_{gate}*V_{gate}$	-0.00002	0.000007	-2.86	0.0188
$ratio*ROC$	0.0000203	0.000008	2.68	0.0253
$ratio*V_{gate}$	0.0000172	0.000008	2.27	0.0497
$ratio*ratio$	0.0000694	0.000018	3.93	0.0035
$L_{gf}*ROC$	0.0000234	0.000008	3.09	0.0129
$L_{gf}*V_{gate}$	-0.000011	0.000008	-1.44	0.1831
$L_{gf}*ratio$	-0.00002	0.000008	-2.68	0.0253
$L_{gf}*L_{gf}$	0.0000115	0.000007	1.66	0.1323
$ratio*ratio*ratio$	0.0000182	0.000004	4.16	0.0024
$ratio*ratio*ratio*ratio$	-0.000026	0.000004	-6.16	0.0002
$L_{gf}*L_{gf}*L_{gf}$	0.0000151	0.000004	3.45	0.0073

Table 6.4.13 Coefficient estimates for the response D.

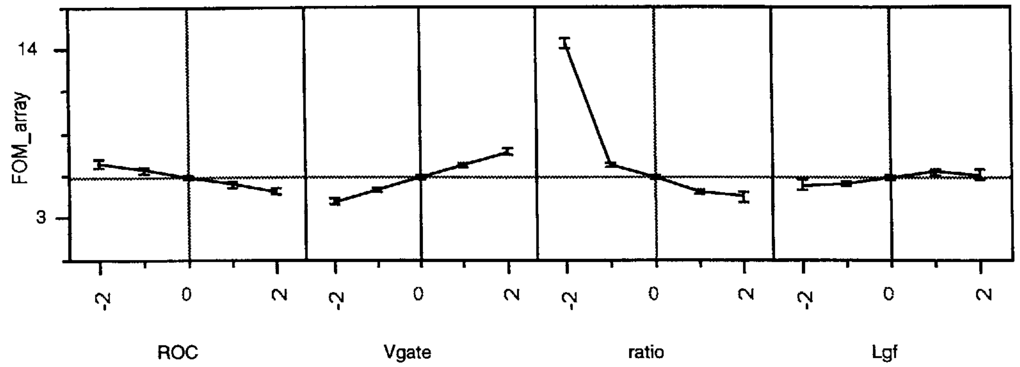


Figure 6.4.7 *FOM* vs. its four statistically significant parameters (ROC, V_{gate}, V_{focus}/V_{gate} and L_{gf}) at their zero levels.

R^2	0.997545
Adjusted R^2	0.994644
Root Mean Square Error	0.159937
Mean of Response	5.787048
Treatments	25

Table 6.4.14 Summary of fit for the response *FOM*.

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	5.6077092	0.071526	78.40	<.0001
ROC	-0.469324	0.032647	-14.38	<.0001
V_{gate}	0.8095038	0.032647	24.80	<.0001
$ratio$	-0.329335	0.056546	-5.82	0.0001
L_{gf}	0.4634217	0.056546	8.20	<.0001
$V_{gate}*ROC$	0.1521167	0.039984	3.80	0.0029
$ratio*V_{gate}$	-0.199555	0.039984	-4.99	0.0004
$ratio*ratio$	-0.467402	0.093011	-5.03	0.0004
$L_{gf}*ROC$	-0.158737	0.039984	-3.97	0.0022
$L_{gf}*V_{gate}$	0.0631326	0.039984	1.58	0.1427
$L_{gf}*L_{gf}$	-0.045268	0.033453	-1.35	0.2032
$ratio*ratio*ratio$	-0.540453	0.023085	-23.41	<.0001
$ratio*ratio*ratio*ratio$	0.3497404	0.022302	15.68	<.0001
$L_{gf}*L_{gf}*L_{gf}$	-0.074228	0.023085	-3.22	0.0082

Table 6.4.15 Coefficient estimates for the response *FOM*.

6.4.3 The Third-Phase Analysis: Prediction of The Best Performance

In this sub-section, the achievable minimum spot size as well as the achievable best figure of merit of the IFE FEDs that can be fabricated by the process facility are explored.

The second-phase experimental design for IFE FEDs indicates that the smallest ratio of V_{focus} to V_{gate} (0.4) gives the best figure of merit as well as the smallest spot size. However, according to the simulation results presented in Figure 4.6.2, the minimum spot size occurs at the point where the electron trajectories are almost parallel to the Y direction shown in Figure 6.4.1. Beyond this point (i.e., smaller focus voltage in Figure 4.6.2), the focus voltage overcompensates for the lateral velocities of the emitted electrons. The electron trajectories cross over the axial line of the emitter, and result in an increase of the spot size.

The goal of the third-phase experimental design is to find the parameter values at which the minimum spot size occurs. This design uses three design parameters: tip ROC, V_{gate} and the ratio of V_{focus} and V_{gate}. As expected, the ratio of V_{focus} to V_{gate} is the dominant effect on the spot size. However, Figure 4.6.2 shows that once the smallest spot size is achieved, about two volts of reduction in IFE voltage results in a zero anode current, because all the emitted electrons are deflected toward the emitter. Consequently, the selection of the range of the parameter *ratio* is essential to locate the turning point. The criteria are:

- values that result in zero anode current are avoided.
- the range of the parameter must be small enough so that the turning point (minimum value) of the spot size can be captured by the regression model.

In this design, the value of tip ROC was reduced. The value of V_{gate} was also reduced accordingly because smaller tip ROC results in lower turn-on voltage.

Finding the design space that captures the minimum spot size as well as the maximum figure of merit does require a few trial-and-error runs. In each trial-and-error run, different values for the five levels of the parameter *ratio* were used, and the other two parameters were kept the same. Fortunately, since there are a total of 15 ($2^3 + 2 \cdot 3 + 1$) treatments in each experimental design, the computational time is about 12 to 14 hours running on a SUN Ultra 30 Model 295 workstation.

