Outdoor computer vision and weed control

Andersen, Hans Jørgen

Publication date:
2001

Document Version
Early version, also known as pre-print

Computer Vision & Media Technology Laboratory

Ph.D. Dissertation

Outdoor Computer Vision and Weed Control

Hans Jørgen Andersen

Computer Vision
& Media Technology Laboratory

Aalborg University 2002

Outdoor Computer Vision and Weed Control

A Ph.D. dissertation
by

Hans Jørgen Andersen

Computer Vision & Media Technology Laboratory
Institute of Health Science and Technology
Aalborg University, Denmark
E-mail: hja@cvmt.dk
URL: http://www.cvmt.dk/~hja

June 2002

This dissertation is submitted in January, 2001 to the Faculty Council of Technical Sciences, Aalborg University, Denmark, in partial fulfillment of the requirements for the Doctor of Philosophy degree.

While the first edition was approved, this second edition includes revisions in accordance with comments from the adjudication committee.

The following adjudication committee was appointed to evaluate the thesis:

Professsor Dr. ir A. W. M. Smeulders (head opponent)
University of Amsterdam
Informatics Institute, Faculty of Sciences
Amsterdam, Netherlands

Professor Josse de Baerdemaeker
Katholieke Universiteit Leuven
Faculty Agricultural and Applied Biological Sciences
Department Agro-Engineering and Economics
Laboratory for Agro Machinery and Processing
Leuven, Belgium

Associate Professor Claus Brøndgaard Madsen (committee chairman)
Computer Vision & Media Technology Laboratory
Aalborg University
Aalborg, Denmark

Abstract

This thesis investigates the potential support of computer vision within outdoor weed control for reduced use of herbicides. The outdoor use of computer vision encounters the problem that the illumination conditions may change from e.g. direct sunshine to a clear blue sky as illumination source. Such changes will radically alter the colours and conditions for the image formation. The focus of the dissertation is to find methods that will make it possible to implement computer vision for robust operation under such changeable outdoor/daylight conditions for the image formation.

The work is divided into two main parts. The first investigates the relation between computer vision and weed control while the second investigates and develops computer vision methods adaptable to varying daylight conditions.

The first part introduces a new concept, "accessibility" for assessment and description of the performance of weed control treatments in a given context. A concept that may combine both primary and secondary parameters influencing on the condition for the weed control treatment. Knowing their influence on the treatment is a necessary prerequisite for successful support of computer vision. The power of the concept is that it may both be used for description of the objectives for the weed control treatment and for assessment of the parameters influencing on it. Hence, it is able to serve as general basis for both planning and execution of the weed control treatment. The concept has also been used for structuring of a comprehensive survey of the literature within this field. The survey shows that many computer vision methods and techniques have been investigated within the application of weed control but only few have taking their point of departure in the process "weed control".

The second part approaches the problem of using computer vision outdoors in two different ways. First by the use of a "standard" RGB camera and second by development of a dedicated sensor for detecting vegetation. Initially the theory behind a comprehensive modelling of the image formation is introduced. The modelling is based on a physical understanding of the image formation involving use of CIE daylight standard, the radiance from a Black Body, camera and reflectance characteristics, and the dichromatic reflection model.

Based on this modelling a taxonomy for categorizing the current illumination characteristics of two light sources by the reflection from an object represented by its intensity independent chromaticities, is presented. The analysis shows that daylight may be found as an important category where the sun may be regarded as a uniform light source and the sky as an ambient light source. The taxonomy is used in two experiments, first for the assessment of the reflection of six uniform objects illuminated by two known light sources in controlled laboratory conditions. In the second for the assessment of the illumination of realistic images in field condition of Barley plants monitored from early morning to late evening in direct sunshine, over cast, and shaded light conditions.

The modelling of the image formation is also used for the development of a segmentation method invariant to spectral variation of daylight for segmentation of vegetation from background regarded as soil. By introducing the intensity as a parameter the method is capable of classifying background areas in shade by the vegetation. These areas will be biased towards the colour of the vegetation as the light is first transmitted through the plants. Instead they may be classified by the feature that they will have a lower intensity than the directly illuminated vegetation.

The other approach involves the design and the evaluation of a dedicated sensor for detecting vegetation. An investigation of the reflectance characteristics of vegetation and soil reported in the literature yields that bands in the green, red, and near infra red would be the best wave bands to use for the classification of vegetation and soil. For the design and evaluation of the sensor with these three wave bands the modelling of daylight is extended into the near infra-red region by the radiance characteristics of a Black Body. The sensor is implemented and evaluated by comparing the modelled characteristic with images captured at different conditions of the image formation. Finally a segmentation based on the physical modelling of the sensor is developed and compared with different combinations of the three wave bands. The analysis shows that the sensor may provide a more robust basis for classification of vegetation under

different conditions of the image formation than a "traditional" camera.

The results of the work demonstrate that a priori knowledge and modelling of the image formation can be of significant importance for the development of image analysis methods and sensors for robust analysis of colour information captured under varying illumination condition. This opens general perspectives for development of computer vision methods and sensors for outdoor use. Within weed control it is especially encouraging that the proposed methods are through improved use of colour information will be able to operate on images with vegetation on an early growth stage. A property of vital importance for future development of new plant growth systems with minimal use of herbicides.

Resume

Denne Ph.D. afhandling omhandler den potentielle brug af computer vision ved udendørs ukrudtskontrol for at reducere brugen af herbicider (ukrudtsmidler). Anvendelsen af computer vision under udendørs forhold byder på det problem at belysningsforholdene kan skifte fra f.eks. direkte solskin til skygge med det blå himmelrum som lyskilde, et skift der ændrer farverne og betingelserne for billedoptagelsen radikalt. Fokus for afhandlingen er at finde metoder, der kan gøre det muligt at implementere computer vision til robust funktion under disse omskiftelige ydre forhold.

Arbejdet kan indeles i to hovedemner. Det første behandler relationen mellem computer vision og ukrudtskontrol, mens det andet behandler udviklingen af computer vision metoder, der kan tilpasse sig varierende dagslysforhold.

Første hovedemne introducerer et nyt koncept "accessibility", til vurdering og beskrivelse af en ukrudtskontrolmetodes præstationsevne i en given sammenhæng. Konceptet gør det muligt at kombinere både primære og sekundære parametre, der influerer på en given ukrudtskontrolmetode. At kende disse parametres indflydelse er nødvendig for bl.a. en vellykket brug af computer vision. Styrken ved accessibility -konceptet ligger i at det både kan anvendes til at beskrive formålet for ukrudtskontrolmetoden og parametrene der influerer på denne. Herved kan "accessibility" anvendes som en generel basis for planlægning og styring af ukrudtskontrol og kulturtekniske tiltag. Konceptet er også brugt til at strukturere en omfattende gennemgang af litteraturen inden for området. En gennemgang der viser at mange computer vision metoder og teknikker er søgt overført til indenfor området, men at mere grundlagsgivende arbejder med udgangspunkt i processen "ukrudtskontrol" kun foreligger i ringe omfang.

Det andet hovedemne angribes fra to forskellige vinkler, først ved brugen af et standard RGB kamera og siden ved udvikling af en dedikeret sensor for detektion af vegetation. Indledningsvis gennemgås den bagved liggende teori for udviklingen af en omfattende modellering af billeddannelsesprocesen. Denne baseres på en beskrivelse af de væsentligste fysiske størrelser involverende brug af CIE's dagslysstandard, udstrålingen fra at sort legeme, kamera og reflektans karakteristikker samt den dikromatiske reflektion model.

Med udgangspunkt i denne modellering udvikles en taxanomi for katagorisering af belysningsbetingelser ved to belysningskilder, baseret på et objekts reflektioner repræsenteret ved dets intensitets uafhængige "kromaticiteter" (chromaticities). Af analysen kan dagslys etableres som en vigtigt kategori, hvor solen kan antages som en uniform lyskilde og himmelrummet som en omgivende lyskilde. Kategoriseringen anvendes i to eksperimenter, først for beskrivelse af reflektionsfordelingen af seks velkendte ensartede objekter belyst med to velkendte lyskilder under kontrollerede laboratorie forhold. Dernæst for vurdering af belysningbetingelserne af realitiske billeder af bygplanter under markforhold, der følges fra tidlig morgen til sen aften i direkte solskin, i overskyet vejr og i skygge.

Modelleringen af den billeddannende proces anvendes yderligere til at udvikle en segmenteringsmetode, der er invariant overfor spektral variation af dagslys, og anvendes til segmentering af vegetation fra baggrund bestående hovedsaglig af jord. Ved yderligere at inddrage intensiteten som parameter udmærker denne metode sig ved at være i stand til at klassificere baggrundsområder i skygge af vegetation. Disse områder vil have tendens til at fremstå grønlige pga. at lyset først er transmitteret gennem vegetationen. De kan udskilles, idet de vil have en lavere intensitet end vegetationen belyst direkte.

Den anden angrebsvinkel indenfor hovedtemaet, behandler udvikling og evaluering af en dedikeret sensor for detektering af vegetation. En underøgelse of reflektanskarakteristikker af vegetation og jord i litteraturen viser at bånd i det grønne, røde og nær-infra-røde område vil være mest velegnede for klassificering af vegetation fra baggrund (jord). For at evaluere sensoren udvides modellen af dagslyset ind i det nær-infra-røde område med udstrålingen fra et sort legeme. Sensoren vurderes ved først at sammenholde dens modellerede karakteristik med billeder optaget ved forskellige betingelser for billededannelsen. Sluttelig udvikles en segmenteringsmetode på basis af den fysiske modellering af sensoren, der sammenlignes med de andre mulige kombinationer af de tre bånd. Analysen viser at denne metode giver den mest robuste basis for klassificering af vegetation under varierende betingelser for billededannelsen.

Generalt viser resultaterne af det samlede arbejde, at a priori viden opnået ved modellering af billeddannelsesprocessen, kan muliggøre robust brug af farveinformation under varierende belysning. Dette åbner generelle perspektiver for udvikling af computer vision teknikker og sensorer til udendørs brug. Inden for ukrudtskontrol er det specielt lovende at de udviklede metoder er i stand til at arbejde med vegetation på et meget tidligt vækststadie, hvilket er af vital betydning for fremtidig udvikling af dyrkningssystemer med minimal brug af herbicider.

Preface

This dissertation is based upon the work and findings of a Ph.D.- scholarship Digital Image Processing (DIP 1), within the national research framework: Optimization and Control in Plant Production of Technology, Implements, Machinery and Systems for Agriculture (OPTIMAL), founded by the Danish Technical Research Council and the Danish Agricultural Veterinary Research Council. The project has been organized as a collaboration between Laboratory of Computer Vision and Media Technology (CVMT), Aalborg University and Department of Agricultural Engineering, Royal Veterinary and Agricultural University (KVL), Copenhagen.

A number of individuals have made this research possible which I am all profoundly grateful to. My supervisor Professor Erik Granum for encouraging and fruitful discussion and tutoring me in the art of research throughout. My colleagues at CVMT and my supervisor Bent S. Bennedsen at KVL for helpful collaboration and support with my experiments.

