IMPROVEMENT OF GFP EXPRESSION QUANTIFICATION

THESIS

Presented in Partial Fulfilment of the Requirements for
The Degree Master of Science in the Graduate
School of the Ohio State University

By

Congling Wu, M.S.

*****

The Ohio State University
2008

Master’s Examination Committee:

Dr. Peter P. Ling, Adviser
Dr. John J. Finer
Dr. Lingying Zhao

Approved by

———
Adviser
Graduate Program in Food,
Agricultural and Biological
Engineering
ABSTRACT

Green fluorescent protein (GFP) has been widely used as a scorable reporter gene to study transgene activities in medical and biological areas. GFP expression quantification is used to evaluate genetic transformation and gene expression. GFP quantification in plant tissues, however, is often complicated by the presence of chlorophyll (excitation at 480 nm and emission around 683 nm) that can interfere with GFP emission. The overarching goal of this thesis is to improve the quantification of the GFP in transgenic embryogenic soybean tissues using biological inhibition, spectral feature selection, image processing approaches.

An innovative approach of using the herbicide isoxaflutole (IFT) to eliminate chlorophyll and to enhance the GFP expression quantification was first developed. The inclusion of this “bleaching” herbicide at 3 or 10 mg/L in a standard soybean embryo proliferation medium resulted in a change in tissue color from green to non-pigmented over the course of a 4-week experiment. Tissue growth remained unaffected. GFP expression in 3 different transgenic soybean clones, representing low to moderate GFP expression levels, was easily detected and quantified using image analysis following culture of the tissues on an IFT-containing medium. Quantification of GFP in tissues from the same clones cultured in the absence of IFT, however, was difficult using image
analysis. The bleaching effects from this herbicide appear to be reversible, which makes IFT and possibly other bleaching herbicides useful for the analysis of GFP expression in tissues. Another approach of using spectral feature selection to extract GFP emission and block other autofluorescence was also investigated. A spectral band (500-590 nm) was identified for extracting the GFP signal and blocking chlorophyll interference. Two commercially available filter sets were evaluated. Results showed that the broad spectral band gave a good image contrast between the tissues and the growth media background, and the narrow spectral band allowed only the GFP signal to pass through. The narrow band images had higher brightness contrast between GFP-expressing and non-expressing areas of the tissue, and lower contrast between the tissue and its growth media background than that of the broad band images. The third effort to improve GFP expression quantification was to apply effective imaging processing techniques. Towards this end, three segmentation algorithms, namely local adaptive thresholding (mean), normalized cut algorithm (Ncuts), and K-means clustering algorithm were evaluated for the GFP quantification task. Among these three techniques, the Ncuts algorithm had the best performance and the highest accuracy in segmenting fluorescence images of soybean tissues, and it was least sensitive to noise.

This study is the first endeavour to systematically investigate three different approaches to improve the GFP expression detection by mitigating the interference from the chlorophyll in transgenic embryogenic soybean tissues. These approaches and
findings will provide opportunities for the new development of the transgenic technology and potentially contribute to the reduction of the world hunger.
Dedicated To

My Grandfather

My Parents

And

My Husband
ACKNOWLEDGMENTS

I thank my advisor, Dr. Peter Ling, who gave me invaluable help and guidance for my research.

I was brought to a new field by Dr. Finer, who provided me a good working environment and taught me lot of biological knowledge. Thank Joe and Cheri for their help with my experiments and teaching me many indispensable research skills.

I should express my grateful appreciation to Dr. Zhao, who generously became one of my committee members when I need her and gave me valuable instructions to my thesis.

I thank Carol and Peggy for their considerate help, which made it easy to adapt the life and work style in US.

I could have felt lonely if I did not have my good friend and officemate Dr. Noriko Takahashi. We spent very happy time together and she gave me lot of help. Thank my good friends both in Columbus and in Wooster, Xinying, Xianwen, Wenjing, Ran, Xiaojian, Jingjing, Shijin, Yiguang, Kaixuan, Li Sun, Yu Chen, Anjali, Anita, Sukhinder, Smitha, Hehe, Xiaolu, Na, Xiulan, Linjian and Xuehui who offered me precious friendship and kindly gave me encouragements. They were my thoughtful listeners when I felt stressed or frustrated.

Last but not the least, I should thank my families for their consistent support and encouragements. My husband has been motivating me to complete this thesis at every stage. I thank my parents and parents-in-law for their everlasting love, care and support.
VITA

Feb 16, 1977 ........Born in Inner Mongolia, China.

July, 2002.............B.S. Computer Science, Inner Mongolia Normal University

July, 2005.............M.S. Agricultural Engineering, China Agricultural University

Sep, 2005-present..........Graduate Research Associate, The Ohio State University

PUBLICATIONS

1. Wu, C., J. M. Chiera, P. P. Ling, and J. Finer. Isoxaflutole treatment leads to reversible tissue bleaching and allows for more effective detection of GFP in transgenic soybean tissues. In Vitro Cellular and Developmental Biology. Accepted.


FIELDS OF STUDY

Major Field: Food, Agricultural and Biological Engineering
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vi</td>
</tr>
<tr>
<td>Vita</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Significance of this work</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Fluorescent marker gene in transgenic plant production</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Fluorescence and fluorescence imaging microscopy</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Challenges of GFP expression quantification</td>
<td>5</td>
</tr>
<tr>
<td>1.5 Goal and objectives of the study</td>
<td>6</td>
</tr>
<tr>
<td>1.6 Structure of the thesis</td>
<td>6</td>
</tr>
<tr>
<td>2. Application of Herbicide IFT to Suppress Chlorophyll Fluorescence</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Materials and Methods</td>
<td>10</td>
</tr>
<tr>
<td>2.2.1 Plant material preparation</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2 IFT treatment</td>
<td>11</td>
</tr>
<tr>
<td>2.2.3 Image acquisition and processing</td>
<td>11</td>
</tr>
<tr>
<td>2.2.4 Recovery from IFT treatment</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Results and Discussion</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Conclusions</td>
<td>22</td>
</tr>
<tr>
<td>3. Spectral Feature Selection for the Fluorescence Microscopic System</td>
<td>24</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>24</td>
</tr>
</tbody>
</table>
3.1.1 GFP and chlorophyll spectral features ..............................................................26
3.1.2 Spectral filters in a fluorescence microscopic system ........................................29
3.2 Materials and Methods ....................................................................................32
  3.2.1 Data collections and measurements of spectral features ..............................32
  3.2.2 Spectral feature selection ............................................................................38
  3.2.3 Spectral filter evaluation ............................................................................39
3.3 Results and Discussion ....................................................................................41
  3.3.1 Chlorophyll emission spectrum ....................................................................41
  3.3.2 Spectral feature selection for emission filter ..............................................42
  3.3.3 Spectral feature evaluation ........................................................................45
3.4 Conclusions .......................................................................................................53
4. Image Segmentation Evaluation for Fluorescence Image ....................................55
  4.1 Introduction .....................................................................................................55
  4.2 Methodology ..................................................................................................57
    4.2.1 Segmentation algorithms description .......................................................57
    4.2.2 Comparison of three algorithms .............................................................59
      4.2.2.1 Accuracy .............................................................................................60
      4.2.2.2 Robustness .........................................................................................62
      4.2.2.3 Running time .....................................................................................64
  4.3 Results and Discussion ..................................................................................64
    4.3.1 Accuracy ..................................................................................................64
    4.3.2 Robustness ...............................................................................................66
    4.3.3 Running time ............................................................................................72
  4.4 Conclusions ....................................................................................................72
5. Conclusions ........................................................................................................74
Bibliography ............................................................................................................77
Appendix ....................................................................................................................84
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Segmentation results of three algorithms: Normalized cut (Ncuts), Adaptive thresholding (Adaptive mean), and K-means clustering. A low contrast image was added salt and pepper noises with various noise density of 0, 0.001, 0.005, 0.01, 0.02, 0.03, 0.05 and 0.1 respectively (The 8 gray images in the first column). Contours indicate the segment boundaries of NCuts algorithm (the second column)</td>
</tr>
<tr>
<td>4.2</td>
<td>Average running time per image for segmenting 20 images with the three methods</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Growth comparison of non-transformed control, and low, low/moderate and moderate-expressing GFP clones cultured on the media containing 0, 3 or 10 mg/L IFT.</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison of the pigment reductions, and the improved GFP detection by culturing the moderately expressing GFP clone on medium containing either 0 (A, B), 3 (C, D), or 10 (E, F) mg/L IFT. Images of tissues were collected (left to right) at 0, 1, 2, 3 and 4 weeks after IFT treatment. Top panels (A, C, and E) show images collected under white light. Bottom panels (B, D, and F) show images of the same tissue taken under blue light for GFP expression observation.</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>GFP expression in tissues of control (non-transformed) and the three transgenic clones with varied expression levels from low to moderate that were grown on media containing 0, 3 or 10 mg/L IFT.</td>
<td>19</td>
</tr>
<tr>
<td>2.4</td>
<td>GFP expression in control (non-transformed) tissues and the three independent GFP expressing clones, grown on medium containing either a) no IFT or b) 10 mg/L IFT.</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Recovery of tissue following isoxaflutole treatment: non-pigmented tissue as result of the treatment (left) and green coloration in new growth, 2 weeks following subculture of the tissue onto a medium without isoxaflutole (right).</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>Excitation and emission spectra of GFP (Data from Molecular Probes, Inc.)</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Excitation and emission spectra of chlorophyll a and b.</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>Light path inside a fluorescence microscopic imaging system.</td>
<td>30</td>
</tr>
<tr>
<td>3.4</td>
<td>Schematic of taking white reference before measurements, the probe and the light source should be kept the same angle (45°) from the white standard.</td>
<td>35</td>
</tr>
</tbody>
</table>
3.5 The spectroradiometer was calibrated by measuring reflectance of a white Standard, which should be a complete reflectance .................................................................36

3.6 Schematic of the measurements of transmission efficiency of the spectral filters, the filters was placed in front of the probe..........................36

3.7 Transmission efficiency of the planapochromatic objective 1x (Leica Microsystems Ltd.). ........................................................................................................37

3.8 Quantum efficiencies of the sensor of the fluorescent imaging system. The monochrome CCD sensor has quantum efficiency of approximately 45% in the GFP expressing band (around 510 nm), and less than 30% in the chlorophyll fluorescence range (600-800 nm). As results of color filtering, quantum efficiency of the sensor reduced to 12% (green channel) and 5% (red channel) for the detection of GFP expressing and chlorophyll fluorescence signals ........................................................................38

3.9 A histogram of a typical GFP fluorescence image. The two thresholds separate the histogram into three subdistributions: distributions of growth media background and tissue area, including GFP-expressing and non-expressing tissue areas. Separability was defined as the distance between two modes of two subdistributions brightness divided by the sum of their standard deviations of these two subdistributions........................................41

3.10 Detected chlorophyll fluorescence spectrum in D20 soybean tissues with the spectroradiometer..........................................................42

3.11 Detected GFP and chlorophyll signals in blue, green and red channels of the camera (without filtering by a dichroic mirror and an emission filter). Here the data is from the products of GFP fluorescence data from Molecular Probes, Inc. (Fig. 3.1) and chlorophyll fluorescence data detected in soybean D20 tissues transformed with 35S-sGFP (S65T) (Fig. 3.10) with the collection efficiency and transmission efficiency of the objective as well as the quantum efficiencies of the CCD sensor mounted with three color filters.........44

3.12 Transmission efficiency of the GFP3 emission filter .................................................46

3.13 Transmission efficiency of the GFP2 emission filter .................................................46
3.14 Effects of the GFP3 narrow band (500-550 nm) emission filter on the fluorescent signal strength in red, green, and blue channels of the image acquisition system. GFP fluorescence passed through the emission filter, while chlorophyll fluorescence was blocked. ........................................................47

3.15 Effects of GFP2 emission filter on the fluorescent signal strength in blue, green and red channels of the image acquisition system. Both GFP and chlorophyll fluorescence transmit through the filter. ........................................................48

3.16 Comparisons of separabilities between background and tissue areas (separability1) and between GFP-expressing and non-GFP expressing areas (separability2) of GFP2 and GFP3 images. Both separability1 and separability2 were significantly different in the images of (a) IFT-free transgenic soybean tissues, (b) 3 mg/L IFT-treated transgenic soybean tissues, (c) 10 mg/L IFT-treated transgenic soybean tissues, taken through GFP2 and GFP3 filters. ...........................................................................................................50

3.17 Segmentation results of the GFP-expression area in GFP2 and GFP3 images. Red circle represents the interference of chlorophyll fluorescence. The optimal brightness threshold for GFP2 image is 0.25, and 0.2 for the GFP3 image. a, GFP2 images; b, segmented GFP-expressing area from the GFP2 image; c, GFP3 image; d, segmented GFP-expressing area from the GFP3 image .....................................................................................52

4.1 Dynamic ranges of 24 images which were selected to segment. Their dynamic ranges were from 5 to 66, selected from the images of three clones with low, low/moderate and moderate GFP expressions and non-transformed ‘Jack’ treated with 0, 3 and 10 mg/L IFT concentration levels, in which 12 were taken through the broad band-pass filter (GFP2 images), and 12 got by the narrow band-pass filter (GFP3 images) ..........60

4.2 Procedures of the tissue area extracting program .................................................62

4.3 Average tissue area ratio relative to whole image of 24 images obtained with three automatic segmentation methods: normalized cut (Ncuts), K-means clustering (Kmeans), and adaptive thresholding (Adaptive mean) and interactive thresholding (Manual) ..................................................................................66

4.4 Two types of misclassification error of three algorithms for
segmenting 24 fluorescence images

4.5 Two types of misclassification error of three algorithms for 12 images acquired with the GFP2 filter and 12 images with the GFP3 filter

4.6 Two types of misclassification error of three segmentation methods: K-means clustering (k=2 and k=3), normalized cut (Ncuts), adaptive thresholding (mean) for an image with a dynamic range of 6.9, which was treated with 8 noise density levels, These images were 8 categories (a through h) by the added noises with a noise density d=0, 0.001, 0.005, 0.01, 0.02, 0.03, 0.05, and 0.1, respectively
CHAPTER 1

INTRODUCTION

1.1 Significance of this work

Today, food shortage is still a tremendous challenge for all human beings around the world. It is estimated that roughly 800 million people are still facing hunger and nearly 6 million children under five years old die of malnutrition every year in the developing world (US National Academy of Sciences, 2000). Furthermore, global climate change and alternative use of land reduced regional production of food. Substantial measures must be taken to sustainably address the food shortage problem. In the past decades, pesticides and fertilizers have increased food production, but they are expensive, can potentially damage the environment and are perceived as harmful to human health. Thus, increasing world food production without compromising the well-being of current and future generations is a considerable challenge.