After three trial-and-error runs, the range of the parameter *ratio* that captures the minimum spot size was obtained. Table 6.4.16 shows the values for the three design parameters. Note that the five levels of the parameter *ratio* are between 0.32 and 0.40. For a gate voltage of 50 volts, this range of *ratio* gives the focus voltage range between 16 and 20 volts.

Note that the focus aperture radius is fixed at 800 nm.

Parameters	-	0	+	-α	+α	**unit**
ROC	4	5	6	3.268	6.732	nm
V_{gate}	40	50	60	32.68	67.32	Volt
V_{focus}/V_{gate}	0.34	0.36	0.38	0.3254	0.3946	Volt/Volt

Table 6.4.16 The range of FED design parameters for a central composite design ($\alpha = 1.732$).

Figure 6.4.8, Figure 6.4.9, and Figure 6.4.10 are the responses vs. their statistically significant parameters. Table 6.4.17, Table 6.4.18, Table 6.4.19, Table 6.4.20, Table 6.4.21 and Table 6.4.22 are the summaries of fit and the coefficient estimates for each regression model. Note that the leakage current shown in Figure 6.4.8 is the total currents intercepted by the gate, the focus, and the emitter (returned to the emitter). In the simulations, the emitted electrons that are collected on the gate (focus) result in the gate (focus) leakage current. Similarly, the emitted electrons that are deflected back to the emitter give rise to the emitter leakage current.

The results are summarized as:

- Excellent regression model for I_a. This is predictable because of fewer parameters and narrower parameter ranges.
- Relatively large variation in the regression models of D and FOM because the maximum resolution of spot size of the simulator is 0.00005 m, which is not fine

enough to capture the change in D for the small variation of V_{focus} in the third-phase analysis.

- The error bars shown in Figure 6.4.9 also follow the trend of the response curves, which means there is a real relationship between the input parameters and the response, and the effects override the error (inaccuracy) caused by the insufficient resolution of the simulator.
- Turning point of D occurs at a *ratio* between 0.34 to 0.36.
- High leakage current (10-20% of the anode current). The low value of V_{focus} (due to small value of *ratio*) results in a substantially large portion of emitted electrons that are deflected back.
- The average value of figure of merit is about two times larger than that in the second-phase analysis. This is not only due to the smaller spot size, but also due to the smaller tip radius of curvatures.

The results from this third-phase analysis suggests that the smallest spot size from a 60x60 (0.25×0.25 mm^2) emitter array is about 0.35 mm for anode voltage of 5000 volts and anode-gate separation of 10 mm, and the figure of merit can be significantly improved just by increasing the sharpness of emitter tips.

Due to a relatively large voltage difference between the gate and the focus, a substantial amount of emitted electrons are intercepted by the focus or gate, or are deflected back towards the emitters, creating a large amount of leakage current. From the results of first-phase analysis of IFE FEDs, increasing the aperture radius of the focus can effectively reduce the leakage current. Therefore, we perform a new third-phase analysis with a larger value of IFE aperture radius. In this new simulation, the focus aperture radius is 900 nm. The simulated results are shown in Figure 6.4.11, Figure 6.4.12, Figure 6.4.13, Table 6.4.23, Table 6.4.24, Table 6.4.25, Table 6.4.26, Table 6.4.27, and Table 6.4.28. These results behave similarly to the results from the previous third-phase analysis, but the leakage current is reduced to less than 10 % of the anode current.

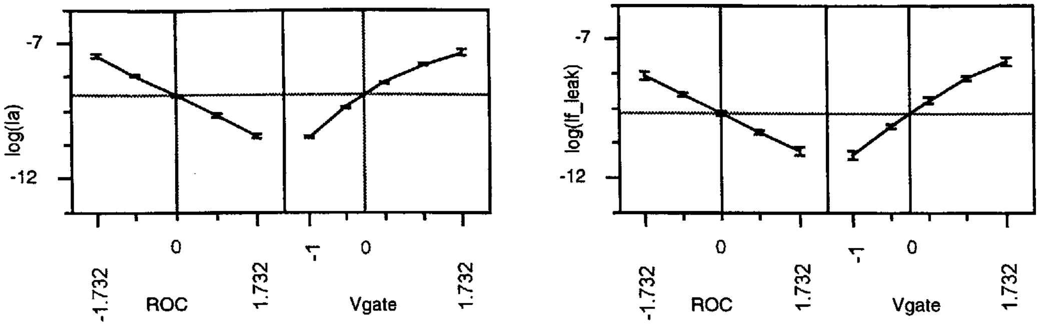


Figure 6.4.8 $\log(I_a)$ and $\log(I_{leak})$ vs. its two significant parameters (ROC and V_{gate}) at their zero levels.

	I_a	I_{leak}
R^2	0.9995	0.998077
Adjusted R^2	0.999166	0.996795
Root Mean Square Error	0.041085	0.081208
Mean of Response	-8.77929	-9.45589
Treatments	11	11

Table 6.4.17 Summary of fit for the response $\log(I_a)$ and $\log(I_{leak})$.

	Coefficients		**Std Error**		**t Ratio**		**Prob>|t|**	
Term	I_a	I_{leak}	I_a	I_{leak}	I_a	I_{leak}	I_a	I_{leak}
Intercept	-8.955	-9.687	0.0179	0.0354	-500.4	-273.8	<.0001	<.0001
ROC	-0.8604	-0.781	0.0121	0.0239	-71.26	-32.74	<.0001	<.0001
V_{gate}	1.3145	1.353	0.0165	0.0327	79.47	41.38	<.0001	<.0001
$V_{gate} * ROC$	0.1731	0.0835	0.0177	0.0349	9.80	2.39	<.0001	0.0539
$V_{gate} * V_{gate}$	-0.2288	-0.152	0.0170	0.0336	-13.46	-4.52	<.0001	0.0040

Table 6.4.18 Coefficient estimates for the response $\log(I_a)$ and $\log(I_{leak})$.