During my study I had the pleasure of being with the people at Silsoe Research Center during the summer 1998. I am deeply grateful to Professor John Marchant and Mrs. Christine Onyango for making this stay very inspiring and for the honour and pleasure of working together with you.

Finally I thank my dear wife and son for supporting me and bearing with me when working late at home and having my thought elsewhere. Without you this would not have been possible.

Hans Jørgen Andersen
Aalborg June 2002

Contents

Chapter 1

Introduction

1.1 Motivating Computer Vision in Weed Control

Increasing concern from the society over the potential environmental side effects of using pesticides has given agriculture worldwide a great challenge in finding ways to lower their use. Already the Brundtland commission [3] recommended that research in alternatives to pesticides should be encouraged and that increased use and development of technical support facilities could contribute to better resource management and more sustainable agricultural production systems. Especially the commission emphasised that the industrial countries should lower their use whereas the developing countries could increase their use of pesticides to achieve a more reliable food production and to a higher degree become self sufficient.

In this context Denmark in 1987 decided to aim towards a 50% reduction by 1997 of the total pesticide use with the period 1981-85 as reference [9]. In autumn 1997 the Danish Ministry of Environment and Energy established a committee to follow up upon the 1987 decision and to investigate the consequences of a total stop of pesticide use (the so called Bichel committee [2]). The committee concluded that the aim from 1987 was only partly fulfilled but more importantly it points out that it is of significant importance to increase research in methods that may contribute to a reduction in use of pesticides.

Herbicides contribute with approximately 50% of the total pesticide use, consequently a reduction in this area will be of significant importance. One way of lowering the use is to apply it selectively, i.e. only where the weed density exceeds the economic threshold [6], rather than to the entire field as common practice is today. Another way is to use various non-chemical weed control methods like hoeing, flaming etc., as described in [7].

To make the slective method applicable it will be necessary to have information about the weed pressure at the individual location to adjust the doze of herbicides. Ultimately, the herbicides may be applied only on the weed plants (spot spraying) which will require that the individual plants should be spatially located. For the non-chemical control methods it will be of significant importance to locate the culture plants for guidance of implements as hoes which will require that the crops spatial position should be located. Further, non-chemical methods may act "spot wise" in which case the spatial position of the weed plants needs to be known. Consequently, spatial information about plant location is an important objective to investigate for development of farming systems with reduced herbicide use.

However, having spatial information of a plant's position may also open other possibilities for development of new sustainable agricultural systems. Fertilisers may be placed optimally in respects to the culture plant, new machinery capable of working in mixed cropping systems (i.e. peas and corn in the same field), autonomous vehicles able to navigate according to a given plant structure which may make is possible to drill optimal according to erosion problems and so forth. Generally speaking, being able to locate the individual position of plants will make it possible to develop agricultural systems able to differentiate

their treatments at the lowest possible level, individual plants, and growing the crops with optimal use of input factors and minimum effect on the environment. Clearly, here are challenging perspectives.

Consequently the question raises of how may the spatial position of plants be found and made available to implements during operation ? We are able visually to distinguish between plants at a very early growth stage an ability used with hand weeding where the weed control may done at individual plant level. However, to base an agricultural system on man power will in the industrial countries be regarded as unrealistic due to both high labor cost and that labor in the future presumably will be a limited resource especially for such trivial tasks. As a result it is of significant interest to investigate potential use of sensors to solve the problem of locating plants. In this respect the use of Computer Vision vision is natural to investigate and indeed various researchers have done so in recent years [4, 5, 10, 8, 1]. However, as of yet, we have not heard of any reports on systems implemented and operating successfully at farm level.

1.2 The Challenge of Outdoor Computer Vision

Weed control and associated cultivation operations are all processes taking place in a highly changeable outdoor environment. The condition may change from the illumination being bright sunshine to a clear blue skylight causing the colours in the scene to change dramatically. Further, it may change from over cast to direct sunshine changing the characteristics of the illumination from a diffuse to uniform light source giving highlights and sharp shadows in the image. Together with potential mist and water drops on the leaves this gives images that will vary tremendously. Further, the vegetation parts of the image may account for only a small fraction of the image which may be sparsely distributed together with other objects as stones, plant residues, manure etc. These circumstances make it difficult for a computer vision system to operate robustly.

One way of solving these problem could be to constrain the environment by for example introducing shields that shades the image field together with artificial light sources. However, introduction of any such devices will make the system more vulnerable to mechanical breakdowns especially likely to happen in an agricultural environment. Also from a design perspective it is clearly better and more versatile to have an unconstrained sensor able to adapt to the current condition.

The main objective of this thesis is to investigate and develop methods that will make it possible for a computer vision system to adapt to the varying condition for the image formation. This will be done by investigating the intrinsic formation of images of vegetation in daylight and exploring this knowledge for the development of potential solutions for robust image acquisition and segmentation.

The approach is by a comprehensive modelling of the image formation process to get the necessary information for adapting to the current illumination condition. The modelling process involves the use of CIE daylight standard, radiation from a black Body, reflectance characteristics of objects, camera characteristics, and reflection models. The understanding obtained by the modelling enables the development of computer vision methods and sensors adjusted towards the situations under which they are expected to operate and the problems they are expected to solve. By using this a priori knowledge we show that it is possible to develop computer vision methods and sensors adaptable to outdoor illumination conditions.

1.3 Outline and Contents of Thesis

The thesis is outlined in two main parts **Computer Vision in Relation to Weed Control** and **Physic Based Analysis of Outdoor Images of Vegetation**. In the following a brief summary of the two parts will be given:

Part I. Computer Vision in Relation to Weed Control

Chapter 2. Computer Vision for Weed Control in Row Crops, an Analysis: Introducing Accessibility

In this paper a categorization of different weed control methods is established and their relation to computer vision is presented. Further, a taxonomy introducing the new concept *accessibility* for assessment and description of the performance of weed control treatments in various contexts, is introduced. It is a concept that may combine both primary and secondary parameters influencing the condition for the weed control treatment and this knowledge of their influence on the treatment is a necessary prerequisite for successful support of computer vision. The power of the concept is that it may both be used for assessment of the parameters influencing the weed control treatment and for the description of the objective for the treatments. Hence, it is able to serve as general basis for planning and execution of the weed control treatment.

Chapter 3. Computer Vision for Weed Control in Row Crops, a Survey

This paper reviews the "state of the art" of the field within the subject. It shows that the development is most advanced with techniques for detecting row structures and that the changing condition of the illumination in general is one of the major limiting factors for successful implementation of computer vision within weed control.

Part II. Physics Based Analysis of Outdoor Images of Vegetation

Chapter 4. On Physical Colour Image Understanding and Illumination Changes

A brief tutorial into the fundamentals of physical colour image understanding is presented. It especially focuses on the dichromatic reflection model. Next, a review is introduced of the CIE daylight standard and of how changing illumination conditions influence on the behavior of the dichromatic reflection model. This gives a natural introduction to the problem of colour constancy, for which alternative approaches are described.

Chapter 5. Classifying Illumination Condition from Two Light Sources by Colour Histogram Assessment

This chapter introduces a comprehensive model of the condition for the image formation for colour images acquired in a two light sources environment. The work gives insight into the dispersion of dielectrical objects' chromaticities under different illumination conditions. It may be used as support for methods relying on specific illumination condition as spectroscopic studies which in general rely on that the received light from the object of interest is diffuse (body) reflection. The paper includes two experiments, one on regular objects in a well controlled laboratory environment, demonstrating that the pixel points distribution is as expected. The second demonstration is potentially useful for supporting spectroscopic analysis of vegetation assessing the varying illumination condition of barley plants in the outdoor scene.

Chapter 6. A Model Based Daylight and Chroma Adaptive Segmentation Method

This paper introduces a segmentation method based on physical modelling of the image formation. The RGB-values are transformed into their chromaticities and intensities, the so called rgI-space. In this space it is possible to locate a plane approximately invariant to changes of the spectral composition of daylight (i.e. its correlated colour temperature). This representation of the pixel points makes it possible to classify also soil parts in shade where the colour will be biased towards that of the vegetation. The

reflection clusters in the rgI-space will typically be structured such that a plane "hinged" at the surface reflection locus may be used for segmenting vegetation from brown and greenish soil. The plane is rotated until the statistic Wilks lambda predicts maximum separation. This allows the method to adapt to not only the secondary reflection but also to small chroma changes of the vegetation.

Chapter 7. The Design and Operation of an Imaging Sensor for Detecting Vegetation

This paper investigates a sensor using the green, red and near infrared band (NIR) for detection of vegetation. It is well known that vegetation has a pronounced reflectance in the near infra red region which together with is low reflectance in the red gives a high contrast against soil. This property is widely used within remote sensing and for development of sensing devices for detection of vegetation. However, the background for images acquired in a natural environment does not solely consist of soil. There may be plant residues, stones, manure etc. present as well as objects that may have a reflectance characteristic almost similar to that of the vegetation. On the other hand these objects often appear less green than "living" vegetation so by using the green band it should be possible to get a sensor with a better basis for segmentation. The paper investigates the sensors' capability together with a physical modelling which includes an extrapolation of the CIE standard for daylight in to the near infra red region by radiation from a black body.

Chapter 8. Evaluation of an Imaging Sensor for Detecting Vegetation Using Different Wave band Combinations

This chapter presents a comparative study of which wave bands that will be most appropriate to use for classification of vegetation from background based on the dedicated sensor. A classification method (the "alpha" method) based on an understanding of the imaging processes is introduced. This is compared with two ratio methods (red/NIR and red/green) and two single band methods (NIR and green intensity). The Receiver Operating Characteristic Curve (ROC) method is used to evaluate the classifications on realistic test images. The alpha and ratio methods all perform reasonably well whereas the single band methods all perform significantly less well. The alpha and ratio methods have "best" thresholds that correspond with detectable histogram features when there is a significant amount of vegetation in the image. The physical basis for the alpha method means that there is a detectable mode in the histogram that corresponds with the "best" threshold even when there is only a small amount of vegetation. The single band methods do not produce histograms which could easily be analysed and so their use should be confined to simple images.

Chapter 9. Summary and Conclusion

In this final chapter a summary is given and the main contributions of the dissertation is outlined.

References

[1] J.V. Benlloch, A Sanchez, S Christensen, and M Walger. Weed mapping in cereal crops using image analysis techniques. In *AgEng96*, volume 2, pages 1059–1060, 1996.

[2] S. Bichel. Rapport fra hovedudvalget. Technical report, Danish Environmental Protection Agency, Danish Environmental Protection Agency, Copenhagen, Mars 1998. In danish.

[3] G. H. Brundtland. *Our Common Future*. Oxford University Press, 1987.

[4] R.J. Hagger, C. J. Stent, and S. Isaac. A prototype hand-help patch sprayer for killing weeds, activated by spectral differences in crop/weed canopies. *Journal of Agricultural Engineering Research*, (28):449–358, 1983.