Transgenic technology, first developed for medical applications during the 1970s, offers hope to address this great challenge. This technology introduces a gene or genes from one individual into the cellular DNA of another individual, even when the genes come from different species. The gene or genes are inserted into a fertilized egg, or cell, and become integrated into the DNA of that cell. Transgenic technology has
been utilized to develop transgenic crop plant varieties with novel traits serving human needs, such as improved yield, improved nutritional quality, herbicide resistance, and other tolerance to pests, disease, salinity and drought stresses. Soybean is the most commonly grown transgenic crop, and the transgenic soybean covers 54.4 million hectares in 2005, making up 60 percent of worldwide soybean production (US National Academy of Sciences, 2000).

1.2 Fluorescent marker gene in transgenic plant production

A successful transgenic product means that the inserted gene can be stably integrated without harmful effects to other plant functions, product quality or the ecosystem. Hence the transgene will be expected to active only at proper moments in the right locations. Therefore biologists intend to observe the activity of the introduced gene and test its effects as well as overall performance in the future. In order to make this possible, marker genes are added to the gene "construct" in order to identify plant cells or tissues that have been successfully targeted with transgene and to study gene expression. β-galactosidase, firefly luciferase, chloramphenicol acetyltransferase (CAT) and β-glucuronidase (GUS) are all marker genes. Among them, however, green fluorescent protein (GFP), isolated from the jellyfish Aequorea Victoria, is currently the most widely-used maker gene due to its stability at pH from 5 to 12 and temperature at 65°C (because its fluorophore locates at the centre of its barrel structure) and its small molecular size (27 k Da) (Lakowicz 1999). Moreover, GFP is the only marker gene that permits non-destructive and continual monitoring over time. For all other markers,
monitoring of expression results in tissue death, either from enzyme extraction or toxins in an in vitro assay mix. GFP fluoresces when illuminated with ultraviolet (395 nm) or blue light (475 nm) (Lakowicz 1999). Due to these advantageous characteristics, many GFP mutants have been developed for optimal expression in particular organisms, such as eGFP, sGFP(S65T), BFP, YFP, RFP, etc (Soboleski et al., 2005). In this thesis, the sGFP(S65T) was studied. It has a single excitation peak at 490 nm, is more resistant to photobleaching, and brighter than the original GFP (Chiu et al., 1996).

GFP has been applied in transgenic research to identify the transformed cells (Takada et al., 1997), to identify pathogens during infection, and to measure gene expression in vivo (Sheen et al., 1995; Chiu et al., 1996). In medical area, GFP has been applied to identify tumours and the progress of gene therapy (Rideout et al., 2002) and to track the location of drug to make sure it is delivered and takes effect at the right places (Watson et al., 2004). GFP fluorescence is a reliable and quantitative reporter of gene expression (Soboleski, 2005), allowing quantitative monitoring of transgene activities over time. Quantitative GFP information of gene expression can be obtained by direct measurement of the intensity of fluorescent emission (Buenrostro-Nava et al. 2005; Chiera et al. 2007). In most of these research mentioned above, fluorescence microscopic imaging instrument was the key instrument for GFP detection.

1.3 Fluorescence and fluorescence imaging microscopy

A nucleus in every atom is orbited by one or more electrons, and each of them has a certain energy level. The electron can absorb some photons to transition to a higher
energy level. When the electron returns back to its original energy level, a photon of light is emitted, and called fluorescence (Sharma and Schulman, 1999). The energy and wavelength of the emitted fluorescence depends on the excitation light which is absorbed by the electron, the electronic structure of the atom, and other environmental factors. The fluorescence wavelength will be longer (have less energy) than the absorbed excitation wavelength, and the difference between their central wavelength is called Stokes Shift (Lakowicz, 1983).

For the past three centuries, microscopes have been used to help see objects that cannot be seen with the eye. However, the microscope alone cannot preserve the image and some analysis cannot be conducted with only visual observation. Substantial efforts have been invested to make microscope even more useful with photographic and computer techniques, which facilitated automatic image acquisition, analysis of large volumes of data, and convenient storage of and access to the data. In transgenic research, automatic imaging microscopy has provided an effective means for investigation of reporter gene expression, transfection efficiency, and cellular fluorescence, etc. The major components of a typical automated imaging system are a light source, a filter wheel coupled with excitation and emission filters, a microscope with automated stage, a camera and a computer with image acquisition and analysis software (Wang and Herman, 1996).
1.4 Challenges of GFP expression quantification

Although GFP has been largely applied in medical area as a quantitative reporter gene, its utility in plant biotechnology has not been fully realized. Plants contain chlorophyll and other components which can interfere with GFP detection (Zhou et al., 2005). There are two possible reasons for this interference, one is that chlorophyll may compete for the excitation energy with GFP, and the other is that chlorophyll emission may mask the GFP fluorescence.

Another challenge comes from the imaging microscopic measurement system itself. In a fluorescence imaging microscopic system, the final detected signal intensity can be expressed in the form of

\[ FI_{\text{det}} = f((\varepsilon \cdot QY_f), [\text{dye}], I_{ex}, CE_o, TE_f, TE_m, QE_{\text{det}}, t) \]  

(1.1)

Where \( FI_{\text{det}} \) is detected fluorescence intensity, \( \varepsilon \) is molar extinction coefficient (the ability of a molecule to absorb light of a particular wavelength is characterized by \( \varepsilon \) ), \( QY_f \) is quantum yield of fluorescence, \([\text{dye}]\) is dye concentration, \( I_{ex} \) is excitation light intensity, \( CE_o \) is collection efficiency of the objective, \( TE_f \) is transmission efficiency of filters including emission and emission filters and dichroic mirror, \( TE_m \) is the transmission efficiency of the microscope (including the objective), \( QE_{\text{det}} \) is the quantum efficiency of the detector, and \( t \) is exposure time (Wang and Herman, 1996). Therefore, the strength of the detected signal is determined by the spectral quality of the microscope, objective, emission filter, dichroic mirror, and detector, other than the fluorescent strength of a sample itself. Hence, the selection of optical filters is critical for the GFP.
detection in order to eliminate unwanted spectral data and at the same time to enable stronger GFP fluorescent signal captured by the CCD detector.

The interference from chlorophyll can be reduced or eliminated by suppressing its emission or optimizing the fluorescence microscopic imaging system. During the course of image processing, segmentation is one of the most difficult problems in image processing, and its accuracy determines the eventual success or failure of image analysis.

1.5 Goal and objectives of the study

The overarching goal of this thesis is to improve the quantification of the green fluorescent protein (GFP) in transgenic embryogenic soybean tissues using biological inhibition and spectral feature selection approaches. This thesis is the first endeavour to accomplish this goal. In order to accomplish this goal, three objectives shown below were pursued:

1. To investigate a new approach of eliminating the chlorophyll interference using the herbicide isoxaflutole (IFT) and observe its effects on tissue growth.
2. To identify spectral features to improve quantification of GFP expression and to reduce chlorophyll autofluorescence interference.
3. To apply and evaluate the image processing algorithm to quantify the GFP.

1.6 Structure of the thesis

In the pages that follow, GFP expression in embryogenic soybean tissues will be quantified and two approaches will be employed to mitigate the influence coming from
chlorophyll: chlorophyll bleaching and spectral feature selection. Furthermore, three state-of-the-art image segmentation algorithms will be applied to segment the fluorescence images with low contrast between regions of interest and background, and the results will be compared with that of an open-source image processing software ImageJ.
CHAPTER 2

APPLICATION OF HERBICIDE IFT TO SUPPRESS
THE CHLOROPHYLL FLUORESCENCE

2.1 Introduction

The green fluorescent protein (GFP), derived from the jellyfish (*Aequorea victoria*), has been widely-used as a scorable marker to study gene expression across a variety of different tissues, of both plant and non-plant origin (Lakowicz, 1999). The wild type GFP emits a bright green fluorescence (509 nm) when illuminated with either ultraviolet (395 nm) or blue (475 nm) light (Heim et al., 1994). The fluorescence from GFP can be observed in real time using the proper instrumentation and the appropriate optical filters (Billington and Knight, 2001). GFP can not only provide qualitative information on spatial/temporal attributes of gene expression under right conditions, but also provide quantitative information of gene expression by direct measurement of the intensity of fluorescent light emission (Urwin et al., 1997; Buenrostro-Nava et al., 2005; Chiera et al., 2007).

In chlorophyll-containing tissues, GFP detection is often masked by the red fluorescence emitted from the large amounts of chlorophyll (Zhou et al., 2005). Chlorophyll has a maximum absorption at 488 nm and a fluorescence emission at 683 nm, large amounts of chlorophyll present in leaf tissue can present problems for
visualization and quantification of GFP. In transgenic medicago and rice leaves containing GFP, reduction of chlorophyll levels by either etiolation or treatment with 95% ethanol allowed GFP to be more easily detected (Zhou et al., 2005).

Isoxaflutole (IFT; 5-cyclo-propyl-1,2-isoxazol-4-yl ααα-trifluoro-2-mesyl p-tolyl ketone) is a pre-emergence herbicide used for control of broadleaf and grass weeds in fields of corn and sugarcane (Luscombe et al., 1995). Use of IFT on susceptible plants causes bleaching of the foliar tissue, resulting in stunting and eventual plant death. This process involves multiple steps beginning with the conversion of IFT to a diketonitrile derivative (DKN) that inhibits the activity of 4-hydroxyphenyl pyruvate dioxigenase (HPPD) (Pallet et al., 1998; 2001). Inhibition of HPPD results in the loss of plastoquinone production which reduces the activity of phytoene desaturase (Pallet et al., 1998; Norris et al., 1995). The reduction of phytoene desaturase activity, in turn, leads to phytoene accumulation, disruption of carotenoid pigment synthesis, and the inhibition of chloroplast development, causing the bleached appearance of plant tissues (Pallett et al., 1998; 2001).

Transgenic approaches have been used to generate plants with resistance to IFT through nuclear targeting of the transgene (Garcia et al., 1999; Matringe et al., 2005) and the production of transplastomic plants (Dufourmantel et al., 2007). In plant tissue culture, carrot cell suspension cultures have been used as a tool to study IFT metabolism (Pallett et al., 1998). IFT was also used as the selection agent to recover green transgene-containing callus in tobacco that over-expressed the Arabidopsis HPPD (Garcia et al., 1999). Studies have not yet been performed, however, on the effects of IFT and other
bleaching herbicides (Wakabayashi and Bögner, 2004) on the growth of plant tissue cultures and as a tool to minimize interference of plant pigments with GFP fluorescence.

The objectives in this study were:

1. To reduce tissue pigmentation and improve methods for quantification of low GFP expression in transgenic embryogenic soybean tissues;
2. To evaluate IFT effects on soybean tissue growth.

2.2 Materials and Methods

The biological experiments were conducted in the department of Horticulture and Crop Science (224 Williams Hall, Wooster Campus). Biological materials were kindly provided by the Plant Transformation Laboratory.

2.2.1 Plant material preparation

Embryogenic cultures of soybean (Glycine max (L.) Merrill cv. ‘Jack’) were initiated and transformed as previously described (Finer and McMullen, 1991) with some modifications. Two different plasmid DNAs, pHytru and pGmupri (2.5 µg each), were co-precipitated onto tungsten particles which were used for particle bombardment. The plasmid pHytru (Chiera et al., 2004) contained the hygromycin phosphotransferase gene, under regulatory control of the cauliflower mosaic virus 35S promoter (CaMV35S) while pGmupri (Chiera et al., 2007) contained the GFP gene (sGFP(S65T); Chiu et al., 1996) under regulatory control of a pre-intronic version of the soybean polyubiquitin promoter. For transformation, embryogenic tissues of soybean (Santarém and Finer, 1999) were subcultured one week prior to bombardment and placed in the center of a Petri dish
containing D20 medium (20 mg/L 2,4-D, Santarém and Finer, 1999). Immediately prior to bombardment, the Petri dish lid was removed for 15 min. to allow partial drying of the tissue in a laminar flow hood. Following bombardment, tissues were placed on D20 medium for one week, prior to placement under hygromycin selection as previously described (Santarém and Finer, 1999). Transgenic clones were recovered and separately cultured on D20 medium containing 30 mg/L hygromycin for a minimum of 2 months. After this time, transgenic tissues were maintained on D20 medium without hygromycin. For evaluation of GFP expression using IFT treatments, proliferating embryogenic tissue from three independent events representing low, low/moderate, and moderate levels of GFP expression were selected.