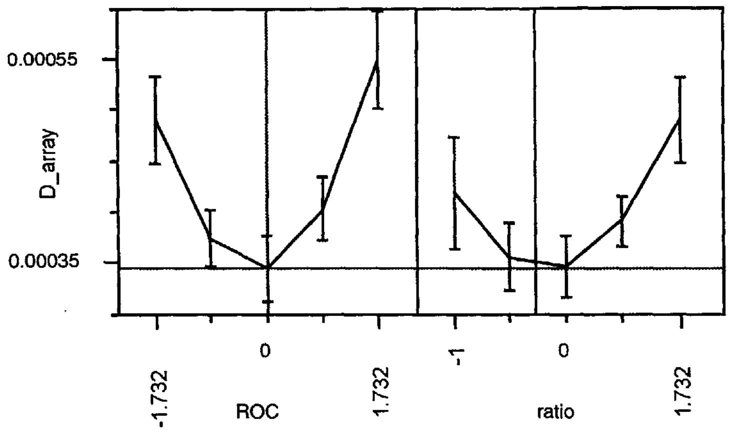


Figure 6.4.9 D vs. its two significant parameters (ROC and $V_{focus}/V_{gate}(=ratio)$) at their zero levels.

R^2	0.978701
Adjusted R^2	0.957402
Root Mean Square Error	0.000019
Mean of Response	0.000427
Treatments	11

Table 6.4.19 Summary of fit for the response D.

Term	**Coefficients**	**Std Error**	**T Ratio**	**Prob>|t|**
Intercept	0.0003444	0.000012	27.91	<.0001
ROC	0.0000168	0.000007	2.28	0.0714
ratio	-0.000016	0.000012	-1.35	0.2352
*ROC*ROC*	0.0000583	0.000006	9.37	0.0002
*ratio*ROC*	-0.000106	0.000012	-9.19	0.0003
*ratio*ratio*	0.0000575	0.00001	5.58	0.0025

Table 6.4.20 Coefficient estimates for the response D.

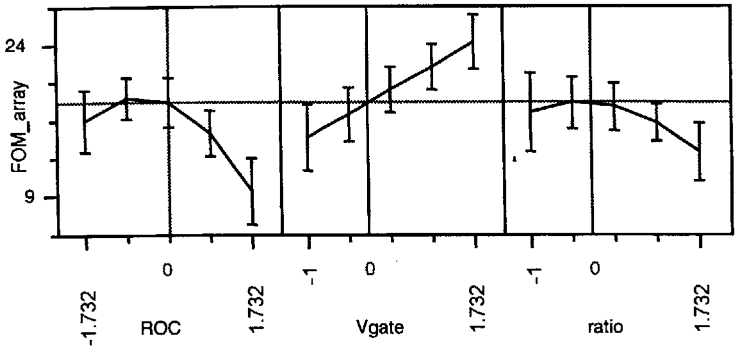


Figure 6.4.10 FOM vs. its three statistically significant parameters (ROC, V_{gate}, V_{focus}/V_{gate}, and L_{gf}) at their zero levels.

R^2	0.994701
Adjusted R^2	0.973504
Root Mean Square Error	0.774548
Mean of Response	16.45875
Treatments	11

Table 6.4.21 Summary of fit for the response FOM.

Term	**Coefficients**	**Std Error**	**t Ratio**	**Prob>|t|**
Intercept	18.171319	0.567982	31.99	0.0010
ROC	-2.008578	0.305724	-6.57	0.0224
V_{gate}	3.4462775	0.324741	10.61	0.0088
$ratio$	-0.466886	0.496528	-0.94	0.4463
$ROC*ROC$	-1.759675	0.274691	-6.41	0.0235
$V_{gate}*ROC$	1.2121388	0.372626	3.25	0.0829
$ratio*ROC$	3.0003822	0.479571	6.26	0.0246
$ratio*V_{gate}$	-0.727614	0.372626	-1.95	0.1901
$ratio*ratio$	-1.329723	0.438478	-3.03	0.0937

Table 6.4.22 Coefficient estimates for the response FOM.

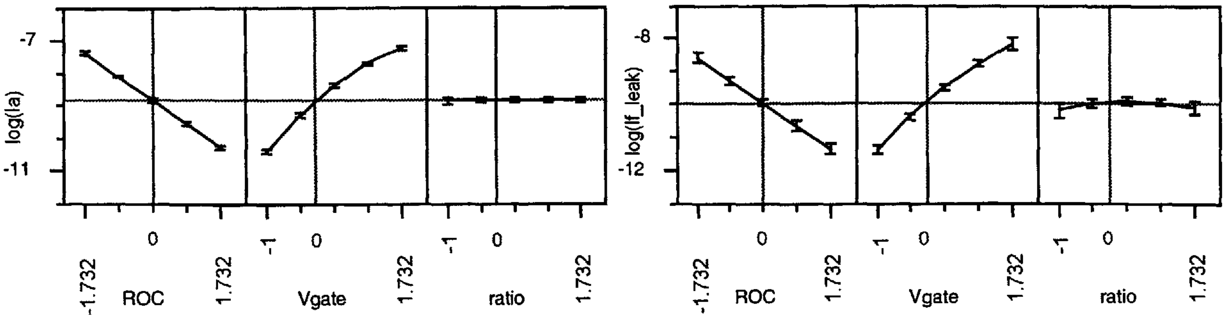


Figure 6.4.11 $\log(I_a)$ and $\log(I_{leak})$ vs. its three significant parameters (ROC, V_{gate}, and $ratio$) at their zero levels.