[5] J.A. Marchant and R. Brivot. Real-time tracking of plant rows using a Hough transform. *Real-Time Imaging*, 1:363–371, 1995.

[6] E. Nordbo, S. Christensen, K. Kristensen, and M. Walter. Patch spraying of weed in cereal crops. *Aspects of Applied Biology*, 40:325–334, 1994.

[7] J. Rasmussen and J. Ascard. *Ecology and Integrated Farming Systems*, chapter Weed Control in Organic Farming Systems, Chap. 5, pages 50–67. John Wiley & Sons Ltd., 1995.

[8] J.M. Roger, E Molto, G Rabatel, and J Blasco. Design of a robotized, non-chemical weed controller. In *Bio-Robotics'97*, pages 229–235, 1997.

[9] K.E. Thonke. Research on pesticide use in denmark to meet political needs. *Aspects Appl. Biol.*, 18:327–329, 1988.

[10] E. Vrindts and J de Baerdemaerker. Optical discrimination of crop, weed and soil for on-line weed detection. In J.V. Stafford, editor, *Precision Agriculture '97*, volume 2, pages 537–544, 1997.

Part I

Computer Vision in Relation to Weed Control

Chapter 2

Computer Vision for Weed Control in Row Crops, an Analysis: Introducing Accessibility

Computer Vision for Weed Control in Row Crops, an Analysis: Introducing Accessibility

H.J. Andersen and E. Granum

Abstract

Development of computer vision systems within weed control has so far resulted in dedicated systems each relying on a particular technique. However, none of these systems are reported as being able to adapt to the varying light conditions for the process. In this paper it is suggested that development of robust computer vision has to be a combination of different techniques supporting each other. To facilitate such a combination, a new concept *accessibility* for assessment and description the performance of weed control treatments in a given context, is introduced. It is introduced as a taxonomy that may combine both primary and secondary parameters influencing the condition for the weed control treatments and knowing their influence on the treatment is a necessary prerequisite for successful support of computer vision. The power of the concept is that it may both be used for assessment of the parameters influencing on the weed control treatment and for description of the objective for it. Hence, it is able to serve as general basis for control of the weed control treatment.

2.1 Introduction

Increasing concern from the society over the potential side effects of using herbicides has given agriculture worldwide a great challenge in finding ways of reducing their use or substituting them by alternative methods. Thonke(1988) [12] reports that Denmark required a 50% reduction by 1997 with the period 1981-85 as reference. However, this was unfortunately not achieved by the end of 1997 [2]. A reduction in use of herbicides may be achieved by applying them selectively where the weed density exceeds the economic threshold rather then the entire field [5] or by using alternative methods as described in [9] .

The recent years' development in computers and sensors has opened new possibilities for use of computer vision within weed control, and research in this area has increased rapidly during the last decade. Methods for row guidance of implements [10, 8, 6, 4], detection of individual plant [13, 14, 7, 3], and weed density estimation [1, 11] have been investigated. These studies indicates that support from computer vision may contribute to reduce the use of herbicides but common for them is that they are restricted to investigation of different particular techniques. However, for development of successful support of computer vision we suggest to combine different sources of information which together describe the set of conditions for weed control. These sources of information may include a priori knowledge, information from various sensors, and in particular from the vision system's own image data.

The point is that through appropriate use of these information sources weed control may be made very selective and efficient with a minimum of environmental consequences. A suitable organization of the relevant information sources and criteria for selecting the most important bits of information for a given practical context is a major prerequisite for this approach.

Hence, the objective of this present chapter is to develop a structured description of the related information space. We aim to provide a taxonomy which covers not only the primary parameters which directly influence the conditions for weed control, but also secondary parameters that indirectly are important as they may determine how well supporting methods like computer vision may work.

The chapter is organized as follows. Section 2.2 characterizes various approaches to weed control such as to identify and describe those which can be improved by support from computer vision through real-time assessment of the condition for the weed control in the field. Section 2.3, introduces the concept of *accessibility* of crops with regard to weed control. We show how this concept through all its dependencies convey a constructive relationship between the weed control results achievable and the handling of the range of influencing parameters. Section 2.4, shows how the concept may be used on two images of barley plants acquired under different conditions for the image formation. Finally, section 2.5 summarizes and concludes the chapter.

2.2 Characterizing weed control treatments

Optimal performance of weed control methods as regards matching the individual need for control is today strongly limited because the information about this need is not made available. Some information may be acquired a priori and if not up to date when performing weed control, statistical methods for prediction may be applied. Preferably, however, it should include real-time and on-site acquired information which can be used to directly adjusting the control to the actual need. Exactly the ability to acquire real-time and on-site information is one of the strengths of computer vision and hence it is of significant interest to investigate its potential for supporting weed control methods.

2.2.1 Spatial resolution and selectivity of the weed control treatment

Two main characteristics of weed control treatments which describe if and how computer vision may be an aid are concerned with how differentiated the method is able to consider 1) the spatial variation of

the vegetation within a field and 2) the spatial location of the vegetation. For this purpose we define the concepts of *spatial resolution* and *spatial selectivity*, respectively.

Spatial resolution of a control treatment defines its ability to divide a field into subfields (sites) for individual treatment. The size of the sites will both be limited downwards by the spatial resolution of the weed control method and of the position system.

Spatial selectivity of the control treatment defines its ability to spatially differentiate its treatment between culture and weed plants. The selectivity will also be constrained by the spatial resolution of the control method and of the position system.

By these two concepts typical categories of weed control treatments may be characterized as shown in table 2.1:

Table 2.1: *Weed control treatments characterized by spatial resolution and selectivity.*

Treatment		Spatial resolution	Spatial selectivity	Knowledge acquisition	Computer vision task
Global		-	-	a priori	-
Site specific		coarse	-	a priori	-
Site adaptive		medium	-	real-time/on-site	weed density
Plant	-row	fine	medium	real-time/on-site	row detection
differential	-plant	fine	fine	real-time/on-site	plant detection

- = none

1) *Global treatment*, where the weed control is applied to the whole field without any adjustments to the need at the individual site. The concepts of spatial resolution and selectivity of the control method are therefore not applicable. This is the weed control method currently used at farm level.

2) *Site specific treatment*, where weed control is adjusted to the predicted need at each individual site according to a priori stored knowledge about the characteristics for each of them. This treatment is slowly emerging by support from GPS positioning systems.

Site specific treatment is clearly superior to global treatment, as regards matching the need for control. However, performance and result is limited to what can be predicted about the need for control from a priori knowledge as compared to having also real-time/on-site information on the condition at the very time of treatment. The following types of treatment are dependent on obtaining such real-time/on-site registered information.

3) *Site adaptive*, where the weed control is adjusted on-line according to the actual need at the particular site (e.g. weed density) at the time of treatment. For this it is necessary to have up to date, and in practical terms real-time measurement at an appropriate spatial resolution for weed control at each location. Support from computer vision may be of significant importance for acquiring this information.

The site adaptive treatment is clearly superior to both global and site specific treatment as regards matching the need for control. However, the control at a site will be limited to what the culture plants are able to withstand. To overcome this limitation it will be necessary to conduct the weed control in a spatially selective way, using some way of having real-time/on-site detection of the spatial location of the culture plants. The following type of treatment is depending on having such information.

4) *Plant differential treatment*, where the weed control is spatially differentiated between the locations of culture and weed plants, respectively. For the weed control method to operate spatially selective it is necessary to obtain the spatial location of either the culture or the weed plants in real-time. For sensing of this information computer vision may be of significant importance. As indicated in table 1 the treatment

may be split further into two sub groups, row and plant, respectively. The first one represents the case where it is possible to locate the plants as well structured rows and the second one where it is possible to locate individual plants.

2.3 Weed Control and Accessibility

From the above characterization of weed control treatments in relation to support from computer vision the main objectives may be identified as i) estimation of weed density, ii) detection of row structures, and iii) detection of plant structures, respectively.

These objectives may be seen as spanning a spectrum of *accessibility*, i.e describing how detailed or "close" one needs access to the culture plants in question to apply a given control method or how detailed or "close" the available weed control technology allows one to come to the culture plants.

The concept of *accessibility* may capture a fundamental aspect of weed control, such that below we will make an attempt to build a comprehensive taxonomy around it. A taxonomy that can accommodate descriptions of what is needed for a given weed control treatment, as well as what is- or could be-achievable under various conditions and using various supporting facilities.

2.3.1 Accessibility

The taxonomy of accessibility will be introduced by considering the accessibility of the row structure on one hand and of the individual plants on the other and both relative to a particular treatment.

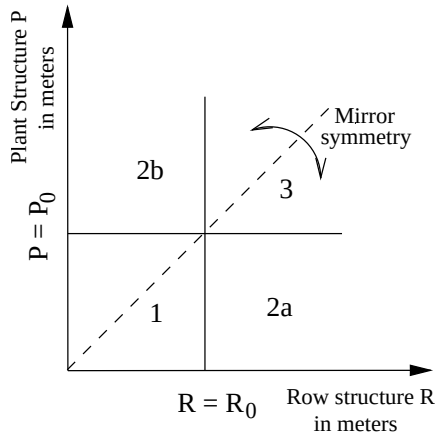


Figure 2.1: *Diagram of accessibility. Row and plant structure for a given crop structure subjected to two discriminant functions indicating if the structure is accessible or not for the control treatment. $R = R_0$ and $P = P_0$ denote the thresholds for the accessibility of the plant and row structure, respectively.*

The basic parameters for accessibility for a row crop structure are given in terms of distance between plants in the row and distance between the rows as it is sown. Figure 2.1 shows a Cartesian coordinate system relating these plant and row structures. For a given crop structure in combination with a given weed control treatment it is possible to determine thresholds in the plant and row structure, i.e. discriminant functions for both parameters, which the crop structure is not accessible to the control treatment.

Such discriminant functions, R_0 for the row structure (distance) R, and P_0 for the plant structure (distance) P, are included in figure 2.1. They divide the accessibility diagram into 4 characteristic areas which are labelled 1, 2a , 2b and 3. Suppose for example that the weed control method, hoeing, demands a row distance $R_0 = 30cm$. Then the accessibility of it will be, none in area 1, where neither row nor crop structure is accessible, and high in area 2a and 3 where row (and crop) are accessible. As may be noticed, area 2b has not been mentioned. It is a rare crop structure that seldom will occur, and it may be considered as mirror symmetrical to 2a around the dashed axis of figure 2.1. Hence areas 2a and 2b may degenerate to area 2.

However, for a weed control method supported by various sensors the accessibility is also dependent on a range of parameter all having different relations to accessibility. As an example consider a situation where the objective is to find a given weed plant. The weed plant is fully visible in the image field but the spatial sampling of the camera is so low that it is impossible to detect it. In this situation an accessibility corresponding to area 3 is required to fulfill the task but the camera is only providing an accessibility corresponding to area 1. Importantly, it is to recognize that the limiting parameter here in this case is not the crop structure but the low resolution of the camera, so the way to increase the accessibility will be to increase the spatial sampling of the camera or reduce the field of view (zoom) so it becomes possible to detect the weed plant.