### 2.2.2 IFT treatment

For the evaluation of IFT on tissue growth and GFP expression, 15 uniform pieces of embryogenic tissue were selected from each of the three clones along with non-transformed soybean tissue (cv ‘Jack’). IFT (kindly provided by B. Pelissier, Bayer CropScience) was dissolved in DMSO and added to D20 medium (Santarém and Finer, 1999) prior to autoclaving to final concentrations of 0 mg/L (control), 3 mg/L and 10 mg/L. Five clumps of each clone and non-transformed ‘Jack’ tissues were placed on each medium in the same Petri dish. The Petri dish format was replicated a total of three times.

### 2.2.3 Image acquisition and processing

Petri dishes were marked to retain the same order of image collection using an automated image collection system (Buenrostro-Nava et al., 2005). The image collection
system consisted of a custom-designed 2-dimensional robotics platform for holding Petri dishes, along with a programmable spot CCD camera (SPOT-RT, Diagnostic Instruments Inc., Sterling Heights, MI) mounted on a Leica MZFLIII stereomicroscope (Leica, Heerbrugg, Switzerland). The stereomicroscope was equipped with a broad band-pass GFP2 filter set containing a 480 nm ± 20 nm excitation filter and a long-pass emission filter with a cutoff at 510 nm. The whole imaging system was placed in a laminar hood, in a temperature-controlled tissue culture room (25°C). Condensation on the lids of Petri dishes, which could interfere with image analysis, was controlled by using polymer discs in place of the standard Petri dish lid (Finer and Finer, 2007). Images were acquired weekly for four consecutive weeks, including an image collection at time 0. Between image collection time points, all cultures were maintained in the tissue culture room under a 16 hour light (40 µEm²s⁻¹) and 8 hour dark photoperiod.

The magnification of the stereomicroscope was fixed at 1X to retain the tissues within the field of view during the four-week experimental period. Images were acquired using a programmable color camera, which provided intensity information in the red, green, and blue (RGB) channels, under either blue light illumination (with the GFP2 filter set) or white light illumination (brightfield). For white light image collections, exposure times of the red, green and blue channels were 29, 39 and 151 milliseconds, respectively. Images collected under blue light for GFP detection had red, green and blue exposure times of 37, 17 and 10 seconds, respectively. For each weekly image acquisition, separate images under blue light and white light illumination were collected. The software package ImageJ was utilized for image analyses (Rasband, 1997-2006).
In order to evaluate the effects of IFT on the tissue growth, tissue size was determined using the white light images. The number of pixels in the 2-D projection of the 3-D tissue volume was used to estimate tissue growth. For tissue area values, RGB color images were first converted to grey-scale images. Tissue blobs were then segmented from the background (i.e. growth media background) in each image by adjusting the grey level intensity threshold value (typically 18-31) to determine the number of pixels in the tissue area. The percent relative growth of tissue at each time point was calculated by subtracting the initial area of the tissue from the measured area at each time point, dividing by the initial area and then multiplying that value by 100.

The effects of IFT on GFP detection were quantified using the color images collected under the blue light. To obtain a high contrast GFP signal, the total red channel value was subtracted from the corresponding total green channel value, then divided by the pixel number of the tissue area to generate the $\frac{G-R}{R}$ value. The $\frac{G-R}{R}$ value was calculated only for the tissue area, as determined above. The relative increase in GFP expression was calculated by subtracting the $\frac{G-R}{R}$ at time 0 from the $\frac{G-R}{R}$ at each time point for each set of collected images.

2.2.4 Recovery from IFT treatment

After tissues were grown on the IFT-containing media for 4 weeks, tissue pieces were subcultured and placed back on D20 medium without IFT. After two weeks, brightfield images of the embryogenic tissues were collected. Recovery of embryogenic
tissue from the IFT treatments was evaluated by the re-growth of tissues having green pigmentation.

2.3 Results and Discussion

The calculated relative growth for each clone showed that IFT had no significant effect on growth of transgenic and non-transformed D20 tissue (Fig. 2.1). Tissue growth rates on media containing either of the IFT levels were unchanged relative to the medium without IFT. Over the 4-week time course of the experiment, the tissues increased in size by approximately 50% every week, which is relatively slow for tissue cultured cells in general (Gamborg et al., 1968) but consistent for slow-growing embryogenic soybean tissues (Finer and Nagasawa, 1988). In addition, there were no significant differences between the relative growth of the control and the two IFT treatment groups (p=0.41, 0.97, 0.55 and 0.84 for control, low, low/moderate and moderate expression groups respectively by one-way ANOVA). Fresh weights of excised cucumber cotyledons were similarly unaffected (Kushwaha and Bhowmik, 1999), even though IFT levels were 36-fold higher than those levels used in the present study.
Figure 2.1: Growth comparison of non-transformed control, and low, low/moderate and moderate-expressing GFP clones cultured on the media containing 0, 3, or 10 mg/L IFT.
From visual observation of tissues grown on media containing the different levels of IFT, obvious differences were easily discerned between 0, 3 and 10 mg/L IFT-treated groups (Fig. 2.2). When embryogenic soybean tissues were cultured on media with IFT, the tissues gradually lost their green pigmentation and became “bleached” (Fig. 2.2). When viewed under blue light, less red fluorescence from chlorophyll was observed and more green fluorescence from GFP was visualized. From a purely visual perspective, IFT enormously reduced production of plant pigments, which interferes with GFP detection (Zhou et al., 2005). More extensive analyses were undertaken to further evaluate the effect of IFT on quantification of GFP expression.
Figure 2.2: Comparison of the pigment reductions, and the improved GFP detection by culturing the moderately expressing GFP clone on medium containing either 0 (A, B), 3 (C, D), or 10 (E, F) mg/L IFT. Images of tissues were collected (left to right) at 0, 1, 2, 3 and 4 weeks after IFT treatment. Top panels (A, C, and E) show images collected under white light. Bottom panels (B, D, and F) show images of the same tissue taken under blue light for GFP expression observation.
At each time point, the mean of the differences between green and red grey-channel values ($\overline{G-R}$) was calculated from the total tissue area, normalized based on the $\overline{G-R}$ measurement at time 0 and reported as GFP expression. Inclusion of IFT in the medium, used for the growth of embryogenic soybean tissue, reduced chlorophyll fluorescence and tremendously improved GFP detection (Fig. 2.3). Even with the appropriate filters, interference of chlorophyll with GFP detection could occur through an overabundance of red fluorescence or competition between GFP and chlorophyll for the same excitation wavelengths of light (Zhou et al., 2005). Zhou et al (2005) further showed that titration of a GFP solution with chlorophyll resulted in a gradual loss in GFP detection. Although the mechanism for interference of chlorophyll with GFP detection is not precisely known, it is clear that GFP can be difficult to visualize and evaluate in green plant tissues. In this study, GFP expression in a low-expressing clone was easily observed and quantified when tissues were grown on media containing IFT (Figs 2.2 and 2.3). For the 3 clones, which displayed low through moderate GFP expression, there were no significant differences in GFP expression in tissues cultured on media containing either of the 2 IFT concentrations. GFP detection levels in tissues cultured on the medium without IFT, however, was significantly different from detection levels in tissues cultured on the medium with the two IFT treatments.
Figure 2.3: GFP expression in tissues of control (non-transformed) and the three transgenic clones with varied expression levels from low to moderate that were grown on media containing 0, 3 or 10 mg/L IFT.
A comparison of GFP detection across clones and the non-transformed control, clearly demonstrated a remarkable improvement in GFP detection following culture of the tissue in media containing IFT (Fig. 2.4). For all tissues grown on the medium without IFT, the measured GFP expression was generally quite low over time with a moderate decline in GFP expression from the second time point. The decline in the measured GFP expression may be due to the greening of tissues, resulting in increased mean red values, from greater chlorophyll fluorescence. With constant green values as the experiment progressed, the increasing red values eventually exceed the constant green values resulting in a negative GFP expression (Fig. 2.4). However, the differences in GFP detection were clear for the control and the transgenic tissues grown in the presence of IFT. Each clone had a distinct level of expression and GFP was easily identified and quantified in the IFT-treated tissues (Fig. 2.4).
Figure 2.4: GFP expression in control (non-transformed) tissues and the three independent GFP expressing clones, grown on medium containing either a) no IFT or b) 10 mg/L IFT.

After transferring the IFT-treated “bleached” tissue pieces onto D20 medium without IFT, new growth appeared bright green within two weeks (Fig. 2.5). Apparently, the growth of the embryogenic soybean tissue was not negatively affected from residual IFT during the recovery process. The effects of the isoxaflutole were therefore fully reversible in soybean tissues, at least when used at the tested concentrations. Plants differ in their susceptibility to IFT and tolerance appears to be dependent on the ability to detoxify DKN (Pallet et al., 1998; 2001). Exposure to as little as 0.5 mg/L IFT resulted in reduced shoot growth in *Phalaris* whereas wheat shoots were unaffected (Kaur et al., 2004). Even in a tolerant plant such as corn, there was a wide range of tolerance observed
between hybrids (Simmons and Kells, 2003). Therefore, the IFT concentrations used for this type of study must be empirically determined for the specific tissue application.

Figure 2.5: Recovery of tissue following IFT treatment: non-pigmented tissue as result of the treatment (left) and green coloration in new growth, 2 weeks following subculture of the tissue onto a medium without IFT (right).

2.4 Conclusions

Use of IFT as a medium addendum provides a valuable tool for chlorophyll elimination in plant tissue cultures, allowing easier visualization of GFP expression in plant tissues. This is especially useful in situations where GFP expression is low and chlorophyll content is high. GFP detection was similarly improved by short-term (6 h) ethanol extraction of leaf tissues of transgenic medicago, rice and Arabidopsis (Zhou et al., 2005). Unfortunately this ethanol treatment is toxic to plant tissues, preventing
continued observation GFP expression in those tissues. Other herbicides and inhibitors that reduce chlorophyll content in green plant tissues (Wakabayashi and Bögner, 2004) may have similarly useful applications. In addition, the effects of IFT are fully reversible with rapid chlorophyll production within two weeks after transferring tissues to an IFT-free medium. The use of IFT proved to be effective in eliminating pigment interference and should be of interest to those working with GFP or other visible reporter genes in plant tissues.
CHAPTER 3

SPECTRAL FEATURE SELECTION FOR THE FLUORESCENCE MICROSCOPIC SYSTEM

3.1 Introduction

As stated in the previous chapters, GFP can be used as a valuable tool to study gene expression in a variety of heterogeneous tissues. However, for plants, GFP is still a relatively new reporter gene and much remains unclear about its activity. Accurate quantification of its expression intensity is useful for many applications, such as detection of promoter activities, visualization of protein targeting, and extraction of transformation efficiency, etc. Due to the interference from chlorophyll, GFP is not easily visualized and accurately quantified. In chapter 2, IFT was shown to be an effective “bleaching” medium addendum. Although IFT has no obvious negative influence on plant tissue in culture, it needs extra procedures to apply and may be potentially toxic to plant crops. If some modifications to the microscopic instrument improves the GFP signal strength relative to non-expressing area and blocks unwanted chlorophyll fluorescence, it would be useful for improving quantification of GFP expression.

The use of efficacious promoters is needed for successful application of transgenic technologies to crop improvement. Strong promoters are needed to promote the expression of marker genes in order to identify transformed cells and transgenic
plants during the transformation process. GFP is beginning to be used for in vivo promoter analysis. Nehlin et al. (2000) fused several promoters to sGPP-S65T and bombarded the complex into Brassicanapus microspores to aid in determining which promoter might be optimal for the development of a microspore transformation method (Nehlin et al., 2000). To quantitatively analyze activities of several promoters, GFP fluorescence intensity was observed and transformation frequencies was examined for evaluation of enhanced CaMV 35S promoter, CaMV 35S, the ACT2 promoter of Arabidopsis thaliana, CsVMV, and an enhanced version of the CsVMV promoter in transgenic grape plants (Li et al., 2001). Tang and Newton quantitatively analyzed promoter properties by GFP fluorescence from transgenic cell lines (Tang and Newton, 2004). Expression of the GFP gene, under regulatory control of either the constitutive 35S promoter or the developmentally-regulated lectin promoter was monitored and quantified using a newly-developed automated tracking system (Buenrostro-Nava, 2006).

For chlorophyll interference problem, several reports could be found about a GFP fluorescence decrease with leaf maturation from a similar 35S-mGFP5er construct in transgenic tobacco (Harper & Stewart, 2000) and oilseed rape (Halfhill et al., 2003). Zhou et al. suggested that chlorophyll was a major culprit in the decreasing of GFP fluorescence as leaves matured (Zhou et al., 2005). The novel finding of the interference of chlorophyll with GFP fluorescence revealed that the application of GFP to plant research deserved careful evaluation. Chlorophyll b exhibited a strong negative effect on GFP fluorescence, while a milder effect was caused by chlorophyll a. As chlorophylls,
especially chlorophyll b, absorb at the excitation wavelength (488 nm) of GFP, which makes them compete with GFP for the excitation light (Zhou et al., 2005).