	I_a	I_{leak}
R^2	0.999919	0.99891
Adjusted R^2	0.99973	0.997276
Root Mean Square Error	0.023343	0.072157
Mean of Response	-8.68591	-9.7679
Treatments	11	11

Table 6.4.23 Summary of fit for the response $\log(I_a)$ and $\log(I_{leak})$.

	Coefficients		**Std Error**		**t Ratio**		**Prob>|t|**	
Term	I_a	I_{leak}	I_a	I_{leak}	I_a	I_{leak}	I_a	I_{leak}
Intercept	-8.885	-9.976	0.0193	0.0402	-460.9	-248.2	<.0001	<.0001
ROC	-0.8448	-0.799	0.0083	0.0286	-102.2	-27.91	<.0001	<.0001
V_{gate}	1.3163	1.2922	0.0095	0.0278	138.48	46.34	<.0001	<.0001
$ratio$	0.0034	0.1137	0.0113	0.0463	0.30	2.45	0.7829	0.0701
$ROC*ROC$	0.0135	NA	0.0084	NA	1.60	NA	0.2085	NA
$V_{gate}*ROC$	0.1739	NA	0.0100	NA	17.31	NA	0.0004	NA
$V_{gate}*V_{gate}$	-0.212	-0.138	0.0107	0.0299	-19.81	-4.61	0.0003	0.0099
$ratio*ROC$	-0.0262	0.1733	0.0126	0.0449	-2.08	3.85	0.1286	0.0182
$ratio*ratio$	NA	-0.119	NA	0.0384	NA	-3.11	NA	0.0359

Table 6.4.24 Coefficient estimates for the response $\log(I_a)$ and $\log(I_{leak})$.

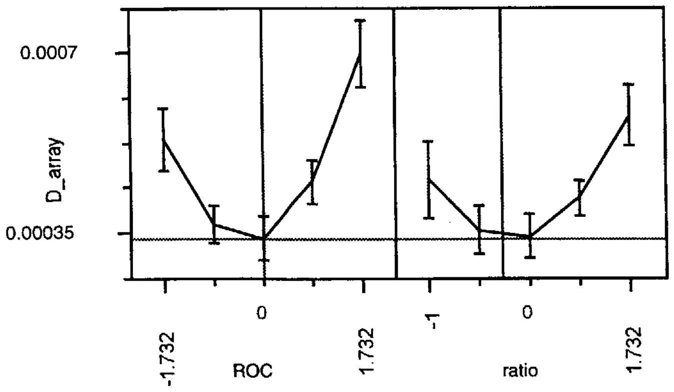


Figure 6.4.12 D vs. its two significant parameters (ROC and $V_{focus}/V_{gate}\,(= ratio)$) at their zero levels.

R^2	0.980456
Adjusted R^2	0.960912
Root Mean Square Error	0.000026
Mean of Response	0.000473
Treatments	11

Table 6.4.25 Summary of fit for the response D.

Term	**Coefficients**	**Std Error**	**t Ratio**	**Prob>|t|**
Intercept	0.0003387	0.000017	20.07	<.0001
ROC	0.0000481	0.00001	4.77	0.0050
ratio	-0.000022	0.000016	-1.33	0.2412
*ROC*ROC*	0.0000917	0.000009	10.77	0.0001
*ratio*ROC*	-0.000139	0.000016	-8.81	0.0003
*ratio*ratio*	0.0000921	0.000014	6.54	0.0012

Table 6.4.26 Coefficient estimates for the response D.

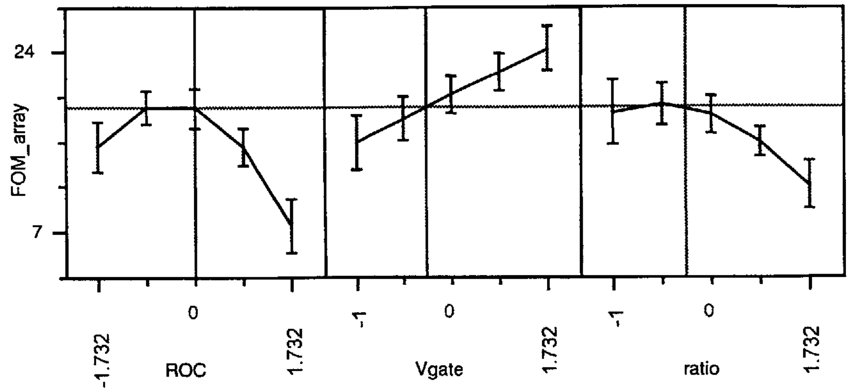


Figure 6.4.13 *FOM* vs. its three statistically significant parameters (ROC, V_{gate}, V_{focus}/V_{gate}, and L_{gf}) at their zero levels.

R^2	0.997563
Adjusted R^2	0.987817
Root Mean Square Error	0.598577
Mean of Response	15.58264
Treatments	11

Table 6.4.27 Summary of fit for the response *FOM* .

Term	Coefficients	Std Error	t Ratio	Prob>|t|
Intercept	18.680549	0.438941	42.56	0.0006
ROC	-2.163769	0.236266	-9.16	0.0117
V_{gate}	3.2582171	0.250963	12.98	0.0059
$ratio$	-1.267784	0.383721	-3.30	0.0807
$ROC*ROC$	-2.466279	0.212284	-11.62	0.0073
$V_{gate}*ROC$	1.0282648	0.287969	3.57	0.0703
$ratio*ROC$	2.8359413	0.370617	7.65	0.0167
$ratio*V_{gate}$	-0.802667	0.287969	-2.79	0.1082
$ratio*ratio$	-1.710459	0.338859	-5.05	0.0371

Table 6.4.28 Coefficient estimates for the response *FOM* .