Table 2.2: *Definition of accessibility together with grading of selectivity for the position system and indication of the potential weed control treatments according to table 2.1, due to the 4 areas defined by the discriminant functions in figure 2.1, indicating if plant or row structure is accessible.*

			Relative	Potential weed control treatments			
	Accessibility		Selectivity of	Site	Site	Plant differential.	
Area	Plant	Row	vision system	Specific	Adaptive	plant	row
1			none	x			
2a		x	medium	x	x		x
2b	x		medium	x	x	x	
3	x	x	high	x	x	x	x

Table 2.2, summarizes the level of accessibility for the different areas of figure 2.1, together with indication of the selectivity of the position system and the potential of conducting weed control treatments according to table 2.2.

As the potential of conducting site adaptive treatment is independent of identification of the crop structure, this may succeed in all areas of the accessibility diagram (figure 2.1). It may be constrained to operation in given area of accessibility but then it is important to notice that it maybe the method chosen for estimation of the weed density that are limiting the success and not the potential of performing the treatment.

This is in contrast to the plant differential treatments that are closely related to and limited by the actual accessibility (table 2.2), which will be the joint of the weed control method and the position system. (i.e. it will be difficult for the control method to perform with a higher spatial selectivity than that given by the position system).

2.3.2 Changes of accessibility

The conditions for the accessibility is depending on a range of parameters like how the crops were sown, the stage of growth, the machinery, current weather conditions etc.. Altogether a rather complex set of parameters determines the actual accessibility for a given weed control treatment at a given time. On the other hand, knowing the accessibility in a given context and being able to specify which parameters have the determining influence, may considerably increase the possibility of optimizing the treatment. In

this section it will be illustrated how such changes may be described in the accessibility diagram together with a classification scheme for the parameters affecting it in relation to support from a vision system.

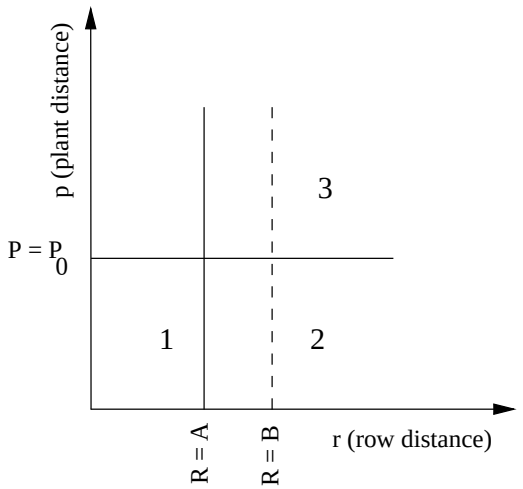


Figure 2.2: *Illustration of a parameter changing the limit of the accessibility in the row direction from A to B*

Figure 2.2 illustrates the accessibility in row direction subjected to a parameter decreasing the potential access of the row structure i.e. the disciminant function in the row direction moves outwards from $R = A$ to $R = B$. As an example this could be due to changes of the illumination condition from over over cast to direct sunlight. A prerequisite for successful implementation of computer vision in relation to weed control is to have control of and methods for description of such parameters affecting the accessibility in a given context. Exactly, this is one of the main problems in development of robust methods. Unfortunately, the parameter space for a complex operation as weed control is almost overwhelming.

In table 2.3 a framework for classification of the various parameter's influence on the accessibility and how they may support a computer vision in adapting to the actual condition, is introduced. In the table, the parameters are first characterized according to whether their variation significantly can change the accessibility or not within a season of growth or a treatment, respectively.

The parameters showing significant variation are first divided according to their variation being temporal (i.e varying by time) or spatial. Both of these groups are further divided into parameters where the variation may be predictable (pred.) or is stochastic (stoch.), respectively.

The first category in table 2.3 consists of parameters that are **given a priori**. They will be constant for a season or treatment, respectively, and as such they will **set** general conditions for the accessibility. Though, it is important to notice that they may be changed for better adaptation of accessibility for next season or next treatment (i.e. the row distance at establishment may be enlarged for better detection of a given crop structure).

Parameters in the next two categories **database information** and **sensor measurements** may all be of support for the vision system for better adjustment to the actual conditions. The difference between them is that the first category is parameters which may be given a priori whereas the second consists of less predictable parameters that may be registered by other sensors.

The last category **vision system adaptation** consists of parameters that the vision system itself must be able to adapt to. This adaptation may also be supported by information from other sensors.

Table 2.3: *Parameters influencing the accessibility are classified according to their variability during a season of growth.*

	Variation of parameters during season of growth					
	None		Significant			
Categories of	Season	Treatment	Temporal		Spatial	
parameters			Pred	Stoch	Pred	Stoch
Given a priori:	**X**	**X**				
- *crop specie*	*					
- *plant structure*	*					
- *implement*		*				
Database information:			**X**		**X**	
- *stage of growth*					*	
- *soil type*					*	
- *inclination of sun*			*		*	
Sensor measurements:				**X**		**X**
- *implement speed*				*		*
- *solar radiation*				*		*
Vision adapting:				**X**		**X**
- *shaded areas*				*		
- *rain/mist*				*		
- *local weed density*						*

Pred = predictable, Stoch = stochastic

2.4 Result: Using Accessibility

The intrinsic power of the concept of accessibility lies in its ability to describe in one coordinated context, the crop specie and how it was sown, growth condition, weed control treatment, the target of the treatment, and the conditions at the very time of treatment. Hence, if all parameters influence on the treatment is transformed into levels of accessibility this will then open the possibility for adaptation to the actual conditions. For illustrations of these considerations and to clarify how the concept accessibility may be used consider the following example demonstrating the influence of angle of tilt ψ (fig. 2.3) of camera and accessibility of individual plants and plant rows.

Figure 2.4 (a) and (b) includes images of barley at growth stage 2 (Feekes scale). In (a) the image plane

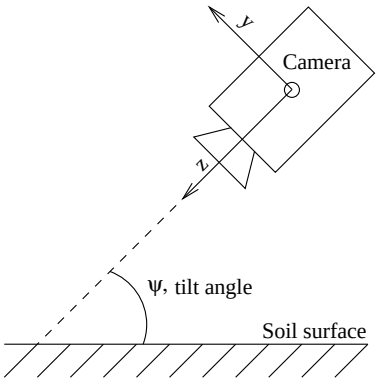


Figure 2.3: *Definition of ψ angle of tilt of camera. x-axis is perpendicular to paper.*

is parallel with the soil surface ($\psi = 90°$). Looking closer at the image it may be noticed (though hard to see) that it is possible to locate the barley and weed plants individually, hence the accessibility will be in area 3 and weed control may potentially be performed as plant differential at plant level.

In (b) the tilt angle ψ has been changed to 60° and the image gets a perspective projection. Due to the overlap of the barley plants in rows it becomes impossible to identify them individually. However, in between the rows it is still possible to identify some plants individually, so in this case the accessibility may be considered individually into in the rows and in between rows (i.e. for culture plants versus weed). Concentrating on accessibility of the culture plants it is clear from the two images that potential detection of rows structures should be easier from (b) that (a), but may potentially be performed in both.

(a) (b)

Figure 2.4: *Images of row sown barley. (a), image plane parallel to soil surface i.e no angle of tilt for the camera. (b), perspective projection of rows.*

In this example it is illustrated how the tilt angle of the camera influences the accessibility of the individual plants. However, many other parameters may also influence the accessibility and the question is how it may be assessed in a given context. For the above segmentation of the image into vegetation and background, the longest perimeters of the connected plants areas (plant blobs) may be a method for assessment of whether the individual plants may not be found in the plants rows or not. Further, combining maximum perimeter length and direction with expected plant height may even make it possible to assess the possible maximum tilt angle at which individual plants may be found.

In chapter 5 of the thesis a metric for assessment of the current illumination condition will be introduced. Looking closer at images in 5.14 (Chapter 5) it is clear that the illumination condition has major impact on the accessibility of potentially weed plants in between the barley rows. With sunlight present in the images of figure 5.14 (a) and (c) the exposure in the shaded areas will be "poor" and the signal to noise ratio high and as a result segmentation of these into vegetation and background will be prone to give faulty classified areas. Looking at images 5.14 (b) and (d) it is evident that the condition for the images formation gives much better exposed areas around and in between the barley plants. Clearly, the changing illumination condition may also have impact on precision of row detection algorithms and other computer vision operations.

2.5 Conclusion

A new concept accessibility is introduced as a taxonomy for adaptation to the actual condition for a vision system and other support tools within weed control. The power of the concept is that it may both be used for assessment of the parameters influencing the weed control treatment and for description of the objective for it.

Specifically, using it within computer vision support of weed control gives a strategy which may be divided into three stages. First we may try to adjust/adapt the current method to the actual conditions, second methods of different and appropriate characteristics may be combined, and finally a change to another more appropriate method may be necessary to solve the specific task of the weed control method in the given context.

Adapting to the concept of accessibility will encourage and motivate a change in the research strategy towards development of systems consisting of combinations of different techniques suitable for specific conditions. We believe that this could be a way forward for development of successful support of weed control by computer vision.

References

[1] C. Andreasen, M. Rudemo, and S. Sevestre. Assessment of weed density at an early stage by use of image processing. *Weed Research*, 37:5–18, 1997.

[2] Anonymous. Status for miljøministerens handlingsplan for nedsættelse af forbruget af bekæmpelsesmidler. Technical report, The Danish Ministry of Environment and Energy, Nov 1997.

[3] M.C. Cardenas-Weber, F.F. Lee, G.E. Miles, and D.E. Guyer. Plant features measurements with machine vision and image processing. *American Society of agricultural Engineers*, fiche no. 88-1541:11, 1988.

[4] J.A. Marchant and R. Brivat. Real time tracking of plant rows using a hough transform. *Real-Time Imaging*, 1:363–371, 1995.

[5] E. Nordbo, S. Christensen, K. Kristensen, and M. Walter. Patch spraying of weed in cereal crops. *Aspects of Applied Biology*, 40:325–334, 1994.

[6] H.J. Olsen. Determination of row position in small-grain crops by analysis of video images. *Computer and electronics in agriculture*, 12(2):147–162, Mar 1995.

[7] W. Petry and W. Kuhbach. Automatisierte unterscheidung von unkrautarten nach formparametern mit hilfe der quantitativen bildanalyse. *Zeitschrift fur Acker- und Pflanzenbau*, 163(5):345–351, Dec 1989.

[8] F. Pla, J.A. Sanchiz, J.A. Marchant, and R. Brivot. Building perspective models to guide a row crop navigation vehicle. *Image and Vision Computing*, 15:465–473, 1997.

[9] J. Rasmussen and J. Ascard. *Ecology and Integrated Farming Systems*, chapter Weed Control in Organic Farming Systems, Chap. 5, pages 50–67. John Wiley & Sons Ltd., 1995.

[10] J.M. Sanchiz, F. Pla, J.A. Marchant, and R. Brivot. Structure from motion techniques applied to crop field mapping. *Image and Vision Computing*, 14:353–363, 1996.