In a microscopic fluorescence instrument, an emission filter is necessary for fluorescence observing. A 515 nm long-pass emission filter was usually used with epifluorescence microscopes by most researchers to study the GFP fluorescence. In using such arrangements, studies have shown that the background fluorescence interfered with observing GFP (Haas et al. 1996; Elliot et al. 1999). Stewart (2001) proposed that GFP observation was significantly improved by choosing the emission filter with a narrower band width. A later study implemented this theory: the GFP was extracted by a GFP-Meter through a 530/35 nm filter, and chlorophyll fluorescence was obtained using a 680/35 nm filter (Millwood et al., 2003). Zhou et al. used a Zeiss SV11 stereomicroscope to record green fluorescence in leaves of plants transgenic for 35SmGFP5er and found that the autofluorescence detected in the control leaves using a 500-nm emission filter was essentially eliminated when a 525-nm emission filter was used (Zhou et al., 2005). Although these studies explored the application of emission filters with different spectral features for the GFP observation, rare research evaluated the filter performance at the aspect of GFP image quality.

3.1.1 GFP and chlorophyll spectral features

The sGFP (S65T), a variant of GFP, was used in this research. Excitation and emission spectra of GFP in Fig. 3.1 show an absorption maximum at 489 nm and a very sharp emission spectrum with a peak at 509 nm.
Chlorophyll can be excited by the same excitation light (especially chlorophyll b) as that of GFP, and the chlorophyll emission peak is at 683 nm (Fig. 3.2). There are two possible reasons for interference of chlorophyll with GFP detection, one is that chlorophyll may compete for the excitation energy with GFP, and the other is that chlorophyll emission may overlap with the GFP fluorescence. The herbicide IFT can resolve the problem from the perspective of excitation competition, while spectral feature selection can analyze and block the interference from chlorophyll fluorescence, even though GFP expression intensity cannot be enhanced.

Although GFP emission is sharp around its peak, it has a tail at the wavelength larger than 600 nm (Fig. 3.1), which overlaps with the chlorophyll’s left emission.
distribution tail (Fig. 3.2). This interference from chlorophyll can make it difficult to extract the GFP signal especially when it is weak, if a broad band-pass filter is used. Spectral feature analysis of the current fluorescence imaging system and appropriate spectral filter selection was proposed to suppress the interference from chlorophyll autofluorescence and to improve the contrast between GFP and non-GFP in the plant tissue for accurate GFP quantification.

Figure 3.2: Excitation and emission spectra of chlorophyll a and b. (http://icecube.berkeley.edu/~bramall/work/astrobiology/fluorescence.htm)
3.1.2 Spectral filters in a fluorescence microscopic system

An optical filter is a device which selectively transmits light of a particular range of wavelengths, while blocking the remainder. Spectral band selection techniques have been applied in hyperspectral remote sensing for material identification (Keshava, 2004), product sorting (Haff and Pearson, 2006), soil characterizing (Henderson et al., 1989; Coleman et al., 1991; Csillag et al., 1993), in which, several sets of filters with distinct central wavelength and band width were used to extract the features of the objects. The filters, which can maximize the separation between the spectral classes, were selected. Spectral selection techniques (Backer et al., 2005; Skurichina et al., 2006) were used for dimensionality reduction and selection optimization based on large amount of spectral data.

However, in fluorescence microscopy, an excitation filter, an emission filter, and a dichroic mirror are necessary either for GFP fluorescence detection by camera or observation. Dichroic mirrors (or dichromatic beam splitters) are short-wave reflecting filters, which are used in fluorescence microscopy to split the light source into two light beams and excitation light is reflected onto specimens while GFP emission light is allowed to pass through (Fig. 3.3). Excitation filters and emission filters are band-pass filters. Band-pass filters can be roughly divided into broadband-pass filters and narrowband-pass filters. There is no definite boundary between these two types of filters. In this study, a broadband-pass filter was a band-pass filter, having the bandwidth of 20% or more of transmission band central wavelength (Macleod, 2001). The spectral features
of both a narrow band-pass filter and a broad band-pass filter were analyzed along with the spectral properties of the microscopic imaging system.

Figure 3.3: Light path inside a fluorescence microscopic imaging system.

In a fluorescence spectroscopic system (Fig. 3.3), the light path can be described as: first, an excitation source sheds light on a tissue specimen which contains GFP after passing through an excitation filter and being reflected by a dichroic mirror; then the photons emitted by GFP and other tissue components along with reflected blue excitation light reaches at the dichroic mirror, which reflects most of the blue excitation light; then the remaining blue light, GFP fluorescence and other autofluorescence pass through an
emission filter, which can block either only excitation light (broad band-pass filter) or exclusively extract the GFP fluorescent photons (narrow band-pass filter); at last the photons are collected by a detector. In the fluorescence detecting process, the energy collected by a detector can be expressed by

\[ F_{\text{det}} = F_{\text{sample}} \cdot CE_{\text{obj}} \cdot TE_{\text{filters}} \cdot TE_{\text{mic}} \cdot QE_{\text{det}} \]  

(3.1)

Where \( F_{\text{det}} \) is detected fluorescence intensity, \( F_{\text{sample}} \) is fluorescence intensity from samples, \( CE_{\text{obj}} \) is the collection efficiency of the objective, \( TE_{\text{filters}} \) is transmission efficiency of filters including emission filter and dichroic mirror, \( TE_{\text{mic}} \) is the transmission efficiency of the microscope (including the objective), and \( QE_{\text{det}} \) is the quantum efficiency of the detector (Wang and Herman, 1996). Therefore, the strength of the detected signal is determined by the spectral quality of the microscope, objective, emission filter, dichroic mirror, and detector, other than the fluorescent strength of a sample itself. The spectral band width of the emission filter is flexible as long as it covers the GFP emission peak. However, for the purpose of extracting GFP expression and eliminating chlorophyll emission, the objective in this chapter was to identify spectral features that could be implemented to improve quantification of GFP expression, specifically, to reduce chlorophyll autofluorescence interference. It was desirable to find a spectral band width that could eliminate detection interference from chlorophyll autofluorescence, and to provide good contrast to identify the plant tissue from its growth media background in the images.
3.2 Materials and Methods
3.2.1 Data collections and measurements of spectral features

The transmittance of the objective and the collection efficiency of the camera were kindly provided by their manufacturers Leica Microsystems (Bannockburn, IL) and Diagnostic Instruments Inc. (Sterling Heights, MI), respectively. In practice, the actual continuous emission spectra $x(\lambda)$ of GFP and chlorophyll are not known, thus the measurements of a spectroradiometer were used instead.

The collection efficiency $CE$ of objective lens can be calculated as

$$CE = NA^2 \cdot T_L$$  \hspace{1cm} (3.2)

where $NA$ is the numerical aperture and $T_L$ is the transmittance of the objective lens (Hejazi and Trauernicht, 1996). The numerical aperture of objective lens is defined by

$$NA = n \cdot \sin \theta$$  \hspace{1cm} (3.3)

where $n$ is the index of refraction of the medium in which the lens is working (1.0 for air, 1.33 for pure water, and up to 1.56 for oils), in this case, $n$ is 1.0. And $\theta$ is the half-angle of the maximum cone of light that can enter or exit the lens. A lens with a larger numerical aperture will be able to visualize finer details than a lens with a smaller numerical aperture. Lenses with larger numerical apertures also collect more light and will generally provide a brighter image. The numerical aperture of the Planapochromatic objective 1x is 0.125 (Leica Microsystems, 2003). Then, the collection efficiency is 0.016 times transmission efficiency of the objective, from which the collection efficiency should be very small, but this small value will not have an effect on the spectral band selection. The most uncertain value in Equation (3.1) is the microscope transmission
efficiency. Generally transmission efficiency of the microscope is not provided by manufacturers (Wang and Herman, 1996) and only the transmission of the objective was considered instead.

To acquire the spectral features of chlorophyll emission and two spectral filters, a portable spectroradiometer (ASD, FieldSpec Pro FR, Boulder, CO) was used for the measurements. The spectrum of this spectroradiometer covers from the visible to near infrared band, about 350 to 2500 nm, although only visible light was used in this study (about 350 to 800 nm) because GFP and chlorophyll emissions were within this range. The sampling interval (the spacing between sample points in the spectrum) and the spectral resolution (the full-width-half-maximum (FWHM) of the instrument response to a monochromatic source) of the spectroradiometer are 1.4 nm and 3 nm respectively for the region 350-800 nm. In the spectroradiometer software, the parameters (sample, white reference and dark current) were set up as described in the user manual. For a quick self-checking, the sample, white reference and dark current averages were set to 10 scans in the control configuration window. Then the instrument was optimized automatically by clicking the command button “OPT” in the Toolbar of the software control panel. Since the measurement was conducted in the laboratory, white reference and dark current averages were set to 60 as the user manual suggested for indoor illumination. Before measuring samples, the spectroradiometer was calibrated with a white reference. When observing samples, the spectrum average was changed to 30 and the spectral saving interval was set to 5 seconds for each sample, totally 10 spectra were saved and averaged later.
A DC powered light, provided by the manufacturer of the spectroradiometer, was used to maintain stability of the excitation light source. A white reference was used for calibration before all experiments, as shown in Fig. 3.4. The DC light source and the optical probe were maintained the same distance and 45° angle to the white standard from left and right sides respectively. The reflectance of the white reference should be kept oscillating around 1(Fig. 3.5).

To detect GFP and chlorophyll fluorescence in embryogenic D20 soybean tissues, embryogenic cultures of soybean (Glycine max (L.)) were initiated and transformed as previously described (Finer and McMullen, 1991) with some modifications. D20 tissues driven by the CaMV35S promoter (Odell et al., 1985) containing high GFP expression were provided by the Plant Transformation Laboratory, Wooster, Ohio. The GFP signal was too weak to be detected by this spectroradiometer because of its limited resolution although the GFP expression in those tissues was high; instead, typical GFP fluorescence signature was obtained from Molecular Probes, Inc. (Fig. 3.1). The spectral features of the chlorophyll emission from these D20 tissues were detected using this spectroradiometer. The tissues could not be kept horizontally underneath the excitation light since the incident angle should be 45°. The excitation light was fixed, therefore, the Petri dish holding the tissues was kept an angle of 45° with the excitation light, and the probe was held with the same angle and distance from another side. The probe and the excitation light should be placed very close to the tissue sample because the fluorescence was very weak. However, the reflected blue excitation light also entered the probe along with chlorophyll emission. Although the central wavelength of the blue light was far
away from the one of the red chlorophyll fluorescence, the chlorophyll fluorescence was obtained by subtracting the tissue growth media reflected blue light, in case that there is a little overlap between their tails.

The spectroradiometer was also utilized to measure the spectral band of the spectral filters. The DC light source was used considering its capability of providing stable light signal for the spectroradiometer. This experiment was conducted in an experimental hood covered with black cloth to avoid stray light.

At first, the DC-powered light source was fixed onto a stand with a downward incident angle of 45°. The radiance of the light source was measured by placing the optical probe of the spectroradiometer in line with the light source. Then the radiance of the light passing through a filter was measured by attaching a filter in front of the optical probe (Fig. 3.6). The transmission efficiency was calculated as the percentage of the radiance passing through the filter.

Figure 3.4: Schematic of taking white reference before measurements, the probe and the light source should be kept the same angle (45°) from the white standard.
Figure 3.5: The spectroradiometer was calibrated by measuring reflectance of a white standard, which should be a complete reflectance.

Figure 3.6: Schematic of the measurements of transmission efficiency of the spectral filters, the filters was placed in front of the probe.

A planapochromatic objective 1x was mounted on the current fluorescence microscopic imaging system, whose transmission efficiency is greater than 90% in the range of 450-700 nm (Fig. 3.7).
Figure 3.7: Transmission efficiency of the planapochromatic objective 1x (Leica Microsystems Ltd.)

A CCD camera (Spot RT Slider, Diagnostic Instruments Inc.) with a monochrome sensor (KAI-2000-M, Kodak) was used to detect the GFP fluorescence. Its quantum efficiency is shown in Fig. 3.8. The transmission efficiencies of the three color filters are very low, in the same way of the collection efficiency of objective lens, which captures very weak signal, but does not affect the spectral filter selection.
Figure 3.8: Quantum efficiency of the sensor of the fluorescent imaging system. The monochrome CCD sensor has quantum efficiency of approximately 45% in the GFP expressing band (around 510 nm), and less than 30% in the chlorophyll fluorescence range (600-800 nm). As results of color filtering, quantum efficiency of the sensor reduced to 12% (green channel) and 5% (red channel) for the detection of GFP expression and chlorophyll fluorescence signals.

3.2.2 Spectral feature selection

To select a proper emission filter band to extract GFP signals exclusively using the automatic microscopic imaging system, the spectral quality of GFP and chlorophyll fluorescence signals in the red, green and blue channels of the camera were estimated. During this process, the transmission efficiency of the dichroic mirror was ignored.
because a dichroic mirror had a high and invariable transmittance in the emission band range, which only scaled down the signals uniformly, but did not change the spectral features of the signals. As the equation 3.1 described, the detected signal could be calculated by the products of fluorescence relative intensity of GFP and chlorophyll, with the collection efficiency and transmission efficiency of the objective, with the quantum efficiencies of the CCD sensor mounted with three color filters.

The subsampling interval of the spectroradiometer data was 1 nm in the range of our interest, but that of the graphs provided by companies was 10 nm (Figs 3.7 and 3.8). To facilitate the spectral selection, the spectroradiometer data were averaged across five samples and the data provided were subsampled at each 5 nm. GFP fluorescence and chlorophyll emission were multiplied by the collection efficiency of the objective, transmission efficiency of the objective, quantum efficiencies of the CCD sensor mounted with three color filters to estimate the detected GFP and chlorophyll signals as Equation 3.1 stated. According to this estimated spectral band, the band range of an emission filter was determined for exclusive extraction of GFP signals.