CHAPTER 7

CONCLUSION

A 3-D CAD tool for modeling and simulating FED/FEA is presented in this thesis. This CAD tool uses I-DEAS and MEMCAD as solid-modelers, and employs the boundary-element-method solver FastLap as an electrostatics solver. The boundary integral theory and numerical techniques for efficient electrostatics simulations are described. Electron trajectory simulation using local and global interpolations of electric field is presented. In this work, the typical boundary-element simulation model of a field emission device has about 5000 to 7000 panels (surface elements), compared with about one hundred thousand elements for three-dimensional finite-element solvers. Neumann-type boundaries are applied in the simulation models to simulate emitter arrays. The computational time for a complete electrostatic simulation is about one hour, versus more than five hours or a few days using three-dimensional finite-element solvers reported by other research groups. Simulation results of the spot size and the Fowler-Nordheim plot are in good agreement with experimental results.

Finding a set of design parameters which gives optimal performance of the device requires a wise strategy in order to minimize the required computation. The technique of DOE explores the responses of the devices by conducting a series of experiments that consists of a structured space of the design parameters. In this study, the experiments are conducted numerically (so called *virtual experiments*) using the CAD tool. Integration of the CAD tool and automation of device solid-model/boundary-condition generation are implemented. About one hundred device models with different configurations can be created, simulated and analyzed in one run. The typical computational time for a 6-parameter central composite design with a total of 77 configurations is about 60 hours.

Two design examples are demonstrated: the proximity focused FED design and the IFE FED design. Their optimal designs and operating conditions are explored. A six-parameter and a four-parameter central composite design are performed for each design example. The six-parameter designs serve as preliminary analyses for exploring the performance in the whole design parameter space. The statistically significant parameters are used to build regression models, which are fairly accurate. Value ranges of parameters, where optimal performance occurs, are also studied. The 4-parameter designs are a refined design of the six-parameter designs. Only statistically (or physically) significant parameters in the 6-parameter design are selected as design parameters, and the value ranges of the parameters are narrowed to the optimal performance ranges. The regression models created by the 4-parameter designs give excellent relationships between the device responses (emission current, spot size and figure of merit) and their statistically significant design parameters (e.g., tip radius of curvature, gate and IFE voltages, gate and IFE aperture radii, anode-gate separation, etc.). Interesting and useful simulation results that can be used for design or operation guidelines are summarized as follows:

- For the proximity focused FEDs, the anode-gate separation is the most (physically) significant parameter. Reducing the separation decreases the spot size significantly, and also slightly increases the emission current density.
- For the IFE focused FEDs, the ratio of the IFE voltage to the gate voltage is the most (physically) significant parameter on the spot size. In FED operation, as long as the ratio is fixed, the spot size is essentially fixed, even though the gate voltage is varied to adjust gray-scale/color of the image.
- Gate thickness does not have any effect on the device performance.
- Smaller gate-aperture radius and tip radius of curvature give higher emission current (lower turn-on voltage). However, the minimum achievable values of these parameters are determined by the fabrication facility and technology.

- For the IFE focused FEDs, the larger the distance between the gate and the focus (L_{gf}), the higher the focusing effect (smaller spot size). Also, a large value of L_{gf} increases the total dielectric strength to prevent insulator breakdown. Its focusing effect becomes flat as L_{gf} increases.

A 3-parameter design was also performed for the IFE FEDs. The purpose of this design is to optimize the IFE FEDs by finding the minimum spot size. Three trial-and-error runs have been performed to find out the appropriate location and range of parameters where the minimum spot size can be observed. The summary of the results is:

- The regression models predict that the smallest achievable spot size for a 60x60 IFE FED array (0.24x0.24 mm^2) is about 0.35 mm for an anode-gate separation of 10 mm and screen voltage of 5000 Volts. The spot can be reduced by using thinner anode-gate separation.
- This minimum value of the spot size occurs at the ratio (V_{focus}/V_{gate}) value in the range of 0.34 - 0.36.
- High leakage current was observed when the minimum spot size occurs. Increasing the radius of focus aperture by 100 nm reduces the leakage current substantially.

This CAD tool and design strategy has been proven to be excellent in characterization and optimization for FEDs.

REFERENCES

[1] D. Cathey, "Field-emission displays," Information Display, vol.11, no.10, Oct. 1995, pp. 16-20.

[2] T. Utsumi, "Vacuum Microelectronics for Future Display Technology," Journal of SID, 1/3, 1993, pp. 313-317.

[3] J. Ghrayeb, T. W. Jackson, R. Daniels, and D. Hopper, "Review of Field Emission Display Potential as a Future (leap-frog) Flat Panel Technology," SPIE vol. 3057, 1997, pp. 237-248.

[4] J. Jung, Y.-H. Kim, B. Lee, and J. D. Lee, "Optimal Design of FEA Using an Evolution Strategy," Technical Digest of IVMC '97, pp. 341-345, Kyongju, Korea, 1997.

[5] W. D. Kesling and C. E. Hunt, "Beam Focusing for Field-Emission Flat-Panel Displays," IEEE Transaction on Electron Devices, Vol. 42, No. 2, pp. 340-347, February 1995.

[6] W. B. Herrmannsfeldt, R. Becker, I. Brodie, A. Rosengreen and C. A. Spindt, "High-resolution Simulation of Field Emission," Nuclear Instruments and Methods in Physics Research, A298, 1990, pp. 39-44.

[7] H. C. Lee and R. S. Huang, "Simulation and Design of Field Emitter Array," IEEE Electron Device Letters, Vol. 11, No. 12, 1990.

[8] W. D. Kesling and C. E. Hunt, "Field Emission Device Modeling for Application to Flat Panel Displays," Journal of Vacuum Science Technology, Vol. B11, No. 2, pp. 518-522, 1993.