[11] J.V. Stafford and J.V. Benlloch. Machine-assisted detection of weeds and weed patches. In J.V. Stafford, editor, *Precision Agriculture '97*, volume 2, pages 511–518, 1997.

[12] K.E. Thonke. Research on pesticide use in denmark to meet political needs. *Aspects Appl. Biol.*, (18):327–329, 1988.

[13] L. Tian. *Knowledge Based Machine Vision System for Outdoor Plant Identification.* PhD thesis, University of California, Davis, Dept. of Biological and Agricultural Engineering, 1995.

[14] D.M. Woebecke, G.E. Meyer, K. von Bargen, and D.A. Mortensen. Shape features for identifying young weeds using image analysis. *Transactions of the ASAE*, 38(1):271–281, Jan 1995.

Chapter 3

Computer Vision for Weed Control in Row Crops, a Survey

Computer Vision for Weed Control in Row Crops, a Survey

H.J. Andersen and E. Granum

Abstract

A survey of the literature within the field computer vision for weed control in row crops, shows that many promising methods are emerging, especially within plant row detection but none is yet successfully implemented at farm level. One of the main limiting factors for successfully implementation may be due to the lack of methods that may adapt to the changing daylight as illumination source. The survey also indicates that combination of methods may be useful and possibly necessary to achieve robust performance.

3.1 Introduction

In chapter 2 the concept *accessibility* was introduced as taxonomy to support the problem of adjusting, combining or selecting the most appropriate weed control methods in a given context. In this chapter a survey through the literature within the field will be presented to give a general picture of "the state of the art" and to see what parameters are limiting the accessibility. The survey concentrates on studies based on computer vision techniques where the images are formed by a digital camera. This is in contrast to studies where the monochromatic reflection from vegetation is used for discrimination between soil, weeds and crops [37, 11, 3] (i.e. spectroscopic analysis).

Previously a survey completely devoted to the use of computer vision within weed control in row crops has not been available, but a couple of more comprehensive surveys including computer vision should be mentioned. A review on the use of computer vision techniques in agricultural processes can be found in Tillett(1991) [36]. It classifies the main areas for image analysis in agriculture and identifies the generic problems that are relevant for biological objects. Another review on guidance of agricultural field machines can be found in Tillett(1991) [35]. This covers the traditional technologies such as leader cables and mechanical methods and more recent methods like computer vision, satellite navigation and other optical systems.

3.2 Survey

This survey will be organized according to the objectives of the individual papers regarding computer vision within weed control in row crops:

- estimation of weed density
- detection of row structures
- detection of individual plants
- basic investigation into computer vision technique for identification of plants

with the first three categories according to the taxonomy introduced in chapter 2. Some papers will fit perfectly into one of the first three categories whereas other will cover several. Especially, research in identifying specific plant species and segmentation of plants from non-plant material is a common problem addressed in many studies. Hence, a fourth category *basic investigation into computer vision technique for identification of plants*, is included.

3.2.1 Basic research in computer vision for identification of plants

Segmentation of living plant material from non-plant background

Segmentation of images into living plant material and non-plant background is one of the important steps in the image processing. A main problem in development of segmentation methods is to get a valid assessment of the ground truth. For scenes of plant and non-plant background this is especially a problem due to the large deviation in colors and spatial distribution of the different classes, traditionally regarded as two, background, and vegetation. Hence, the performance of the methods is often compared with human assessment.

Gausian distribution models
A segmentation algorithm for mono chromatic images based on a Bayes classifier is introduced in Reid and Searcy (1988) [25]. Their assumption is that the pixel distribution of vegetation and soil will form

two different intensity classes that can be regarded as independent Gaussian random variables. As long as the grey-scale distribution was bimodal the classifier gave reasonable results. It could be improved by providing a priori knowledge about the mixture coefficient to handle the cases where the bimodality of the histogram was poor. Andreasen *et al* (1997) [1] also used the assumption that the pixel distribution of plants and soil may be regarded as a mixture of two Gaussian distributions. The histogram was formed by the normalized green stimulus (i.e. the green chromaticity). As long as a sufficient amount of vegetation was present in the image and the lighting condition gave a good contrast between the soil and vegetation, the method was reported to work well.

Near infrared images
Brivor and Marchant (1996) [4] took images in the near-infrared region of cauliflowers under diffuse light condition. They subtracted the gradient of the image from the original image before performing a simple threshold based on the grey-level distribution of the pixel points values. The gradient operation was found to eliminate weeds that was attached to the crops. The performance of the method was assessed by comparing it with human classification. The results showed that a "satisfactory" classification of 92% was attainable for 'good' images and 73% for bad images. The method was later used by Sanchiz *et al.*(1996) [26] who report that it only performed satisfactory in "good" (diffuse) illumination condition.

RGB-values
A segmentation algorithm exploiting the distribution of pixel values in the RGB-cube is introduced in Tian (1995) [34]. To avoid the difference in intensity the RGB-values were transformed into their chromaticities before a k-means cluster analysis (k=2) was performed. The result of the cluster analysis was used to train a Bayesian classifier using the clusters as classes for vegetation and background, respectively. The classifier was found to be stable under the condition in which it was trained, but when the correlated color temperature of the illumination changed significantly, it failed. As a significant change a shift from 5700K to 9500K, was reported.

A study of different color indices for segmentation of plant and non-plant material in a greenhouse environment under natural lighting is reported in Woebbecke *et al.* (1995b) [39]. The mean values of shaded and unshaded surfaces of vegetation and soil manually located in the images were compared using a student t-test. The contrast 2g-r-b (where rgb are the chromaticities of the RGB-values) which corresponds to thresholding the green chromaticity at g=1/3 and an index called "the modified hue" were found to be the best to discriminate between the pixel values of the vegetation and non-vegetation surfaces. Unfortunately, no actual performance of the method was reported.

Recently, Meyer *et al.* (1998) [18] introduced a contrast called Excess Green (ExG):

$$ExG_{x,y} = \sum_{x=1}^{n_x} \sum_{y=1}^{n_y} \frac{2G_{x,y} - R_{x,y} - B_{x,y}}{N},$$

where N is equal to the total number of pixel in the image. x,y denotes pixel location in the image plane. A contrast measure clearly inspired by the study of Woebbecke *et al.* (1995b) but totally empirically based. The method is implemented in connection with a low pass filter to remove small bright soil areas before tresholding the image with a fixed value of 20 on 0-255 scale. The method was used on 480 images all acquired in a greenhouse environment between 11:30am and 12:30pm to ensure good lighting conditions. The reported segmentation result is that it successfully discriminated plants from soil 100% under the constrained condition.

Other wavebands
A related study on the optimum selection of wavelength pair in the interval from 400 to 700nm to differentiate between plant material and soil is reported in Han and Hayes (1990) [12]. The optimum pair was found to be 550 and 700 nm which corresponds to green and red. This is in agreement with the study by Woebbecke *et al.*(1995b) [39] who also consider a contrast between green and red. Another study investigating potential filter pairs is reported in Chaisattapagon and Zhang (1991) [5] that compared the ratio between 14 color filter pairs to discriminate between soil, wheat, sand/stone and weed

leaf and stem. The filters green, red, deep blue, and blue-band reject where found to be the most suitable.

Dichromatic reflection model
Recently, Pla *et al.*(1997) [23] used a segmentation method based on the properties of the dicromatic reflection model, as introduced in Pla *et al.*(1993) [22]. The method utilizes that reflection from dielectrical objects form planes in the RGB-cube spanned by their body- and surface reflection clusters, respectively. Transforming the RGB-values into spherical coordinates and neglecting the radius gives an intensity independent representation of their values by the azimuth and elevation angle which makes it possible to correctly classify areas of an object showing combined surface and body reflection. In the azimuth-elevation plane a classifier was trained by samples of vegetation and background under a given illumination. The method was reported to work satisfactorily in the illumination under which it was trained. As pointed out by the author the classifier may not work if the light changes in spectral composition. Unfortunately, no performance test including its sensitivity to spectral changes of the illumination was reported.

Summary: Segmentation of living plant material from non plant background
Evaluating the performance of the different proposed segmentation methods is in general very difficult due to the lack of information in the studies regarding test condition and the constrained variation of the condition of the image formation. Especially, there seems to be a need for more comprehensive investigation into the consequences of images acquired with variations of the outdoor illumination conditions. The studies of Pla and Tian barely touches the subject but only includes assessments of their classification methods under the condition in which they are trained. Also the study by Woebbecke *et al.* shows promising results but all experiments are done in a green house environment at midday.

Surprisingly, for all studies involving training of classifiers none of them are taking advantage of the reported reflectance characteristic of vegetation in the literature [40, 28]. Possibly, using these together with camera characteristics and CIEs' daylight standard [2] could open for development of more versatile segmentation methods taking the underlying physical process for the image formation into account.

Identification of individual plants

Identification of individual plants may be used both within site adaptive and plant differential treatments, as introduced in chapter 2. Firstly as a support for estimation of the weed density and secondly for detection of individual plants which may be used also for detection of plant rows. A wide variety of approaches for recognition of plant species has been investigated. Generally, the problem is to find a set features that may be used in a classification scheme.

Spatial frequency based properties
Fourier-based methods has been introduced for description and classification of the appearance of leaves, shapes, and canopies. A one dimensional Fourier-transform to quantify leaf contours is reported in Kincaid and Schneider (1983) [13]. The amplitudes of the first 36 frequencies was found the be enough for descriptions of even the most complicated leaf shape. The study involved 1216 not occluded and flat leaves from 7 species. To handle partially occluded leaves Franz *et al.* (1991b) [8] introduced the Fourier-Mellin correlation for matching the curvature function of leaf shape for 5 different plant species. The matching had a success rate of 73% for 124 unknowns. More recently, Critten (1997) [6] introduced fractal dimension relationships and values as a method for discriminating between different plant canopies. Monochrome pictures were taking of seven weeds and three crop species at heights of 2 and 1.25 m, respectively. The results showed that fractal indices derived from three vectors dimension (treating grey level as a contour) have a potential for discrimination between crop and weeds. However, no actual discrimination analysis was performed.

Spectral properties of reflection

Use of spectral properties of leaves for classification of plant species was investigated by Franz *et al.* (1991a) [7]. Along the leaves' margins a set of patches of 2x20 and 10x10 pixels were manually selected. From these the features' mean, variance, and skewness, were calculated in the wavebands 610-1000nm, 480-590nm and 400-500nm, corresponding to the NIR, red and blue band. The features were evaluated for classification first using 48 leaf regions manually extracted giving a correct classification of of 94% and later with an automatically boundary-tracing algorithm which, however, gave an decrease in the classification to 76% using 66 leaf regions.

Semantic shape features
Several studies have involved the scheme of finding suitable semantic features for classification of different plant species. Woebbecke *et al.* (1995a) [38] classifies monocots from dicots with a performance of 74% for plants less than 14 days old using aspect ratio of the leaves as feature and a performance of 89% for plants older than 14 days using the first invariant central moment of the leaves. Petry and Kubach (1989) [21] used 9 features to form 5 canonical parameters to distinguish between 6 weed species. The average rate of correct classification was 82% for a total of 300 weed plants.