3.2.3 Spectral feature evaluation

Two typical filter sets for sGFP (S65T) detection are commercially available from Leica Microsystems: GFP Plus fluorescence (GFP2, excitation filter 480/40nm, broad band-pass emission filter, 510-730 nm), and GFP Plant fluorescence (GFP3, excitation filter 470/40 nm, band pass emission filter 525/50 nm). Their performances were evaluated with the parameter of separability of the images taken through these two filters:
the separability between GFP expressing and non-expressing area of the plant tissue, and the separability between the tissue and its growth media background.

To evaluate the two filter sets, the images of the same 3 clones and non-transformed ‘Jack’ of D20 tissues were acquired as stated in Chapter 2. GFP2 filter set and GFP3 filter set were sequentially used in the last week for evaluation of these two filter sets. Thus for each piece of tissue, 3 images were captured sequentially under white light, GFP2 excitation (480/40 nm) and GFP3 excitation (470/40 nm). Totally 60 GFP3 images and 60 GFP2 images of the 20 herbicide-free, 20 3mg/L IFT-treated, and 20 10 mg/L IFT-treated tissues were obtained. The images were analyzed to get the two separabilities as described below.

In terms of pixel brightness distribution, the pixels in a GFP fluorescence image could be classified as tissue area and growth media background; within the tissue area, pixels could belong to GFP-expressing area or non-expressing area (Fig. 3.9). It was important to separate the tissue area from its growth media background, and isolate the GFP-expressing area from non-GFP expressing area in the tissue for quantification of GFP expression. High contrast between the tissue area and its growth media background, and high contrast between GFP-expressing area and non-GFP expressing area simplify image processing procedure. In this study, separability was defined to evaluate the two contrasts in the images taken through the selected GFP emission band. It was calculated by the distance between the brightness modes ($mode_1$ and $mode_2$) of two comparative distributions divided by the sum of their standard deviations ($\sigma_1$ and $\sigma_2$):

$$\text{separability} = \frac{\text{distance}}{\sigma_1 + \sigma_2} = \frac{mode_2 - mode_1}{\sigma_1 + \sigma_2}$$

(3.4)
Figure 3.9  A histogram of a typical GFP fluorescence image. The two thresholds separate the histogram into three subdistributions: distributions of growth media background and tissue area, including GFP-expressing and non-expressing tissue areas. Separability was defined as the distance between two modes of two subdistributions brightness divided by the sum of their standard deviations of these two subdistributions.

3.3 Results and Discussion

3.3.1 Chlorophyll emission spectrum

The measured spectral band of chlorophyll in embryogenic soybean D20 tissues ranges from 600 to 800 nm, with a peak at 683 nm (Fig. 3.10). It is a combination of chlorophyll a and chlorophyll b (Fig. 3.2).
Figure 3.10: Detected chlorophyll fluorescence spectrum in D20 soybean tissues with the spectroradiometer.

3.3.2 Spectral feature selection for emission filter

To evaluate relative strength of various spectral bands for the GFP expression measurement, the GFP and chlorophyll fluorescent emission signals were multiplied with the collection efficiency and transmittance of the objective and the quantum efficiencies of the CCD sensor mounted with three color filters. After the filtering of the camera color filters, GFP fluorescence could be detected in the red, green and blue channels, even though it was very weak in the red channel, whereas signal from chlorophyll fluorescence was mostly in the red channel and a small part in the green channel (Fig. 3.11). Proper emission filter selection could extract most of the GFP emission energy, while excluding chlorophyll and other autofluorescence. The excitation filter band should cover the
absorption peak of GFP and there should be a distance of at least 10 nm between its higher limit and the lower limit of the emission filter. The dichroic mirror cutoff should be between these two limits. Considering the function of a dichroic mirror, the signal below its cutoff should be zero; therefore, the lower limit of the emission filter should be equal or greater than the dichroic mirror’s cutoff and covers as much GFP signal as possible. The higher limit of the emission filter should be less than 590 nm, where the chlorophyll signal started (Fig. 3.11), to block chlorophyll emission. Thus, the range of 500-590 nm was recommended for an emission filter, and approximately more than 85% of GFP fluorescent signal distributed within this range.

There are three types of emission filters commercially available (Leica, Heerbrugg, Switzerland) for GFP detection: a broad band-pass filter, 510-730 nm (GFP2); a narrow band-pass filter, 500-550 nm (GFP3); and another narrow band-pass filter, 510-530 nm (GFP1). The narrow band-pass GFP3 filter band fell into the selected band (500-590 nm) and approximately 75% of GFP fluorescent signal distributed within the range of 500-550 nm, which met the criteria of the spectral feature selection, which was covering as much as GFP emission band including its peak and blocking chlorophyll emission.
Figure 3.11: Detected GFP and chlorophyll signals in blue, green and red channels of the camera (without filtering by a dichroic mirror and an emission filter). Here the data is from the products of GFP fluorescence data from Molecular Probes, Inc. (Fig. 3.1) and chlorophyll fluorescence data detected in soybean D20 tissues transformed with 35S-sGFP (S65T) (Fig. 3.10) with the collection efficiency and transmission efficiency of the objective as well as the quantum efficiencies of the CCD sensor mounted with three color filters.
3.3.3 Spectral feature evaluation

Customized emission filters with the identified spectral band above were considered for the GFP expression measurement. Nonetheless, commercial filters with similar spectral characteristics were readily available, thus, the spectral feature were evaluated using GFP2 and GFP3 filters provided by the fluorescent microscope provider (Leica, Heerbrugg, Switzerland). Spectral characteristics of the GFP2 and GFP3 were depicted in Figs 3.12 and 3.13, respectively. To get the spectra passing through these two filters, their transmittances were multiplied by the spectra in Fig. 3.11. Results are shown in Figs 3.14 and 3.15. As shown, the GFP3 narrow band-pass filter blocked the chlorophyll signal while attenuating the red channel of GFP expression (Fig. 3.14). The GFP2 broad band-pass filter (510-730 nm) allowed all signals including GFP and autofluorescence to pass through (Fig. 3.15).
Figure 3.12: Transmission efficiency of the GFP3 emission filter.

Figure 3.13: Transmission efficiency of the GFP2 emission filter.
Figure 3.14: Effects of the GFP3 narrow band (500-550 nm) emission filter on the fluorescent signal strength in red, green, and blue channels of the image acquisition system. GFP fluorescence passed through the emission filter, while chlorophyll fluorescence was blocked.
Figure 3.15: Effects of GFP2 emission filter on the fluorescent signal strength in blue, green and red channels of the image acquisition system. Both GFP and chlorophyll fluorescence transmit through the filter. As shown in Fig. 3.14, GFP3 rejected chlorophyll fluorescence beyond 550 nm that made distinguishing GFP expression from other fluorescence sources straightforward. On the other hand, from the images taken through the GFP2 filter, both GFP and chlorophyll fluorescence were collected and the combined signal made it difficult to discern the GFP expression. Although narrower emission filter blocked autofluorescence to be detected by the camera, it attenuated GFP signal, even though the attenuation was minor.
To quantify how well the spectral features can help to identify the objects in the images, the separability between background and tissue area, and the separability between GFP-expressing and non-expressing tissue areas were computed for GFP2 images and GFP3 images of 20 IFT-free (Fig. 3.16a), 20 3 mg/L IFT-treated (Fig. 3.16b) and 20 10 mg/L IFT-treated soybean tissues (Fig. 3.16c). In all three groups, the separability between background and tissue area of GFP2 images was significantly higher than that of GFP3 images and the separability between GFP-expressing and non-expressing area of GFP2 images was significantly lower than that of GFP3 images by two-sample t-tests at a confidence level of 95%.

The separability between the tissue and its growth media background represented the distance between their brightness distributions, or represented the contrast between the tissue area and its growth media. Broad band-pass GFP2 filter could provide a better contrast between the tissue area and its growth media background because most of GFP and chlorophyll fluorescence signal was extracted. However, narrow band-pass GFP3 filter only extracted most GFP fluorescence signal. Similarly, the separability between GFP-expressing and non-expressing area represented the contrast between them. Because of the narrow band of GFP3 filter, only GFP expression area displayed bright color, while chlorophyll fluorescence was blocked and non-expressing area was dark, which presented a higher contrast between the GFP-expressing and non-expressing area than GFP2 images.
Figure 3.16: Comparisons of separabilities between background and tissue areas (separability1) and between GFP-expressing and non-GFP expressing areas (separability2) of GFP2 and GFP3 images. Both separability1 and separability2 were significantly different in the images of (a) IFT-free transgenic soybean tissues, (b) 3 mg/L IFT-treated transgenic soybean tissues, (c) 10 mg/L IFT-treated transgenic soybean tissues, taken through GFP2 and GFP3 filters.
For a direct viewing of the difference of GFP2 and GFP3 filters, both GFP2 and GFP3 images were segmented with double brightness global thresholding in Matlab (The MathWorks Inc., 2005) to isolate the GFP expressing area. For GFP2 images, due to the interference of chlorophyll fluorescence, the GFP expression area could not be identified explicitly with the double thresholding approach. But for the GFP3 image, the GFP expression area could be separated with an optimal brightness threshold of 0.2 (in Matlab, the double class assigns a floating number between 0 and 1 to each pixel of gray-scale images) (Fig. 3.17).
Figure 3.17: Segmentation results of the GFP-expression area in GFP2 and GFP3 images of a piece of transformed soybean D20 tissue with both GFP and chlorophyll fluorescence. Red circle represents the interference of chlorophyll fluorescence. The optimal brightness threshold for GFP2 image is 0.25, and 0.2 for the GFP3 image. a, GFP2 image; b, segmented GFP-expressing area from the GFP2 image; c, GFP3 image; d, segmented GFP-expressing area from the GFP3 image.

Effects of the selected spectral features on the tissue analysis were demonstrated using image analysis. Images were first acquired using the GFP2 and GFP3 emission filters, followed by selecting a threshold value for each of the images before performing a segmentation operation on the green channel images to highlight the GFP expression area. As shown in Fig. 3.17, the colored GFP2 tissue image (Fig. 3.17a) contained both green (mostly GFP expression) and red (mostly chlorophyll autofluorescence) regions while the GFP3 image (Fig. 3.17c) showed only green. After threshold value selection and segmentation operation, a pair of binary images (Figs 3.17b and 3.17d) showed blobs that
represented the GFP expressing areas. As indicated by the red circle in the segmented GFP2 image, part of the non-GFP expressing area was mis-identified as the GFP expressing area. From Fig. 3.17a, yellowish fluorescence could be seen in the faint area, but not bright green GFP fluorescence. It is probably chlorophyll fluorescence because red color displayed yellowish in all the soybean tissue images taken by the Spot RT camera. In Fig. 3.17c, this faint area did not take on discernable fluorescence. In one word, chlorophyll fluorescence could lead to misclassified GFP-expressing area.

Results showed that the broad band-pass filter gave good contrast between the tissues and their background, and the narrow band-pass filter allowed only GFP fluorescent signal to pass through at the cost of a lower contrast between the tissues and their background. However, some GFP expression features quantification is based on the whole tissue area, which can only be extracted easily if there is a good contrast between the tissue area and the medium background.

### 3.4 Conclusions

Spectral features selection could provide good contrasts between GFP fluorescence and autofluorescence from chlorophyll of the plant tissue. The detected GFP signal mainly distributed within the band range of 500-590 nm, and chlorophyll fluorescence started from 590 nm. Therefore, the spectral band of any filter falling within this range and covering most GFP signal would be suitable for exclusively extracting GFP signals.

By evaluation results of the broad band-pass filter and the narrow band-pass filter, the narrow band-pass filter was proved to have improved the contrast between GFP.
expressing and non-expressing area, but it reduced the contrast between tissue area and
background because it attenuated the signal strength and the image brightness was much
lower than the images taken by a broad band-pass filter.

In sum, the broad band-pass filter is suitable for observation of GFP expression,
while the narrow band-pass filter allows accurate quantification of GFP expression
without interference from chlorophyll.
4.1 Introduction

Fluorescence imaging microscopy has provided an effective means for detecting reporter gene expression (Naylor, 1999), transfection efficiency (Chiu et al, 1996), and cellular luminescence (Rizzuto et al, 1998). To extract meaningful information from those microscopic images, automatic image processing algorithms are crucial for processing large number of images. In the imaging processing, segmentation is the first essential and important step, and its accuracy determines the eventual success of image analysis (Pal and Pal, 1993). Image segmentation algorithms are to distinguish regions of interests from background, generally based on either discontinuity or similarity of intensity values (Gonzalez and Woods, 2002).

Threshold techniques, which determine an intensity value based on pixel intensity distribution histogram, are effective when objects’ intensity levels are contracting to those of background. They are especially popular in applications where speed is an important factor (Gonzalez and Woods, 2002). As a powerful segmentation algorithm, the adaptive thresholding was proposed differently from the common global thresholding.
It is effective for uneven illumination by dividing an image into subimages and then utilizes a different threshold to segment each subimage (Gonzalez and Woods, 2002).

Efforts have been dedicated for image segmentation of GFP-containing tissue images. Ridler’s adaptive thresholding was applied (Chi, 2004) to partition multiple transient GFP-expression spots in the lima bean tissues. The ImageJ software was generally employed for most GFP-expression images as an interactive global thresholding method (Rasband 1997-2006). In fact, stable GFP-containing embryogenic tissue images have some characteristics that should be used to select segmentation algorithms. Generally, most of the tissue images have inhomogeneities inside the tissue itself with non-homogeneous illumination. In addition, due to unavoidable stray light, there is quite a lot of noise in both tissue area and background area. Despite those downsides for tissue image processing, these images usually have one large object located in the center and the background is relatively simple except for gradient illumination.