[9] Z. Cui and L. Tong, "Optimum Geometry and Space-Charge Effects in Vacuum Microelectronic Devices," IEEE Transaction on Electron Devices, Vol. 40, No. 2, Feb. 1993, pp. 448-451.

[10] W. D. Kesling and C. E. Hunt, "Field-Emission Display Resolution," SID 93 Digest, 1993, pp. 599-602.

[11] K. C. Choi and N. Chang, "Modeling of a Field Display Using the Adaptive Scheme Method," SID 94 Digest, Santa Ana, CA, 1994, pp.237-40.

[12] D. W. Jenkins, "Emission Area of a Field Emitter Array," IEEE Transaction on Electron Devices, Vol. 40, No. 3, March 1993, pp. 666-672.

[13] E. G. Zaidman, "Simulation of Field Emission Microtriodes," IEEE Transaction on Electron Devices, Vol. 40, No. 5, May 1993, pp. 1009-1016.

[14] T. Su, C. L. Lee, and J. C.-M. Huang, "Electrical and Thermal Modeling of a Gated Field Emission Triode," IEDM 93, Washington D. C., 1993, pp. 765-768.

[15] M. G. Ancona, "Thermomechanical Factors in Molybdenum Field Emitter Operation and Failure," IEDM 94, San Francisco, CA, 1994, pp. 803-806.

[16] R. L. Hartman, W. A. Mackie, and P. R. Davis, "Use of Boundary element methods in field emission computations," Journal of Vacuum Science Technology, Vol. B12, pp. 754-758, 1994.

[17] H. Y. Ahn, C. G. Lee and J. D. Lee, "Numerical Analysis on Field Emission for the Effects of the Gate Insulators," Journal of Vacuum Science Technology, Vol. B13, No. 2, pp. 540-544, 1995.

[18] R. L. Hartman, W. A. Mackie, and P. R. Davis, "Three Dimensional Axisymmetric Space Charge Simulation via Boundary Elements and Emitted Particles," Journal of Vacuum Science Technology, Vol. B14, pp. 1952-1957, 1996.

[19] Y. Hori, K. Koga, K. Sakiyama, W. Kanemaru and J. Itoh, "Tower Structure Si Field Emitter Arrays with Large Emission Current," IEDM 95, pp. 393-396.

[20] R. B. Marcus, et al, "Simulation and Design of Field Emitters", IEEE Transaction on Electron Devices, Vol. 37, No. 6, June 1990.

[21] B. Lee, S.-W. Min, J.-H. Bang, D. S. Hwang, and T.-J. Kwon, "Modeling of Tunneling Current at Field Emission Display Tips : For Non-metallic Tips and an Analytical Method of Solving for Electric Fields," IVMC'96. 9th International Vacuum Microelectronics Conference. Technical Digest, 1996, pp. 97-101.

[22] K. L. Jensen, E. G. Zaidman, M. A. Kodis, B. Goplen and D. N. Smithe, "Analytical and Semi-numerical Models for Gated FEA'S," IVMC '95. Eighth International Vacuum Microelectronics Conference. Technical Digest, New York, NY, 1995, pp.216-20.

[23] J. S. Park and J. I Han, "A Theoretical Estimation of Beam Spot Size in a Single Emitter," IVMC 97, pp. 331-335.

[24] MEMCAD 3.2 User Manual, Microcosm Technologies Inc., Cambridge, MA, 1997.

[25] K. Nabors, T. Korsmeyer, F. T. Leighton and J. White, "Preconditioned, Adaptive, and Multipole-Accelerated Iterative Methods for Three-Dimensional First-Kind Integral Equations of Potential Theory," SIAM J. Sci. Comp., Vol. 15, No. 3, pp. 713-735, May 1994.

[26] J. Trevelyan, Boundary Elements for Engineers: Theory and Applications, Computational Mechanics Publications, Boston, 1994.

[27] D. C. Montgomery, Introduction to Statistical Quality Control, John Wiley & Sons Inc., New York, 1996.

[28] A. Modinos, Field, Thermionic, and Secondary Electron Emission Spectroscopy, Plenum Press, New York, 1984.

[29] H. Busta, "Vacuum Microelectronics-1992," Journal of Micromechanics and Macroengineering, Vol. 2, No. 2, June 1992, pp. 43-74.

[30] R. H. Good jr. and E. W. Muller, "Field Emission," Handbuch der Physik Spring, 1956, Vol. XXI.

[31] P. A. Keller, The Cathode-Ray Tube: Technology, History and Applications, Palisades Institute for Research Services, Inc., New York, 1991.

[32] R. H. Fowler and L. W. Nordheim, "Electron Emission in Intense Electric Fields," Proc. R. Soc. London, Vol. A119, 1928, p173.

[33] P. H. Cutler, T. E. Feuchtwang, Z. Huang and T. E. Sullivan, "Tunnelling Theory and Vacuum Microelectronics," Vacuum Microelectronics 89, Inst. Phys. Conf. Ser. No. 99: section 5, pp. 121-132.

[34] P. H. Cutler, J. He, J. Miller, N. M. Miskovsky, B. Weiss and T. E. Sullivan, "Theory of Electron Emission in High Fields from Atomically Sharp Emitters: Validity of the Fowler-Nordheim Equation," Progress in Surface Science, Vol. 42, 1993, pp. 169-185.

[35] R. G. Forbes, "New Theory for the Derivation of Emission Area from A Fowler-Nordheim Plot," Eleventh International Vacuum Microelectronics Conference. IVMC'98, New York, 1998, pp.3-4.

[36] C. A. Spindt, I. Brodie, L. Humphrey, and E. R. Westerberg, "Physical Properties of Thin-film Field Emission Cathodes with Molybdenum Cones," Journal of Applied Physics, Vol. 47, No. 12, Dec 1976, pp. 5248-5263.

[37] H. F. Gray, "The Field-emitter Display," Information Display, Vol. 9. No. 3, 1993, pp. 9-14.