Semantic and Syntactic shape features
Tian (1995) [34] presents a comprehensive study recognizing tomato plants captured in field condition with several or none plants present in the image. It involves use of semantic features combined with a syntactic method for whole plant identification. First, the tomatoes cotyledons are found in the image using the semantic features compactness, elongation, logarithm of the ratio height to width and the ratio of length to perimeter using a Bayes classifier. Afterwards the tomato plant's stem location is found by introducing a syntactic method to describe the structure of the plant. The algorithm was able to find and identify 61-82% of all the individual crop plants in 270 frames. The range in classification success was found to be mainly dependent of the quality of the raw images (i.e. illumination conditions, wind etc).

Texture
An example of plant identification by color co-occurrence matrices is introduced in Shearer and Holmes (1990) [27]. As input in the co-occurrence matrix a variety of different texture features calculated for intensity, saturation and hue is used. The best classification was achieved using 3 intensity and 4 hue features with a classification accuracy of 91% for 7 plant species.

Others
Studies of a more "specialized nature" are to be found in Franz *et al.* (1995) [9] and McDonald and Chen (1990) [17]. The first developed algorithms for extraction of leaf boundary for either complete visible or partially occluded leaves. The latter introduced mathematical morphology procedures for image processing in agriculture with an example of discriminating african violet from ivy leaves.

Summary: Identification of individual plants
Common for most of the studies reported is that they typically are performed under well-controlled experimental conditions with the plants carefully located in image. However, images of plants taking under field conditions are influenced by many factors disturbing the image formation of the plants. This may both be environmental factors as wind, illumination condition and lack of regularities in the crop geometry.

Further, no study has yet involved the use of 3D representation of the plants. This might be a way forward to be able to move from the well controlled laboratory condition to the uncertainties of outdoor scenes.

3.2.2 Estimation of weed density

A necessary prerequisite for site adaptive treatment is to have an estimate of the weed density at the particularly location. The idea of making a distinction between "real-time" assessment or an a priori "mapping" of the density (as in site specific treatment as introduced in Chapter 2) is also introduced in Nordby*el al.*(1994) [19]. However, they only treat the latter by subdividing the field sites into 20x20 meter squares where they a priori determine the weed pressure.

Thompson *et al.* (1991) [33] points to the possible use of computer vision for weed density estimation but argue that it will not be suitable for real-time control of a selective sprayer in cereal crops, because the crop will severely cover the potential space for weed detection at a very early stage in a crop growth cycle. Thompson *et al.* [32] showed that an autumn-sown crop within the first few weeks after germination will prevent weeds from being detected in 60-90% of the field area.

Andreasen *et al.* (1997) [1] investigates weed density assessment in field condition comparing an automatic method with an interactive approach. The automatic estimator is based on two features, segment size with and without border cutting. It was concluded that these two features were insufficient to give a reliable estimate of the weed density. In the final conclusion Andreasen *et al.* states that utilization of computer vision for reliable estimation of weed density on farm level cannot be expected in the near future but the use of the method as an aid in research activities may be realistic within a few years.

Estimation of weed density in near-ground images in cereal crops is reported in Stafford and Benlloch (1997) [31]. Classification of weed plants are based upon whether they are located in the cereal crop row and that their shape features are distinct from those of cereal leaves. To distinguish between the two classes several geometric features were used to train a Bayes classifier. The results showed a correlation of 85% between the computer vision based estimation and manual visual surveying. The computer vision based technique tended to underestimate at high weed densities and overestimate at low densities. The sources for misclassification where found to be irregularities in the crop geometry, overlap and coverage between weed seedlings, and the fact that not all weeds are embedded among the crop plants. Also the natural light condition was found to be a severe source for misclassification due to highlights and shadows which made the segmentation process difficult.

Summary: Estimation of weed density
Reported studies within this area are restricted to very limited conditions. The major problems for successful implementation may be changing conditions for the image formation and irregular and/or occluded appearance of the plants.

Common for all studies are that they try to find weed plants and estimate the weed pressure without any use of a priori knowledge of what species are to be expected at the current location. Clearly, support from weed maps together with information about expected growth stage may be of significant support for a computer vision system in "narrowing down" which weeds to search for (i.e. characteristic weeds for the location) and on what growth stage. To build weed maps comprehensive off-line analysis of acquired images may be done. Hence, a more holistic approach trying to combine this information with the on-site real-time acquired information may be a way forward.

3.2.3 Detection of row structure

This section presents the studies that assumes that the crop forms rows or if they find position of individual plants they only make use of this information for determination of plant row locations. The methods are divided into two categories *Perspective* and *Orthogonal*, related to whether the camera has an angle of tilt relative to the soil surface which is regarded as being locally plane, as illustrated in chapter 1, fig. 2.3. .

"Perspective" methods
The Hough transformation has in several studies been used to find row structures in field images of plant

rows. One of the advantages of the transformation is that it is able to find predominant line structures in an image. This makes it able to detect incomplete lines which often is the case for row crops.

Reid and Searcy (1986) [24] investigated the possibilities to use the Hough transformation to find the row structure in cotton fields. Assuming that the soil surface is plane and that the camera doesn't pitch or roll and has a constant angle of tilt relative to the ground, the images will be taking in a perspective view in which it is possible to estimate both the lateral off-set and heading angle. The study both involved an evaluation of the method on simulated rows and on real images of cotton. Their conclusion was that it should be possible to guide a tractor by the method but no actual implementation was done.

A more dedicated version of the Hough transformation that takes advantage of the a priori knowledge about number of rows and row space in the image is introduced in Marchant and Brivot (1995) [15]. The algorithm is evaluated in Marchant (1996) [14] in three different crops: cauliflowers, sugar beet and widely spaced double rows of wheat. The typical errors reported when compared with a human assessment of row position in the images was 18 mm of lateral offset and 1° heading angle. An implementation of the method fused with odometry data using a Kalman filter on an autonomous vehicle working in transplanted cauliflowers, is reported in Marchant(1997) [16]. Trials showed that the standard deviation of the vehicle's path with respect to the row structure was 20 mm for the whole set of runs, 2 beds of 70 m. The error where within a band of ±45 mm with one short excursion to 66 mm. This accuracy was concluded to be sufficient for operation in cauliflowers.

Recently, a skeleton method for determination of plant row location, is introduced in Pla *et al.* (1997) [23]. The method utilizes that perspective images of plant rows will introduce a vanishing point. First the center of the soil region between the plant rows is found by thinning. To remove noise branches *hairs* and *double branches* a skeleton pruning algorithm is applied. After this operation the vanishing point is determined by a cluster analysis clustering the coordinates of all possible points resulting form intersections of lines. From this list of clusters the one supported by most points is assumed to come from the structure of the parallel plant rows and hence represent the vanishing point. The method was evaluated in straight lines of row crops with good performance. The major drawback of the method is that it for reliable detection of the vanishing point needs a long forward sight and plant rows forming straight lines. This may in many practical situations not be possible due to the topography of the field, that the rows are not drilled in completely straight lines, or that at the ends of the rows the length will not be sufficient to build the perspective model.

"Orthogonal" methods
Another approach for finding the lateral offset of an interrow-weeder relative to the plant rows is presented in Olsen (1995) [20]. The sum of the pixel element's grey values in the travelling direction is calculated. To this sum-curve two methods for finding the plant rows position are introduced: 1) a least square fit of a sinusoidal curve and 2) a FFT low-pass filtering algorithm where the maxima of the back shifted signal is used for control. The methods were evaluated in cereals and sugar beets using an interrow-weeder. The result showed that the methods worked with an accuracy of about 10 mm in cereal but none of them were found to be applicable in sugar beets.

An algorithm to determine the centerline of a plant row formed by individual plants is described in Slaughter *et al.* (1995) [29]. For every plant in the plant row a histogram was calculated in the travelling direction and its median was found. If the median of a given plant differed significantly from the average of all plant medians it was disregarded. The centerline was then based upon the remaining medians and the location in the image of the corresponding plants. The method was evaluated in Giles and Slaughter (1997) [10] that used it for guidance of a band sprayer in tomatoes and lettuce. It was found to have a 95% confidence limit for the desired tool position in tomatoes of ±12 and ±9 mm for a plant height of 38 and 76 mm, respectively, and a travelling speed of 8 km/h. It allowed for a decrease in spray application rate by 66 to 80% compared with traditional broadcast spraying.

Summary: Detection of row structure

It may be noticed that the activity within development of methods for row detection have been quit intensive the last years. Many methods are emerging operating with a satisfactory precision for many crops with an acceptable forward speed. However, none are yet reported implemented at farm level.

In the survey methods have been divided into two approaches for detection of the row position related to the angle of tilt of the camera relative to the soil surface. Clearly, both have advantages and disadvantages. The first category of *perspective methods* are relying on the rows being drilled in straight lines with small deviation, that the soil surface may be regarded as being locally plane, and that the field topography enables a long forward sight. The most prominent for the method proposed by Pla *et. al.* The methods advantage is that they are less sensitive for missing plants and locally high weed density. This is in contrast to the *orthogonal methods* operating more locally in front of the implement. These are less sensitive for deviation in the precision of the drilling and the topography but more sensitive for missing plants and locally high weed density that may occlude the culture plants.

Common for the studies reported are the constrained conditions for which there are evaluated. None, is reported adaptive besides the condition in which it is trained. A major objective for future research will be to evaluate and/or develop methods that are able to adapt to the variation for the image formation. Especially, variation of the climate (illumination, wind, mist etc) and locally condition for the crop growth may be recognized as limiting factors that needs to be investigated. However, the recent progress within the area clearly indicate that it should be possible to develop methods applicable for implementation at farm level.

3.2.4 Detection of plant structure

This section consider studies that only focus on identification of individual plants in a given plant row. Many of the studies in the paragraphs on "Identification of individual plants" treats the problem of determining presence of individual plant species but they do not consider the problem of finding their spatial location. Also, some of the studies in the previous section explicitly finds the locations of the individual plants but only uses the information for determination of plant rows and not individual plant location.

One exception though made in the study by Tian (1995) [34], which also investigates the potential use of a priori knowledge about the plant spacing for better locating the stem of tomatoes. One preliminary experiment where the tomatoes was planted with 50 mm spacing (the plants did not touch each other) indicated that the method could be used with success. The drawback of it is the need for a special precision planter. Another, method also having the position of the individual plants implicit is the centerline method, presented in Slaughter *et al.* (1995) [29].

Navigation of an autonomous vehicle through a field of individual plants by a tracking algorithm for 2 dimensional objects is introduced in Sanchiz *et al.*(1995) [26]. The study introduces a clustering and shape description method that is able to handle splitting and merging of regions in subsequent images. The actual displacement in subsequence images of the plants is found by a Hough transform method estimating the best match for the rotation and translation that is later used in a minimization algorithm to find the motion parameters of the vehicle. The method was evaluated navigating a vehicle through a field with 9-12 individual plants in the image and was found to be able to recover the plant structure "quite" accurately. The main limiting factors are that the individual plants have to form separate cluster and shall at least appear as a complete plant in two consecutive frames.