Considering those special characteristics of the tissue images, three segmentation algorithms, namely adaptive thresholding, K-means clustering (MacQueen, 1967), and a graph-based normalized cut algorithm (Shi and Malik, 2000) were selected to accomplish the objectives in this chapter:

1. To program for automatic image reading, processing and data saving;
2. To quantitatively measure the segmentation quality of three automatic segmentation algorithms based on their accuracy, noise tolerance and running time;
3. To find a segmentation algorithm for separating tissue area from the background with high accuracy and high tolerance to noises.

4.2 Methodology

4.2.1 Three segmentation algorithms description

Adaptive thresholding

The adaptive thresholding is used in uneven lighting conditions when a lighter foreground object need to be segmented from its background (Nir Milstein, 1998). Adaptive thresholding is based on each pixel with respect to its local neighbourhood (Gonzalez and Woods, 2002). This allows each pixel to be considered in a more adaptive environment. The algorithm calculates the mean, median, minimum, maximum, minmax \(((\text{min}+\text{max})/2)\), or other statistics of the local neighbourhood. Thus, other than a source image, statistic and ‘window size’ of the neighbourhood are two necessary parameters. The pixel window should be large enough to reduce unwanted lighting.

K-means clustering

As stated in Chapter 3, there are only a few classes (two or three) in GFP fluorescence images of embryogenic soybean tissues: medium background, non-expressing tissue area, and GFP-expressing area. The K-means algorithm is an iterative process and good at partitioning an image into several clusters (Kanungo, 2002). Its procedures can be described as:

1. Decide a fixed number of clusters, \(k\);

2. Initial \(k\) center means randomly;
3. Assign each pixel to the center mean with a minimized variance;
4. Re-estimate the centers by averaging every pixel in the cluster;
5. Repeats steps 3 and 4 until it converges (no pixels change clusters).

The algorithm is guaranteed to converge, but an optimal solution may not be attained, which depends on the initial sets of centres and the number of clusters.

**N-cut algorithm**

The method of normalized cut algorithm can be used to segment out regions based on pixel intensity and locations (Shi and Malik, 2000). In this method, each image is regarded as a weighted undirected graph. Each pixel is a node in the graph, and an edge is formed between each two pixels. The weight of an edge is a value describing the similarity between the pixels. An image is partitioned into different groups by finding the minimum weights of the edges between these groups, which was called “cut”. The N-cut algorithm minimizes the “normalized cut”, which is the ratio of the cut to all of the edges in the set. The algorithm can be described in equation (4.1):

\[
\text{Weight } W = w(i, j) = e^{-\frac{\|F(i) - F(j)\|^2}{2\sigma^2}} \times \begin{cases} 
  e^{-\frac{\|X(i) - X(j)\|^2}{2\sigma^2}}, & \text{if } \|X(i) - X(j)\|_2 < r \\
  0, & \text{otherwise,}
\end{cases}
\]

(4.1)

where \(X(i)\) is the spatial location of pixel \(i\), and \(F(i)\) is a feature vector based on intensity at that pixel. \(w_{ij} = 0\) for any pair of pixels \(i\) and \(j\) if they are more than \(r\) pixels apart. For an image \(G=(V, E)\), the cut and normalized cut was given as:

\[
\text{cut}(A, B) = \sum_{u \in A, v \in B} w(u, v)
\]

(4.2)

\[
\text{Ncut}(A, B) = \frac{\text{cut}(A, B)}{\text{assoc}(A, V)} + \frac{\text{cut}(A, B)}{\text{assoc}(B, V)}
\]

(4.3)

\[
\text{Assoc}(A, V) = \sum_{u \in A, v \in V} w(u, v)
\]

(4.4)
Then the optimum partition is found by computing the minimum normalized cuts using eigenvectors.

4.2.2 Comparison of three algorithms

Three algorithms were compared in the same computer with an Intel Core 2 CPU 6600, 2.39GHz and 2.4GHz, as well as a 2G RAM and all programs were modified and executed in the environment of Matlab (MathWorks, 2005).

The local adaptive thresholding had two main parameters: local window size and the statistic (mean, median, minimum, maximum, or (min+max)/2). All of these statistics and window size (0-160 with a step of 10) were tried to segment the fluorescence images, finally a window size of 120 and local mean were chosen for adaptive thresholding because the most satisfying results were obtained by this choice. The parameter of the K-means clustering algorithm was the number of classes, which was tried with 2 and 3 in this study. For normalized cut algorithm, the segmented group number was selected 2 as an input parameter. To compare the performance of these three algorithms for segmenting low-contrast images with substantial noise, three parameters, i.e., accuracy, robustness, and running time, were compared.

To evaluate the accuracy, robustness and running time of those automatic segmentation algorithms, 24 images with dynamic ranges from 5 to 66 were selected from the images of three clones with low, low/moderate and moderate GFP expressions and non-transformed ‘Jack’ treated with 0, 3 and 10 mg/L IFT concentration levels, in which 12 were taken through the broad band-pass filter (GFP2 images), and 12 were obtained by the narrow band-pass filter (GFP3 images) (Fig. 4.1). This selection covered
most image contrast conditions among embryogenic soybean tissue fluorescence images.

In Fig. 4.1, those images with a lower dynamic range were either IFT-free tissue images or images taken by the narrow band-pass filter, as stated in Chapter 3.

![Dynamic range of images](image)

**Figure 4.1:** Dynamic ranges of 24 images which were selected to segment. Their dynamic ranges were from 5 to 66, selected from the images of three clones with low, low/moderate and moderate GFP expressions and non-transformed ‘Jack’ treated with 0, 3 and 10 mg/L IFT concentration levels, in which 12 were taken through the broad band-pass filter (GFP2 images), and 12 were obtained through the narrow band-pass filter (GFP3 images).

### 4.2.2.1 Accuracy

To evaluate the accuracy of the three algorithms, the tissue area as a feature was extracted by the three algorithms and the results were compared with interactive thresholding. Interactive thresholding segmented the images with original size 60...
(1200*1600 pixels), whereas for those three algorithms implemented in Matlab, the images were resized to 120*160 pixels for fast execution. Because of different size of the segmented images, tissue area ratio relative to the whole image was compared. A baseline was attained by interactively thresholding the 24 images with software ImageJ. First, each 1200*1600 RGB image was smoothed and split into blue, green and red gray images. Since the green channel image had the highest contrast, it was thresholded and the pixel number within tissue area was counted. A ratio was computed between pixel numbers in tissue area and that of the whole image.

The three algorithms were modified to implement automatic image reading, segmentation, computation and data saving. Since the image size was too large to compute in Matlab, the images were resized to 120*160 by bicubic interpolation for fast computations and the procedure of the tissue area ratio extraction is shown in Fig. 4.2.
4.2.2.2 Robustness

To compare three algorithms’ performance in separating the tissue from background, two types of misclassification error were defined in equations (4.5-4.6). Type I error was defined as the ratio of the number of pixels in background falsely segmented to tissue group to the total number of pixels within the background area, and
type II error was the ratio of the number of pixels in tissue area misclassified to background area to the total number of pixels within the tissue area.

\[
Type \ I \ error = \frac{\text{number of misclassified tissue pixels}}{\text{number of background pixels}} \tag{4.5}
\]

\[
Type \ II \ error = \frac{\text{number of misclassified background pixels}}{\text{number of tissue pixels}} \tag{4.6}
\]

The process of computing two types of error was developed as described in the following four steps. First, the fluorescence images were interactively smoothed because of large amount of noise in the background area. Each smoothed image was segmented with interactive thresholding, and then a binary image was saved as a standard. The pixel numbers in the tissue area and background area were counted respectively. Second, after segmentation of those images by each of the three methods, a mask (segmented result) was saved, in which the partitioned areas were marked by “1” and “2”. Take the normalized cut algorithm for example, in its final mask, background was labelled by “1” and tissue area was marked by “2”, therefore, the object and the background in the corresponding standard binary image got in the first step was labelled in the same way. Third, the standard image subtracted the segmented mask. As a result, most pixels would get zero values if they were correctly segmented. However, the pixels with a difference of “-1” were the ones in the background area, but were falsely classified as tissue pixels, and the difference of “1” represented those pixels within the tissue area, but were misclassified as background. Lastly, the numbers of these “1”s and “-1”s were counted respectively, and then the ratio of the number of “1”s to the number of background pixels got in the first step (type I error) and the ratio of the number of “-1”s to the number of tissue pixels (type II error) were computed as two types of misclassification error.
As a test of noise tolerance of each algorithm, using the function of `imnoise` in Matlab, black and white noises (the type of salt and pepper noise) were added to one typical fluorescence image with a dynamic range of 6.9 with various noise densities of 0, 0.1%, 0.5%, 1%, 2%, 3%, 5%, and 10%, which represent the percentage of the number of pixels in the image which was noise, and the size of these images were saved as 160×120 for fast executions in Matlab. Then the speckled images were segmented with normalized cut algorithm, K-means clustering algorithm (k was tried with 2 and 3), and adaptive thresholding. Two types of misclassification were computed to compare their tolerance to noises.

4.2.2.3 Running time

To compare the running speed of the three algorithms, running time was measured for the segmentation process in each algorithm using the stopwatch timer functions (`tic` and `toc`) in Matlab and automatically saved into Excel. Twenty images were executed by three algorithms and the average running time was compared.

4.3 Results and Discussion

4.3.1 Accuracy

The average ratio, based on 24 images, of the two dimensional projected tissue area relative to the whole image obtained by the three algorithms and interactive thresholding in ImageJ are shown in Fig. 4.3. The average area ratio obtained by the interactive thresholding was 14%. The result normalized cut was the most close to that of
the interactive thresholding. Adaptive thresholding took the second place, and the last one was K-means clustering. These automatic algorithms all had a larger standard deviation of tissue area ratio than that of the interactive thresholding, among which, that of K-means was the largest, approximately 14%.

For some images with large amount of noise, K-means and adaptive thresholding grouped lots of bright background pixels to tissue area. Meanwhile, K-means missed some pixels in the tissue area. Adaptive thresholding did not miss tissue pixels, however, it included large amount of background pixels and their neighbours due to their high intensity. The normalized cut algorithm could separate the tissue area, without including background pixels outside of the tissue contour, but on the tissue edge, some bright background pixels and dark tissue pixels were misclassified.
Figure 4.3: Average tissue area ratio relative to whole image of 24 images obtained with three automatic segmentation methods: normalized cut (Ncuts), K-means Clustering (Kmeans), and adaptive thresholding (Adaptive mean) and interactive thresholding (Manual).

4.3.2 Robustness

The results of two types of misclassification errors of four segmentation algorithms for 24 images were shown in Fig. 4.4. Normalized cut got the smallest misclassification error, under 6%. For normalized cut and K-means clustering, type II error was larger than type I error, and type I error was less than 1%, and type II error was lower than 10%. Adaptive thresholding had larger type I error than type II error as it included large amount of bright background pixels and their neighbourhoods as tissue pixels, although it did not miss much tissue pixels.
Figure 4.4: Two types of misclassification error of three segmentation algorithms K-means clustering (Kmeans), adaptive thresholding (Adaptive mean) and normalized cut (Ncuts) for segmenting 24 fluorescence images.

The algorithms’ performance on GFP2 and GFP3 images were also compared. Because of the lower contrast between the tissue area and background area as stated in Chapter 3, GFP3 images were more difficult to be segmented than GFP2 images. Therefore, the misclassification error of GFP3 images was larger than that of GFP2 images for each algorithm (Fig. 4.5). Normalized cut performed the best, with type I error under 2% and type II error under 6%. Moreover, this algorithm was most insensitive to noises, thus there was not a big difference between its misclassification error for GFP2 images and GFP3 images. However, K-means clustering and adaptive thresholding both had worse performance for GFP3 images than GFP2 images.
Figure 4.5: Two types of misclassification error of three algorithms for 12 images acquired with the GFP2 filter and 12 images with the GFP3 filter.

For a further test, a fluorescence image was treated with salt and pepper noise (Table 4.1) and the speckled images were segmented with the three algorithms. Normalized cut had the best performance, which was the most insensitive to noise, but it was influenced when the noise density increased to 10%. Adaptive thresholding was good at extracting the complete tissue area, however, it was sensitive to noise, and the
bright spots in background were separated as tissue pixels. K-means missed many pixels in the tissue area. When k was 3 to segment the image into 3 classes, some missed pixels by the case of k=2 were mitigated, but it was still severely influenced by the noises. This result was also proved by the two types of misclassification error of these three algorithms in Fig. 4.6. Normalized cut had the lowest error of background pixels misclassified as tissue (type I error) and it was very stable as the noise density increased, which is a very important attribute since there was a low signal to noise ratio in fluorescence images. It was found that adaptive thresholding had the highest type I error, and its error ratios exceeded the artificial noise ratios at each data point. K-means had almost the same error, and was very sensitive to noise, especially when noise density increased to 5% and 10%. Normalized cut falsely classified some tissue pixels to background area at the tissue boundary, but its type II error was stable with increasing noise level too. K-means had very high type II error, which was more than 10%, even when the noise density was less than 5%. For different parameters of K-means, when 2 classes were classified, the type II error was largely improved, but still got a high type I error when noise increased to 5%.
Table 4.1: Segmentation results of three algorithms: normalized cut (Ncuts), adaptive thresholding (Adaptive mean), and K-means clustering. A low contrast image was added salt and pepper noise with various noise density of 0, 0.001, 0.005, 0.01, 0.02, 0.03, 0.05 and 0.1 respectively (The 8 gray images in the first column), then the images with various levels of noise were segmented with the three algorithms. Contours indicate the segment boundaries of Ncuts algorithm (the second column).