[38] T. Utsumi, "Keynote Address, Vacuum Microelectronics: What's New and Exciting", IEEE Transaction on Electron Devices, Vol. 38, No. 10, Oct. 1991, pp. 2276-2283.

[39] A. I. Akinwande, B. R. Johnson, J. O. Holmen, D. Murphy, and D. K. Arch, "Field-emission Lamp for Avionic AMLCD Backlighting," SID 96 Digest, 1996, pp. 745-748.

[40] C. Py, J. Itoh, T. Hiran, S. Kanemaru. "Beam focusing characteristics of silicon microtips with an in-plane lens," IEEE Transactions on Electron Devices, vol.44, no.3, March 1997, pp.498-502.

[41] Y. Tima, S. Kanemaru and J. Itoh, "Electron Beam Characteristics of Double-gated Si. Field Emitter Arrays," IVMC `95. Eighth International Vacuum Microelectronics Conference. Technical Digest, New York, NY, 1995, pp.9-13.

[42] J. Itoh, Y. Tohma, K. Morikawa, S. Kanemaru and K. Shimizu, "Fabrication of Double-gated Si Field emitter Arrays for Focused Electron Beam Generation," Journal of Vacuum Science Technology, Vol. B13, No. 5, 1995, pp. 1970-1972.

[43] C.-M. Tang and T. A. Swyden, "Beam Collimation from Field-Emitter Arrays with Linear Planar Lenses", SID 97 Digest, 1997, pp. 115-118.

[44] C. H. Tsai, W. C. Wang, J. T. Lai, C. H. Wang, C. H. Tien and K. L. Tsai, "A CRT-Like FED," SID 97 Digest, pp. 119-122.

[45] A. Palevsky, G. Gammie and P. Koufopoulos, "Field-Emission Displays : A 10,000-fL High-Efficiency Field-Emission Display," SID 94, Digest of Technical Papers, Vol. XXV, pp. 55-57, San Jose, CA, June, 1994.

[46] L. Dvorson, Field Emitters with Integrated Focusing Electrode, Master Thesis, Massachusetts Institute of Technology, Sep., 1998.

[47] J. Ghrayeb, T. W. Jackson, D. G. Hopper, and R. Daniels, “Approach of Field Emission Display toward Technology Status”, SPIE vol. 3363, 1998, pp. 244-255.

[48] F. Charbonnier, “Voltage Breakdown in vacuum Microelectronics Microwave Devices Using Emitter Arrays : Causes, Possible Solutions and Recent Progress,” IVMC'97. 1997 10th International Vacuum Microelectronics Conference. Technical Digest, Seoul, South Korea, 1997, pp. 7-13.

[49] B. R. Chalamala, R. H. Reuss, T. A. Trottier, C. W. Penn and Yi Wei, “Device reliability issues in field emission displays”, SPIE Int. Soc. Opt. Eng. Proceedings of Spie the International Society for Optical Engineering, Vol. 3363, No. 1998, pp. 15-17.

[50] K. R. Sarma, T. Akinwande, “Flat Panel Display for Portable Systems,” Journal of VLSI Signal Processing Systems 13, 1996, pp. 165-190.

[51] T. J. Nelson and J. R. Wullert II, “Electronic Information Display Technology,” World Scientific, New Jersey, 1997.

[52] S. Matsumoto et. al., Electronic Display Devices, John Wiley & Sons, New York, 1990.

[53] S. Sherr, Electronic Displays, 2nd ed., John Wiley & Sons, New York, 1993.

[54] A. Ghis, R. Meyer, P. Rambaud, F. Levy, and T. Leroux, “Sealed Vacuum Devices: Fluorescent Microtip Displays,” IEEE Transactions on Electron Devices, Vol. 38, No. 10, 1991, pp. 2320-2322.

[55] C. A. Spindt, C. E. Holland, I. Brodie, J. Mooney and E. R. Westerberg, “Field-emitter Arrays Applied to Vacuum Fluorescent Display,” IEEE Transaction on Electron Devices, Vol. 36, No. 1 Jan. 1989, pp. 225-227.

[56] A. I. Akinwande, 1993 Presentation.

[57] C. A. Spindt, C. E. Holland, A. Rosengreen, and I. Brodie, “Field-emitter Arrays for Vacuum Microelectronics,” IEEE Transactions on Electron Devices, Vol. 38, No. 10, 1991, 2355-2363.

[58] R. E. Neidert, P. M. Philips, S. T. Smith, and C. A. Spindt, “Field Emission Triodes,” IEEE Transactions on Electron Devices, Vol. 38, No. 3, 1991, pp. 661-665.

[59] K. Yokoo and H. Ishizuka, “RF Applications Using Field Emitter Arrays (FEAs),” IVMC 96, St Petersburg, 1996, pp. 490-500.

[60] S.-C. Lu, J. C.-M. Huang, C.-L. Lee, J.-M. Wang, J.-G. Peng, W.-C. Wang, J.-H. Tsai and D. Liu, “A Robust Gated-Field-Emission Triode,” IEDM 95, pp. 397-400.

[61] C. A. Spindt, C. E. Holland, P. R. Schwoebel, and I. Brodie, "Field-emitter-array Development for Microwave Applications," IEDM 95, 1995, pp. 389-342.

[62] C. H. Liu, et. al., "A high-sensitivity micromachined accelerometer based on electron tunneling transducers," Microelectromechanical System (MEMS), 1997 Micromechanical Systems Microscale Energy Transport. The 1997 ASME International Mechanical Engineering Congress and Exposition. ASME. 1997, pp.13-20.

[63] P. W. Kruse, Uncooled Infrared Imaging, Semiconductors and Semimetals, Vol. 47, Academic Press, Boston, 1997.