Southall *et. al* [30] introduce a method using a model of the a priori known plant pattern for both navigation of an autonomous vehicle and for positioning of individual plants for a herbicide treatment system. The method includes a self-starting routine also using the model of the plant pattern. For navigation it was compared with the method proposed by Marchant [14] and human assessment of the plants locations. It is reported to work as well as the previously implemented Hough transform method for navigation and for localization of individual plants to be within a mean difference of 14 mm with plant spacing ranging from 600- 1700 mm, which is as accurate as found in human assessment.

Summary: Detection of plant structure
The problem of finding plant structures has not yet received as must attention as the row structures. Obviously, the problem is more complex but as indicated in Tian and Southall having a plant pattern model may significantly support the task. It is an area that needs more investigation as many of the crops, where it is of interest to find the individual plant position, is in "high" value crops and/or in crops of significant environment concern for the consumer (sugar beets, vegetables etc.) hence having a greater potential for paying the necessary investment in technology.

3.3 Discussion and recommendation

Generally, the survey shows that research within the area is investigations of many different "traditionally" used computer vision methods applied within weed control but almost all constrained to very specific conditions of the weed control operation. Clearly, there is a need to look into more holistic approaches combining a priori knowledge with on-site/real-time observations for development of a more versatile system that may also be used for other crop cultivation purposes i.e. fertilization, pest management etc.

The most promising area for implementation at farm level may be recognized as being detection of row structures for guidance of hoes and navigation of autonomous vehicles.

A common limiting factor for all studies is the limited evaluation under varying condition of the image formation. Especially adaptation to changing illumination condition as occurring for outdoor images needs deeper investigation for development of robust methods.

As stated in chapter 2 we suggest and we encourage for a change in the research strategy towards development of systems which uses combination of techniques suitable for each their specific condition. We believe that this will be the only way forward for development of successful support of computer vision within weed control.

References

[1] C. Andreasen, M. Rudemo, and S. Sevestre. Assessment of weed density at an early stage by use of image processing. *Weed Research*, 37:5–18, 1997.

[2] Anonymous. Colorimetry. Technical Report 2. edition, Commision Internationale de L'Eclairage (CIE), Vienna, 1986.

[3] T. Borregaard. *Application of imaging spectroscopy and multivariate methods in crop-weed discrimination*. PhD thesis, The Royal Veterinary and Agricultural University, Copenhagen, Denmark, 1997.

[4] R. Brivot and J.A. Marchant. Segmentation of plants and weeds using infrared images. In *Proceedings of the Institution of Electrical engineers, Vision, Image and Signal processing*, volume 143, pages 118–124, 1996.

[5] C. Chaisattapagon and N. Zhang. Weed detection using machine vision. *American Society of Agricultural Engineers*, (91-3508):15, 1991.

[6] D.L. Critten. Fractal dimension relation.ship and values with certain plant canopies. *Journal of Agricultural Engineering Research*, 67:61–72, 1997.

[7] E. Franz, M.R. Gebhardt, and K.B Unklesbay. The use of local spectral properties of leaves as an aid for identifying weed seedlings in digital image. *Transaction of the ASAE*, 34(2):682–687, Mar 1991a.

[8] E. Franz, M.R. Gebhardt, and K.B. Unklesbay. Shape description of complete visible and partially occluded leaves for identifying plants in digital images. *Transaction of the ASAE*, 34(2):673–682, Mar 1991b.

[9] E. Franz, M.R. Gebhardt, and K.B. Unklesbay. Algorithms for extracting leaf boundary information from digital images of plant foliage. *Transactions of the ASAE*, 38(2):625–633, 1995.

[10] D.K. Giles and D.C. Slaughter. Precision band spraying with machine-vision guidance and adjustable yaw nozzles. *Transactions of the ASAE*, 40(1):29–36, 1997.

[11] R.J. Hagger, C.J. Stent, and S. Isaac. A prototype handheld patch sprayer for killing weeds activated by spectral differences in crop/weed canopies. *J. of Agric. Res.*, 28:349–358, 1983.

[12] Y.J. Han and J.C. Hayes. Soil cover determination using color image analysis. *Transactions of the ASAE*, 33(4):1402–1408, 1990.

[13] D.T. Kincaid and R.B. Schneider. Quantification of leaf shape with o microcomputer and fourier transform. *Canadian journal of Botany*, 61:2333–2342, 1983.

[14] J.A. Marchant. Tracking of row structure in three crops using image analysis. *Computers and electronics in agriculture*, 15:161–179, 1996.

[15] J.A. Marchant and B. Brivot. Real-time tracking of plants rows using a hough trasnform. *Real-Time Imaging*, 1:363–371, 1995.

[16] J.A. Marchant, T. Hague, and N.D. Tillett. Row following accuracy of an autonomous vision guided agricultural vehicle. *Computer and Electronics in Agriculture*, 16:165–175, 1997.

[17] T. McDonald and Y.R. Chen. Application of morphological image processing in agriculture. *Transactions af the ASAE*, 33(4):1345–1351, 1990.

[18] G.E. Meyer, T. Mehta, M.F. Kocher, D.A. Mortensen, and A. Samal. Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. *Transactions of the ASAE*, 41(4):1189–1197, 1998.

[19] E. Nordbo, S. Christensen, K. Kristensen, and M. Walter. Patch spraying of weed in cereal crops. *Aspects of Applied Biology*, 40:325–334, 1994.

[20] H.J. Olsen. Determination of row position in small-grain crops by analysis of video images. *Computer and electronics in agriculture*, 12(2):147–162, Mar 1995.

[21] W. Petry and W. Kuhbach. Automatisierte unterscheidung von unkrautarten nach formparametern mit hilfe der quantitativen bildanalyse. *Zeitschrift fur Acker- und Pflanzenbau*, 163(5):345–351, Dec 1989.

[22] F. Pla, F. Ferri, and M. Vicens. Colour segmentation based on a light reflection model to locate citrus for robotic harvesting. *Computers and Electronics in Agriculture*, 9:53–70, 1993.

[23] F. Pla, J.A. Sanchiz, J.A. Marchant, and R. Brivot. Building perspective models to guide a row crop navigation vehicle. *Image and Vision Computing*, 15:465–473, 1997.

[24] J.F. Reid and S.W. Searcy. Detecting crops rows using the hough transform. *ASAE Summer Meeting*, Paper No. 86-3042, 1986.

[25] J.F. Reid and S.W. Searcy. An algoritm for separating guidance information from row crop images. *Transaction of the ASAE*, 31:1624–1632, 1988.

[26] J.M. Sanchiz, F. Pla, J.A. Marchant, and R. Brivot. Structure from motion techniques applied to crop field mapping. *Image and Vision Computing*, 14:353–363, 1996.

[27] S.A. Shearer and R.G. Holmes. Plant identification using color co-occurrence matrices. *Transactions of the ASAE*, 33(6):2037–2044, Nov 1990.

[28] C.A. Shull. A spectrophotometric study of plant reflection of light from leaf surfaces. *BOT. Gaz.*, (87):583–607, 1929.

[29] D.C. Slaughter, R.C. Curley, P Chen, and D.K. Giles. Robotic cultivator. *U.S. Patent No. 5,442,552. Washington, D.C.: United States Patent & Trademark Office*, 1995.

[30] B. Southall, J.A. Marchant, T. Hague, and B.F. Buxton. Model based tracking for navigation and segmentation. In *5th European Conference on Computer Vision Freiburg, Germany*, volume 1. European Conference on Computer Vision, 1998.

[31] J.V. Stafford and J.V. Benlloch. Machine-assisted detection of weeds and weed patches. In J.V. Stafford, editor, *Precision Agriculture '97*, volume 2, pages 511–518, 1997.

[32] J.F Thompson, J.V. Stafford, and B Ambler. Weed detection in cereal crops. *ASAE Paper No. 90-7516*, 1990.

[33] J.F. Thompson, J.V. Stafford, and P.C.H. Miller. Potential for automatic weed detection and selective herbicide applicatio. *Crop Protection*, 10:255–259, 1991.

[34] L. Tian. *Knowledge Based Machine Vision System for Outdoor Plant Identification.* PhD thesis, University of California, Davis, Dept. of Biological and Agricultural Engineering, 1995.

[35] N.D. Tillett. Automated guidance sensors for agricultural field machines: a review. *Journal of agricultural engineering research*, 50(4):167–187, Nov 1991.

[36] R.D. Tillett. Image analysis for agricultural processes: a review of potential opportunities. *Journal of agricultural engineering research*, 50(4):247–258, Nov 1991.

[37] E. Vrindts and J de Baerdemaerker. Optical discrimination of crop, weed and soil for on-line weed detection. In J.V. Stafford, editor, *Precision Agriculture '97*, volume 2, pages 537–544, 1997.

[38] D.M. Woebbecke, G.E. Meyer, K Von Bargen, and D.A. Mortensen. Shape features for identifying young weeds using image analysis. *Transactions of the ASAE*, 38(1):271–281, 1995a.

[39] D.M. Woebbecke, G.E. Meyer, K. von Bergen, and D.A. Mortensen. Color indices for weed identification under various soil, residue and lighting conditions. *Transactions of the ASAE*, 38(1), 1995b.

[40] J.T. Wooley. Reflectance and transmittance of light by leaves. *Plant Physiology*, 47:656–662, 1971.

Part II

Physics Based Analysis of Outdoor Images of Vegetation

Chapter 4

On Physical Colour Image Understanding and Illumination Changes

4.1 Physical Colour Image Understanding

As introduced weed control is an outdoor operation taking place in a highly changeable environment. To adapt to the changing condition we approach the image analysis task by the philosophy of physic based image analysis. In the beginning of computer vision research analysis of images was generally divided into to phases a low segmentation or feature extraction phase and a higher level reasoning phase in which image features are related to object features described in object models of the scene. More recently research in computer vision has started to pay attention to the physical process behind the images and tried to incorporate this knowledge into image analysis methods. Series of intrinsic models has been introduced each describing how one or more particular process will influence on the captured images. As processes geometric aspects of objects shape from shadows, photometric aspects of objects shape for shading, motion analysis, optical flow etc., have been investigated.

The advantages of using intrinsic models prior to the traditional approach includes several aspects as introduced in Klinker [15]. First, having knowledge and appropriate models of the process which one wants to capture enables the researcher to design the measurements space necessary to capture the set of adequately characteristics. Second, intrinsic models can be extended to evolve more and more physical processes. Third, models of sensor characteristics (i.e. sensor uncertainty) may be included together with sensors influence on the measurement space. Finally, physical modelling of the processes in the scene provides the means to evaluate its strengths and limitations and makes it possible to predict which scenes that can be analysis by it.

This chapter is outlined as a tutorial into physical colour image understanding and illumination changes. It gives a more detailed and thorough presentation of the models and theory behind the methods used in the following chapters of this part.