*d is noise density.
Figure 4.6: Two types of misclassification error of three segmentation methods: K-means clustering (k=2 and k=3), normalized cut (Ncuts), adaptive thresholding (mean) for an image with a dynamic range of 6.9, which was treated with 8 noise density levels. These images were 8 categories (a through h) of noise levels: noise density d=0, 0.001, 0.005, 0.01, 0.02, 0.03, 0.05, and 0.1, respectively.
The reason for good performance of normalized cut was because that this algorithm considered both brightness and location features of each pixel. As a result, if there were some pixels on the tissue boundary with low intensity, they were easily misclassified as background pixels.

4.3.3 Running time

Running time of three algorithms were compared in Table 4.2. The K-means clustering and adaptive thresholding algorithms were both faster than the N-cut algorithm. The N-cut algorithm required a long running time, more than 150 times that of other algorithms.

<table>
<thead>
<tr>
<th>Method</th>
<th>K-means clustering</th>
<th>Adaptive thresholding</th>
<th>Normalized cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running time (sec)</td>
<td>0.040672</td>
<td>0.096065</td>
<td>14.85808</td>
</tr>
</tbody>
</table>

Table 4.2: Average running time per image for segmenting 20 images with the three methods.

4.4 Conclusions

Automatic image reading, processing and data saving were implemented with Matlab programming. Normalized cut algorithm was proved to be a suitable method for segmenting GFP fluorescence images, and it was the most insensitive to noises. It classified pixels based on brightness and coordinates features of pixels, and rarely misclassified tissue pixels to background area. Thus it did not miss pixels because of their
low brightness and the coordinate information could also give dominant information. For
low-contrast embryogenic soybean tissue images with large amount of noise, normalized
cut algorithm was shown to be a proper segmentation method, if its slow running speed
was not a critical problem.
CHAPTER 5

CONCLUSIONS

This thesis was a comprehensive study of using three approaches, namely herbicide IFT bleaching, spectral feature selection, and optimal image segmentation, to improve the GFP quantification in transgenic embryogenic soybean tissues.

The herbicide IFT approach, which was first developed in this thesis, was shown to be an effective way to eliminate chlorophyll emission and other possible autofluorescence from the pigments in plant tissue cultures, allowing easier visualization of GFP expression in moderate or low-expressing plant tissues. In addition, this study found that the IFT did not have negative effect on the tissue growth and the bleaching effect was completely reversible. IFT not only reduced chlorophyll and other autofluorescence interference, but also enhanced GFP expression intensity, which made two contrasts observably high: the one between the tissue area and the medium background and the one between the GFP expressing and non-GFP expressing areas. Therefore, the bleached tissue images could be easily segmented.

When the herbicide IFT is not applied in the plant material preparation, spectral feature selection could provide a method to reduce the interference from autofluorescence including chlorophyll from the emission detection perspective during the process of image acquisition. In this thesis, chlorophyll and other possible autofluorescence were
considered, which were detected in the soybean tissues with the spectroradiometer. According to the spectral characteristics of GFP and autofluorescence spectra as well as the fluorescence imaging system, autofluorescence can be blocked if the spectral band is within 500-590 nm. For accurate GFP detection, the spectral band should be narrow enough to block other autofluorescence. At the same time, it cannot be too narrow because the narrow band filter greatly attenuates the GFP signal intensity. The narrow band-pass GFP3 filter was evaluated and it blocked most chlorophyll fluorescence and improved the contrast between GFP expressing and non-expressing area, providing an access to accurate GFP quantification. Meanwhile, the narrow band-pass filter led to a darker image with low contrast between tissue and its medium background, which attenuated the fine detail for human visualization and increased the difficulty for image segmentation.

In order to find a suitable segmentation method for low contrast fluorescence images, three powerful image segmentation algorithms were compared and evaluated. It turned out that the state-of-the-art normalized cut (Ncuts) algorithm was the best algorithm for segmenting GFP fluorescence images, and furthermore, it was least sensitive to noise. Its good performance was attributed to the fact that it classified pixels based on brightness and coordinates features of pixels, and rarely misclassified tissue pixels to background area. For low-contrast embryogenic soybean tissue images with large amount of noise, normalize cut algorithm appeared to be the best choice among three all effective algorithms. One disadvantage of the Ncut algorithm is its relative longer processing time (roughly 14 seconds/image) which was due to the large
computation task. However, this would not be a critical problem for most static soybean tissue fluorescence image processing when the accuracy outweighs the speed. Automatic image reading, processing and data saving were also implemented by using Matlab programming in this study.

Although this thesis has made the valuable exploration of using three different approaches to improve the GFP quantification in tissue culture, there are several issues need to be considered if future work will be conducted: 1) To acquire monochrome images with narrow band-pass filters to avoid the signal attenuation by the color filters in the CCD camera. 2) To identify other herbicides that reduce chlorophyll content for GFP expression improvement. 3) To identify other segmentation algorithms with both improved accuracy and processing speed.
BIBLIOGRAPHY


APPENDIX A

DATA FOR SPECTRAL FEATURE SELECTION
<table>
<thead>
<tr>
<th>Wavelength</th>
<th>Chlorophyll Emission</th>
<th>GFP Emission</th>
<th>Collection Efficiency of Objective</th>
<th>Transmission Efficiency of Objective</th>
<th>Quantum Efficiency of CCD with blue filter</th>
<th>Quantum Efficiency of CCD with green filter</th>
<th>Quantum Efficiency of CCD with red filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>350</td>
<td>0.071023</td>
<td>4.55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>355</td>
<td>0.154959</td>
<td>9.92</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>360</td>
<td>0.290548</td>
<td>18.60</td>
<td>0.002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>365</td>
<td>0.41968</td>
<td>26.86</td>
<td>0.003</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>0.555269</td>
<td>35.54</td>
<td>0.004</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>375</td>
<td>0.697314</td>
<td>44.63</td>
<td>0.0044</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>380</td>
<td>0.826446</td>
<td>52.89</td>
<td>0.0048</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>385</td>
<td>0.923295</td>
<td>59.00</td>
<td>0.0054</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>390</td>
<td>1.020145</td>
<td>65.29</td>
<td>0.006</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>395</td>
<td>1.110537</td>
<td>71.07</td>
<td>0.0074</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>1.181302</td>
<td>75.60</td>
<td>0.0088</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>405</td>
<td>1.226756</td>
<td>78.51</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>410</td>
<td>1.276409</td>
<td>81.82</td>
<td>0.028</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>415</td>
<td>1.323605</td>
<td>84.71</td>
<td>0.044</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>420</td>
<td>1.355888</td>
<td>86.78</td>
<td>0.064</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>425</td>
<td>1.375258</td>
<td>88.02</td>
<td>0.074</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>430</td>
<td>1.401085</td>
<td>89.67</td>
<td>0.084</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>435</td>
<td>1.426911</td>
<td>91.32</td>
<td>0.092</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>440</td>
<td>1.433368</td>
<td>91.74</td>
<td>0.102</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>445</td>
<td>1.456612</td>
<td>93.22</td>
<td>0.102</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>450</td>
<td>1.470816</td>
<td>94.13</td>
<td>0.1008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>455</td>
<td>1.485021</td>
<td>95.04</td>
<td>0.1044</td>
<td>0.005</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>460</td>
<td>1.485021</td>
<td>95.04</td>
<td>0.1004</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>465</td>
<td>1.50439</td>
<td>96.28</td>
<td>0.0916</td>
<td>0.024</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>470</td>
<td>1.499871</td>
<td>95.99</td>
<td>0.078</td>
<td>0.044</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>475</td>
<td>0.029926</td>
<td>1.51986</td>
<td>0.068</td>
<td>0.072</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>480</td>
<td>0.069202</td>
<td>1.52177</td>
<td>0.062</td>
<td>0.078</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>485</td>
<td>0.147824</td>
<td>1.52234</td>
<td>0.05</td>
<td>0.096</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>490</td>
<td>0.283023</td>
<td>1.52376</td>
<td>0.038</td>
<td>0.104</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>495</td>
<td>0.486674</td>
<td>1.53409</td>
<td>0.0236</td>
<td>0.118</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.756043</td>
<td>1.535382</td>
<td>0.014</td>
<td>0.124</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>505</td>
<td>0.963555</td>
<td>1.540548</td>
<td>0.011</td>
<td>0.127</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>510</td>
<td>1</td>
<td>1.541839</td>
<td>0.006</td>
<td>0.126</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>515</td>
<td>0.865315</td>
<td>1.541968</td>
<td>0.005</td>
<td>0.128</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>520</td>
<td>0.67646</td>
<td>1.542097</td>
<td>0.004</td>
<td>0.130</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>525</td>
<td>0.531796</td>
<td>1.545713</td>
<td>0.003</td>
<td>0.134</td>
<td>0.002</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Continued
| 530 | 0.443535 | 1.547004 | 99.01 | 0.002 | 0.132 | 0.004 |
| 535 | 0.386451 | 1.547004 | 99.01 | 0 | 0.126 | 0.006 |
| 540 | 0.3448 | 1.547133 | 99.02 | 0 | 0.12 | 0.008 |
| 545 | 0.305051 | 1.552169 | 99.34 | 0 | 0.11 | 0.0128 |
| 550 | 0.258522 | 1.555398 | 99.55 | 0 | 0.1 | 0.02 |
| 555 | 0.208894 | 1.555139 | 99.53 | 0 | 0.088 | 0.028 |
| 560 | 0.164355 | 1.555139 | 99.53 | 0 | 0.076 | 0.036 |
| 565 | 0.129466 | 1.554881 | 99.51 | 0 | 0.064 | 0.048 |
| 570 | 0.10282 | 1.554752 | 99.5 | 0 | 0.052 | 0.054 |
| 575 | 0.08249 | 1.553401 | 99.42 | 0 | 0.0388 | 0.0568 |
| 580 | 0.069189 | 1.552169 | 99.34 | 0 | 0.032 | 0.058 |
| 585 | 0.057741 | 1.550878 | 99.26 | 0 | 0.024 | 0.062 |
| 590 | 0.04824 | 1.549587 | 99.17 | 0 | 0.018 | 0.068 |
| 595 | 0.040052 | 1.54313 | 98.76 | 0 | 0.012 | 0.072 |
| 600 | 0.025896 | 0.032873 | 1.547004 | 99.01 | 0 | 0.008 | 0.0692 |
| 605 | 0.020648 | 0.025677 | 1.545713 | 98.93 | 0 | 0.006 | 0.0684 |
| 610 | 0.017997 | 0.021501 | 1.545713 | 98.93 | 0 | 0.004 | 0.0668 |
| 615 | 0.015505 | 0.016864 | 1.545713 | 98.93 | 0 | 0 | 0.0604 |
| 620 | 0.014786 | 0.013457 | 1.548295 | 99.09 | 0 | 0 | 0.0508 |
| 625 | 0.01348 | 0.011015 | 1.548295 | 99.09 | 0 | 0 | 0.0516 |
| 630 | 0.015066 | 0.009333 | 1.548295 | 99.09 | 0 | 0 | 0.046 |
| 635 | 0.014752 | 0.00986 | 1.548037 | 99.07 | 0 | 0 | 0.04 |
| 640 | 0.018674 | 0.006707 | 1.547004 | 99.01 | 0 | 0 | 0.036 |
| 645 | 0.025067 | 0.005643 | 1.537965 | 98.43 | 0 | 0 | 0.0316 |
| 650 | 0.040294 | 0.004958 | 1.536674 | 98.35 | 0 | 0 | 0.028 |
| 655 | 0.062234 | 1.530217 | 97.93 | 0 | 0 | 0.0208 |
| 660 | 0.104591 | 1.525052 | 97.6 | 0 | 0 | 0.0172 |
| 665 | 0.180596 | 1.524406 | 97.56 | 0 | 0 | 0.0128 |
| 670 | 0.338147 | 1.52376 | 97.52 | 0 | 0 | 0.01 |
| 675 | 0.589968 | 1.519886 | 97.27 | 0 | 0 | 0.0076 |
| 680 | 0.89517 | 1.514721 | 96.94 | 0 | 0 | 0.006 |
| 685 | 1.964664 | 1.506973 | 96.45 | 0 | 0 | 0.0044 |
| 690 | 0.730549 | 1.491477 | 95.45 | 0 | 0 | 0.0028 |
| 695 | 0.512057 | 1.486312 | 95.12 | 0 | 0 | 0.0014 |
| 700 | 0.382609 | 1.478564 | 94.63 | 0 | 0 | 0 |
| 705 | 0.310632 | 1.466942 | 93.88 | 0 | 0 | 0 |
| 710 | 0.275299 | 1.460486 | 93.47 | 0 | 0 | 0 |
Continued