[64] M. G. Ancona, "Thermomechanical Analysis of Failure of Metal Field Emitters," Journal of Vacuum Science Technology, Vol. B13, No. 6, 1995, pp. 2206-2214.

[65] B. W. Chui et. al., "Low-stiffness silicon cantilevers with integrated heaters and piezoresistive sensors for high-density AFM thermomechanical data storage," Journal of Microelectromechanical Systems, Vol. 7, No.1, March 1998, pp.69-78.

[66] J. S. Sovey et. al., "Development status of the NASA 30-cm ion thruster and power processor," Collection of Technical Papers. 29th Intersociety Energy, AIAA. Vol.2, 1994, pp. 756-63.

[67] S. Z. Katai et. al., "Ion beam mass spectrometer for compositional analysis of plasma assisted surface processes in the pressure range of 1-50 mbar," Review of Scientific Instruments, vol.70, no.8, Aug. 1999, pp. 3324-8.

[68] Y. Saad and M. H. Schultz, "GMRES: A generalized minimum algorithm for solving nonsymmetric linear systems," SIAM J. Sci. Stat. Comp., Vol. 7, pp. 856-869, July 1986.

[69] C. A. Brebbia, J. Telles, L. C. Wrobel, Boundary Element Techniques, Springer-Verlag, New York, 1984.

[70] W. S. Hall, The Boundary Element Method, Kluwer Academic Publishers, Boston, 1994.

[71] C. A. Brebbia and J. Dominguez, boundary Elements: An Introductory Course, Computational Mechanics Publications, Boston, 1992.

[72] K. E. Atkinson, "A Survey of Boundary Integral equation Methods for the Numerical Solution of Laplace's Equation in Three Dimensions," Numerical Solution of Integral Equations, Plenum Press, New York, 1990, pp. 1-34.

[73] R. Kress, Linear Integral Equations, Spring-Verlag, Berlin, 1989.

[74] D. Colton and R. Kress, "Integral Equation Methods in Scattering Theory", Krieger, Malabar, Florida, 1992.

[75] Y.-J. Yang, F. T. Korsmeyer, V. Rabinovich, M. Ding, S. D. Senturia, and A. I. Akinwande, "An Efficient 3-Dimensional CAD Tool for Field-Emission Devices", International Electron Devices Meeting (IEDM) 1998, San Francisco, Dec 1998, pp. 863-866.

[76] I-DEAS Master Series 5, SDRC, Ohio.

[77] Z. Liu and J. Ximen, "Space Charge Effects in Miniaturized Field Emission Systems," IVMC'97. 1997 10th International Vacuum Microelectronics Conference. Technical Digest, Seoul, South Korea, 1997, pp. 353-356.

[78] F. T. Korsmeyer, K. Nabors and J. White, FastLap Version 2.0 Manual, Massachusetts Institute of Technology, 1996.

[79] F. T. Korsmeyer, D. K. P. Yue, K. Nabors and J. White, "Multipole-Accelerated Preconditioned Iterative Methods for Three-Dimensional Potential Problems," BEM 15, Worcester, MA, 1993.

[80] K. S. Nabors, "Efficient Three-Dimensional Capacitance Calculation," Ph.D. Dissertation, Massachusetts Institute of Technology, 1993.

[81] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, Numerical Recipes in C, Cambridge University Press, 1988.

[82] A. Hosono, S. Kawabuchi, S. Horibata, S. Okuda. H. Harada, and M. Takai, "High Emission Current Double-gated Field Emitter Arrays," Technical Digest of IVMC '98, pp. 97-98, Asheville, North Carolina, July 1998.

[83] D. C. Montgomery, Design and Analysis of Experiments, 4th ed., John Wiley & Sons Inc., New York, 1996.

[84] G. E. P. Box, W. G. Hunter and J. S. Hunter, Statistics for Experimenters, John Wiley & Sons, Inc., New York, 1978.

[85] S. Wolf and R. N. Tauber, Silicon Processing for the VLSI Era, Vol. 1, Lattice Press, Sunset Beach, CA, 1986.

[86] A. Leon-Garcia, Probability and Random Processes for Electrical Engineering, 2nd Ed., Addison-Wesley Publishing Company, 1994.

[87] L. S. Myers and D. C. Montgomery, Response Surface Methodology : Process and Product Optimization Using Designed Experiments, Wiley, New York, 1995.

[88] A. I. Khuri and J. A. Cornell, Response Surfaces – Designs and Analyses, 2nd ed., Marcel Dekker Inc., 1996.

[89] S. D. Senturia, "CAD challenges for microsensors, microactuators, and microsystems," Proceedings of the IEEE, Vol. 86, No.8, Aug. 1998, pp. 1611-1626.

[90] D. S. Boning, P. K. Mozumder, "DOE/Opt: A System for Design of Experiments, Response Surface Modeling, and Optimization Using Process and Device Simulation", IEEE Transaction on Semiconductor Manufacturing, Vol. 7, No. 2, May 1994, pp. 133-244.

[91] TECPLOT Version 7.0 User's Manual, Amtec Engineering, Inc., 1996.

[92] JMP Introduction Guide, version 3 of JMP, SAS Institute Inc., Cary, NC, 1998.

[93] C. G. Lee, H. Y. Ahn, and J. D. Lee, "Scaling-down of Cone-like Field Emitter Using LOCOS," IEDM 95, pp. 401-404.

[94] D. G. Pflug, M. Schattenburg, H. I. Smith, and A. I. Akinwande, "100nm Aperture Field Emitter Arrays for Low Voltage Applications," International Electron Devices Meeting (IEDM) 1998, San Francisco, Dec 1998, pp. 855-858.

[95] M. Tanaka, K. Takayama, A. Azeta, K. Yano, and T. Kishino, "A New Structure and Driving Syetem for Full-color FEDs," SID 97 digest, pp. 47-51.

S230-20