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>715</td>
<td>0.262308</td>
<td>1.452738</td>
<td>92.98</td>
</tr>
<tr>
<td>720</td>
<td>0.259838</td>
<td>1.446281</td>
<td>92.56</td>
</tr>
<tr>
<td>725</td>
<td>0.261263</td>
<td>1.433368</td>
<td>91.74</td>
</tr>
<tr>
<td>730</td>
<td>0.264052</td>
<td>1.420455</td>
<td>90.91</td>
</tr>
<tr>
<td>735</td>
<td>0.269417</td>
<td>1.404959</td>
<td>89.92</td>
</tr>
<tr>
<td>740</td>
<td>0.268551</td>
<td>1.394628</td>
<td>89.26</td>
</tr>
<tr>
<td>745</td>
<td>0.251269</td>
<td>1.384298</td>
<td>88.60</td>
</tr>
<tr>
<td>750</td>
<td>0.221641</td>
<td>1.372676</td>
<td>87.85</td>
</tr>
<tr>
<td>755</td>
<td>0.184712</td>
<td>1.362345</td>
<td>87.19</td>
</tr>
<tr>
<td>760</td>
<td>0.15581</td>
<td>1.349432</td>
<td>86.36</td>
</tr>
<tr>
<td>765</td>
<td>0.124057</td>
<td>1.330062</td>
<td>85.12</td>
</tr>
<tr>
<td>770</td>
<td>0.09697</td>
<td>1.319731</td>
<td>84.46</td>
</tr>
<tr>
<td>775</td>
<td>0.077293</td>
<td>1.304236</td>
<td>83.47</td>
</tr>
<tr>
<td>780</td>
<td>0.060864</td>
<td>1.28874</td>
<td>82.48</td>
</tr>
<tr>
<td>785</td>
<td>0.052462</td>
<td>1.278409</td>
<td>81.82</td>
</tr>
<tr>
<td>790</td>
<td>0.044864</td>
<td>1.268079</td>
<td>81.16</td>
</tr>
<tr>
<td>795</td>
<td>0.039075</td>
<td>1.25</td>
<td>80.00</td>
</tr>
<tr>
<td>800</td>
<td>0.036991</td>
<td>1.233213</td>
<td>78.93</td>
</tr>
</tbody>
</table>
APPENDIX B

MATLAB SOURCE CODE FOR MISCLASSIFICATION ERROR COMPUTATION FOR THREE ALGORITHMS
function GFPNcutMisclassification;

% automatically compute two types of misclassification error
% automatically read images from folder

nr=120;nc=160;
M=70;N=77;nm=80;
for imagenb=M:N
    imageaddress=['specific_NcutImage_files/jpg_images/gfp/'
                  num2str(imagenb),'.tif']
    I = imread_ncut(imageaddress,nr,nc); % resized brightness image
    nbSegments =2; % the image will be separated into 2 parts
    disp('computing Ncut eigenvectors ...');
    tic;
    [SegLabel,NcutDiscrete,NcutEigenvectors,NcutEigenvalues,W,imageEdges]=
     NcutImage(I,nbSegments);
    disp(['The computation took ' num2str(toc) ' seconds on the '
          num2str(size(I,1)) ' x ' num2str(size(I,2)) ' image']);
    t(imagenb)=toc;

    %compute type I error
    StdMask=imread(['c:\FABEGrad Desktop\USER FILES HERE ONLY\congling
                    wu\Segmentation\NcutImage_7_AMD64\NcutImage_7_AMD64\specific_NcutImage_files\jpg_images\gfp\' num2str(nm),'.tif']);
    % background area
    bpixelnb=length(find(StdMask(:,:)==255));
    tpixelnb=length(find(StdMask(:,:)==0));
    if SegLabel(1,1)==1
        StdMask(find(StdMask(:,:)==255))=1;
        StdMask(find(StdMask(:,:)==0))=2;
        Dff=double(StdMask)-SegLabel;
        er1(imagenb)=length(find(Dff(:,:)==-1));
        er2(imagenb)=length(find(Dff(:,:)==1));
    elseif SegLabel(1,1)==2
        StdMask(find(StdMask(:,:)==255))=2;
        StdMask(find(StdMask(:,:)==0))=1;
        Dff=double(StdMask)-SegLabel;
        er1(imagenb)=length(find(Dff(:,:)==1));
        er2(imagenb)=length(find(Dff(:,:)==-1));
    end
    er1(imagenb)=er1(imagenb)/bpixelnb;
    er2(imagenb)=er2(imagenb)/tpixelnb;
    figure(imagenb);
    imagesc(SegLabel);colormap(gray);axis off;
end
xlswrite('NcutsError.xls',[er1',er2'],'Misclassification-no noise');

function
[SegLabel,NcutDiscrete,NcutEigenvectors,NcutEigenvalues,W,imageEdges]=
 NcutImage(I,nbSegments);

89
if nargin < 2,
    nbSegments = 10;
end

[W, imageEdges] = ICgraph(I);

[NcutDiscrete, NcutEigenvectors, NcutEigenvalues] = ncutW(W, nbSegments);

%% generate segmentation label map
[nr, nc, nb] = size(I);
SegLabel = zeros(nr, nc);
for j = 1:size(NcutDiscrete, 2),
    SegLabel = SegLabel + j * reshape(NcutDiscrete(:, j), nr, nc);
end

function [Eigenvectors, Eigenvalues] = ncut(W, nbEigenValues, dataNcut);
if nargin < 2
    nbEigenValues = 8;
end
if nargin < 3
    dataNcut.offset = 5e-1;
    dataNcut.verbose = 0;
    dataNcut.maxiterations = 100;
    dataNcut.eigsErrorTolerance = 1e-6;
    dataNcut.valeurMin = 1e-6;
end

% make W matrix sparse
W = sparsifyc(W, dataNcut.valeurMin);

% check for matrix symmetry
if max(max(abs(W - W'))) > 1e-10 % voir (-12)
    disp(max(max(abs(W - W'))));
    error('W not symmetric');
end

n = size(W, 1);
nbEigenValues = min(nbEigenValues, n);
offset = dataNcut.offset;

% degrees and regularization
d = sum(abs(W), 2);
dr = 0.5 * (d - sum(W, 2));
d = d + offset * 2;
offset = dr + offset;
W = W + spdiags(dr, 0, n, n);
Dinvsqrt = 1./sqrt(d+eps);
P = spmtimesd(W,Dinvsqrt,Dinvsqrt);
clear W;

options.issym = 1;

if dataNcut.verbose
    options.disp = 3;
else
    options.disp = 0;
end
options.maxit = dataNcut.maxiterations;
options.tol = dataNcut.eigsErrorTolerance;

options.v0 = ones(size(P,1),1);
options.p = max(35,2*nbEigenValues);
options.p = min(options.p,n);

[vbar,s,convergence] =
eigs2(@mex_w_times_x_symmetric,size(P,1),nbEigenValues,'LA',options,tri l(P));

s = real(diag(s));
[x,y] = sort(-s);
Eigenvalues = -x;
vbar = vbar(:,y);
Eigenvectors = spdiags(Dinvsqrt,0,n,n) * vbar;

for i=1:size(Eigenvectors,2)
    Eigenvectors(:,i) = (Eigenvectors(:,i) / norm(Eigenvectors(:,i)) ) * norm(ones(n,1));
    if Eigenvectors(1,i)~=0
        Eigenvectors(:,i) = - Eigenvectors(:,i) * sign(Eigenvectors(1,i));
    end
end

function [NcutDiscrete,NcutEigenvectors,NcutEigenvalues] =
ncutW(W,nbcluster);

[NcutEigenvectors,NcutEigenvalues] = ncut(W,nbcluster);

% compute discretize Ncut vectors
[NcutDiscrete,NcutEigenvalues] =discretisation(NcutEigenvalues);

NcutDiscrete = full(NcutDiscrete);
function

[EigenvectorsDiscrete,Eigenvectors] = discretisation(Eigenvectors)

[n,k] = size(EigenVectors);

vm = sqrt(sum(EigenVectors.*EigenVectors,2));
EigenVectors = EigenVectors./repmat(vm,1,k);

R = zeros(k);
R(:,1) = EigenVectors(1+round(rand(1)*(n-1)),:);
c = zeros(n,1);
for j = 2:k
    c = c + abs(EigenVectors*R(:,j-1));
    [minimum,i] = min(c);
    R(:,j) = EigenVectors(i,:);
end

lastObjectiveValue = 0;
exitLoop = 0;
nbIterationsDiscretisation = 0;
nbIterationsDiscretisationMax = 20; % voir
while exitLoop == 0
    nbIterationsDiscretisation = nbIterationsDiscretisation + 1;
    EigenvectorsDiscrete = discretisationEigenvectorData(EigenVectors*R);
    [U,S,V] = svd(EigenvectorsDiscrete'*EigenVectors,0);
    NcutValue = 2*(n - trace(S));
    if abs(NcutValue - lastObjectiveValue) < eps | nbIterationsDiscretisation > nbIterationsDiscretisationMax
        exitLoop = 1;
    else
        lastObjectiveValue = NcutValue;
        R = V*U';
    end
end

function k-means;

% automatically read images from folder
nr = 120; nc = 160;
M = 70; N = 77; nm = 80;
for imagenb = M:N
    imageaddress = ['c:\FABEGrad Desktop\USER FILES HERE ONLY\congling wu\Segmentation\NcutImage_7_AMD64\NcutImage_7_AMD64\specific_NcutImage_files\jpg_images\gfp' num2str(imagenb),'.tif']
    I1 = imread(imageaddress); % original RGB image
    I1 = imresize(I1, [nr, nc], 'bicubic'); % resized RGB image
    ima = rgb2gray(I1);
    % I = imread_ncut(imageaddress, nr, nc); % resized brightness image
    ima = medfilt2(ima, [3 3]);
    ima = double(ima);
copy=ima; % make a copy
ima=ima(:); % vectorize ima
mi=min(ima); % deal with negative
ima=ima-mi+1; % and zero values

s=length(ima);

% create image histogram

m=max(ima)+1;
h=zeros(1,m);
hc=zeros(1,m);

for i=1:s
    if(ima(i)>0) h(ima(i))=h(ima(i))+1;end;
end
ind=find(h);
hl=length(ind);

% initiate centroids
mu=(1:k)*m/(k+1);

% start process
tic;
while(true)
it=0;
    oldmu=mu;
    % current classification
    for i=1:hl
        c=abs(ind(i)-mu);
        cc=find(c==min(c));
        hc(ind(i))=cc(1);
    end

    % recalculation of means
    for i=1:k,
        a=find(hc==i);
        mu(i)=sum(a.*h(a))/sum(h(a));
    end

    if(mu==oldmu) break;end;
it=it+1;
end

display(['this image took' num2str(it) 'iterations']);

% calculate mask
s=size(copy);
mask=zeros(s);
for i=1:s(1),

93
for j=1:s(2),
c=abs(copy(i,j)-mu);
a=find(c==min(c));
mask(i,j)=a(1);
end
end

mu=mu+mi-1;   % recover real range
t(imgenb)=toc;
bw=edge(mask,0.02);
imagesc(bw);axis off;
figure;imagesc(mask);axis off;title('After Segmentation');
colormap(gray(64));
imadr=['c:\FABEGrad Desktop\USER FILES HERE ONLY\congling
wu\Segmentation\NcutImage_7_AMD64\NcutImage_7_AMD64\specific_NcutImage_
files\jpg_images\gfp\' num2str(nm),'.tif']
StdMask=imread(imadr);

% k=2 misclassification ratio
bpixelnb=length(find(StdMask(:,:,)==255));
tpixelnb=length(find(StdMask(:,:,)==0));
StdMask(find(StdMask(:,:,)==255))=1;
StdMask(find(StdMask(:,:,)==0))=2;
Dff=double(StdMask)-mask;
er1(imgenb)=length(find(Dff(:,:,)==-1));
er2(imgenb)=length(find(Dff(:,:,)==1));
er1(imgenb)=er1(imgenb)/bpixelnb;
er2(imgenb)=er2(imgenb)/tpixelnb;
end
xlswrite('KmeanserrorKis2.xls',[er1',er2'],'Misclassification');

function adaptiveMisclassification;
% automatically read images from folder
nr=120;nc=160;
M=70;N=77;nm=80;
for imagenb=M:N
imageaddress=['c:\FABEGrad Desktop\USER FILES HERE ONLY\congling
wu\Segmentation\NcutImage_7_AMD64\NcutImage_7_AMD64\specific_NcutImage_
files\jpg_images\gfp\' num2str(imagenb),'.tif']
I1=imread(imageaddress);%original RGB image
I1=imresize(I1,[nr,nc],'bicubic'); % resized RGB image
I1=rgb2gray(I1);
tic;
bwim1=adaptivethreshold(I1,120,0.002,0);
t(imgenb)=toc;
figure(imgenb);imagesc(bwim1);colormap(gray);axis off;
imadr=['c:\FABEGrad Desktop\USER FILES HERE ONLY\congling
wu\Segmentation\NcutImage_7_AMD64\NcutImage_7_AMD64\specific_NcutImage_
files\jpg_images\gfp\' num2str(nm),'.tif']
StdMask=imread(imadr);
% missclassification ratios
bpixelnb=length(find(StdMask(:,)==255));
tpixelnb=length(find(StdMask(:,)==0));
StdMask(find(StdMask(:,)==0))=1;
StdMask(find(StdMask(:,)==255))=0;%background
Dff=double(StdMask)-bwim1;
er1(imagenb)=length(find(Dff(:,)==-1));
er2(imagenb)=length(find(Dff(:,)==1));
er1(imagenb)=er1(imagenb)/bpixelnb;
er2(imagenb)=er2(imagenb)/tpixelnb;
end
xlswrite('AdaptivemeanError.xls',[er1',er2'],"Misclassification-
nonoise");
% figure;imagesc(bwim2);colormap(gray);axis off;

function bw=adaptivethreshold(IM,ws,C,tm)
if (nargin<3)
    error('You must provide the image IM, the window size ws, and C.');
elseif (nargin==3)
    tm=0;
elseif (tm~=0 && tm~=1)
    error('tm must be 0 or 1.);
end
IM=mat2gray(IM);
if tm==0
    mIM=imfilter(IM,fspecial('average',ws),'replicate');
else
    mIM=medfilt2(IM,[ws ws]);
end
sIM=mIM-IM-C;
bw=im2bw(sIM,0);
bw=imcomplement(bw);