4.1.1 Image Formation and Representation

Formation of images are in general described by the process of illumination sources emitting light striking objects' which reflect the light towards an observer, i.e. the process involves the three parts; light sources, objects, and observers. Figure 4.1 illustrates the three components and the travel of the light from light

sources to the observer.

In addition to these three "main" parts also the medium which the light travel through and the background surrounding the objects will influence the image formation process. In this study we will concentrate on images captured in open air but for other computer vision purposes as for example images captured underwater the medium has be taken into account for an appropriate description of the image formation process.

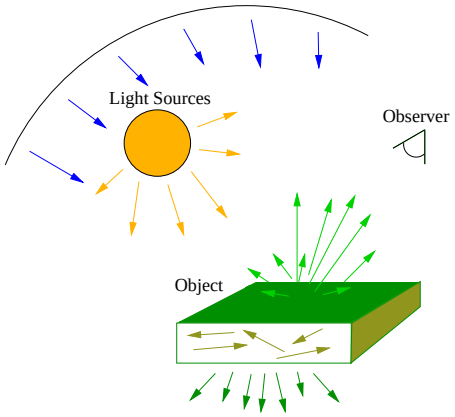


Figure 4.1: *Illustration of the image formation process*

Illumination

Several illumination sources may illuminate the objects of interest giving very complex images to analysis, as illustrated in figure 4.1. Not only may the sources be different in intensity but also in spectral composition.

Light consist of photons travelling round in the world. A photon is quantum of light energy moving in a single direction unless something like matter or a strong gravity field affects its motion. Thanks to the sun and other light sources there are many photons travelling around. Collection of photons travelling in the same direction at the same place and time constitutes a light ray. As photons move they oscillate about their direction of travel making a spectrum of wavelength λ specifying their distance of travel at a single oscillation.

Light may be characterized by its spectrum and polarization state. The spectrum specify its radiant power at a given wavelength λ. Concentrating on light in the human visible region from about 380nm (deep blue) to 760nm (deep red) a light spectrum with its radiant power concentrated from about 600nm to 760nm will we as humans perceive as reddish light whereas light with its power concentrated from 400nm to 500nm will appear bluish.

The polarization state specifies the photons oscillation and orientation in the direction of travel. Polarization is normally quantified with Stokes vectors. Stokes parameters are four quantities of and electro magnetic wave specifying the amount of photons travel in the same orientation and oscillation.

Objects

When light strikes matter three "transfer" processes may occur, reflection, absorption, and transmission. Denote the incident light spectrum at location (x, y) by $E(\lambda)$ then:

$$E(\lambda) = L_{reflection}(\lambda) + L_{absorption}(\lambda) + L_{transmission}(\lambda) \tag{4.1}$$

Dividing equation 4.1 by $E(\lambda)$ gives the relative reflections of the incident light denoted by ρ as:

$$1 = \rho_{reflection}(\lambda) + \rho_{absorption}(\lambda) + \rho_{transmission}(\lambda) \tag{4.2}$$

where $(\rho_{reflection}(\lambda))$ and $(\rho_{transmission}(\lambda))$ is the reflectance and transmittance of the object, respectively.

Reflectance
Within computer vision focus naturally are on the reflected light of objects. Further, within gray scale images is has been limited to the analysis of the intensity image of the objects in the scene without further consideration to the image formation process. However, within colour image analysis and especially physics based image understanding, the process happening in the material is of great interest and have been used for development of reflection models [27, 34]. The community has adopted reflection model from other fields. Among the eldest is Lamberts law [13] used for description of body (diffuse) reflection and utilized within vision methods, as shape from shading. For description of surface (specular) reflection from optical rough surfaces the model of Torrance & Sparrow [32] has been used. Later, a more thorough description of some of the proposed reflectance models will be presented.

Further, reflectance characteristics of objects are of major interest for sensor design where they may be used to specify the sensitivity characteristic of the sensor for optimal detection of specific objects. Traditionally, reflectance characteristics are regarded as being constant over time. However, for biological objects this may not be the case. Think of skin colour at time t_1 a person may appear white then after a hard run he or she at time t_2 may appear more reddish due to increased blood circulation. Hence, for biological it is more reasonable to specify approximate reflectance characteristic which may be used as guidance for design of sensors and image analysis methods.

Absorption
Some of the light striking an object may be absorbed and possibly used by an internal process as photosynthesis in vegetation. As the sum of the relative reflection, absorption, and transmission coefficients add up to one, absorption characteristics are also of interest for development of sensors as an object's absorption in a given wave band will influence the transmittance and reflectance in the same wave band.

Transmittance
Many, especially thin objects, will transmit part of the light striking it. Transmittance characteristics may as absorption be useful for design of sensors but also for modelling of second order reflection due to light first passing through the object. As absorption is a function of λ, the transmitted light may have a spectral composition very different from the incident light.

Observer

The reflected light from an object at location (x, y) towards an observer will be equal to the incident light $E(\lambda)$ weighted by its reflectance $\rho(\lambda)$ as:

$$L(\lambda) = E(\lambda)\rho(\lambda) \tag{4.3}$$

Equation 4.3 is defined over a continuous spectrum which in a sensing device such as a camera is reduced by the linear transformation spectral integration [26]:

$$\mathbf{C}_f = \int_\lambda L(\lambda)\tau_f(\lambda)d\lambda \;=\; \int_\lambda E(\lambda)\rho(\lambda)\tau_f(\lambda)d\lambda \tag{4.4}$$

Depending on the filter characteristics $f = 1, 2 \ldots n$ of the observer the spectrum received is now reduced to n dimensions. In a standard colour camera the dimension is traditionally three ($f = 3$) corresponding to a response in the red, green, and blue, giving the so called RGB triplet. The colour vector $\mathbf{C}$ as specified

in equation 4.4 is the sum of the incident light $E(\lambda)$ weighted for every wavelength by the reflectance $\rho(\lambda)$ of the objects and spectral sensitivity of the camera $\tau(\lambda)$. A process that has the unfortunately ambiguity *metamerism* [36] i.e. that two objects having different reflectance characteristics viewed under one light source may give the same RGB values whereas under a other light source they may give different RGB values.

Chromaticity

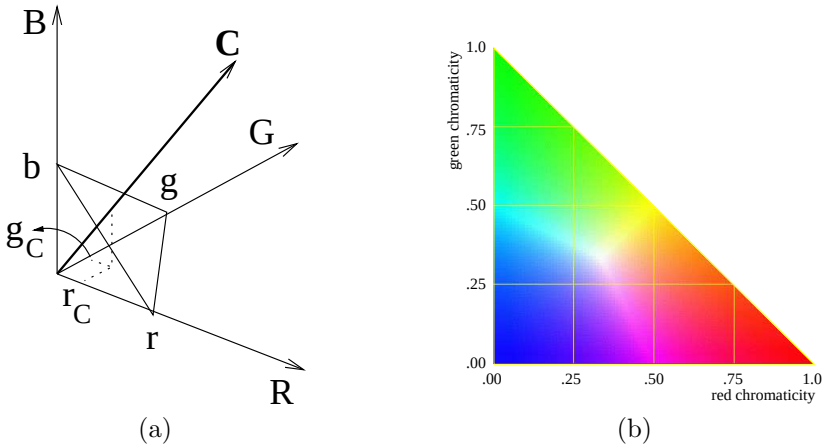
Concentrating on a RGB camera then equation 4.4 denotes a *colour vector,* **C** in the RGB colour cube. The magnitude of the vector will be depending on the intensity of the illumination and settings of the camera (aperture, exposure time).

Many methods for representation of a colour vector has been proposed. Basically they may be divided into two categories, a) transforms that intent to adapt to the visual perception of humans (CIE $L^*u^*v^*$ and CIE $L^*a^*b^*$), and b) transforms suitable for visualization purposes and computer vision methods (HSI, YIQ, rgb, I1I2I3 etc.). A thorough discussion of several of the proposed methods are presented in [21].

In this presentation only the convenient and traditional transformation of the RGB colour vector into its intensity independent chromaticities, will be introduced. This is done by dividing the colour vectors with its first norm:

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \quad \text{and} \quad r+g+b = 1 \tag{4.5}$$

As the sum of the chomaticities is equal to one it is sufficient to represent a colour vector by two of them. Figure 4.2(a) illustrates the intersection of a colour vector with the chromaticity triangle and then further its projection onto the rg-chromaticity plane.



(a) (b)

Figure 4.2: *Projection of a colour vector* **C** *onto the rg-chromaticity plane. (a), the RGB-space with the location of colour vector* **C***, its intersection with chromaticity triangle and its projection onto the rg-plane. (b), colour distribution on the rg-chromaticity plane.*

In figure 4.2(b), the distribution of colours onto the rg-chromaticity plane is illustrated. Notice that the gray level line is projected onto the point $(r, g) = (1/3, 1/3)$. The chromaticity is not defined for $R+G+B = 0$, however this has no influence in practical situations as pixels with very low intensity are often not considered due to the "poor" signal to noise ratio.

4.2 Dichromatic Reflection Model

In recent years the dichromatic reflection model introduced by Shafer (1985) [27] has almost become synonym with physical colour image understanding. For analysis of colour images it have been used for image segmentation [16, 25, 20], analysis of highlights [18, 24], estimating scene properties [22] and colour of illumination [31, 28], and for generation of computer graphics rendering [30].

As the title of the model states the reflected light from dielectrical objects (i.e. non metal) may be split into two (di) distinct reflection components from the **body** and the **surface**, respectively. Figure 4.3 (a) illustrates the reflection components from a dieletrical object.

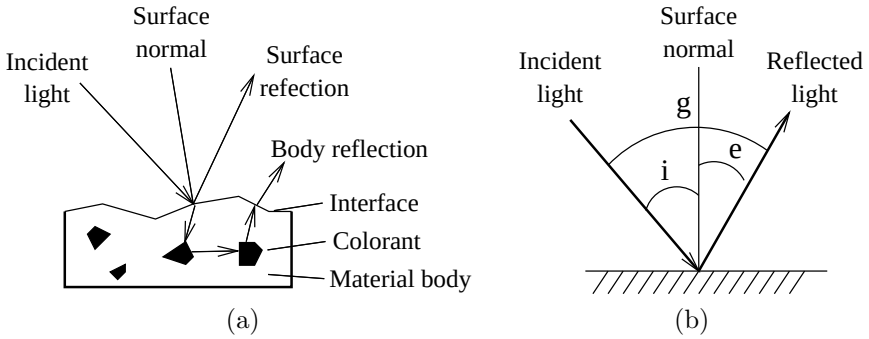


Figure 4.3: *Illustration of reflection components from a dielectrical object. (a), reflection and material components. (b), photometric angles*

Body reflection is formed by light penetrating into the material body where it traverses around hitting pigments, fibers and other materials. Some percentage of the various wavelengths will be absorbed by the material, others may be transmitted and finally some will be refracted at the material surface and reflected as body reflection. A reflection also refereed to as *diffuse* reflection. It is the reflection that primary will the colour of the object i.e. green, red, brown etc.

When light hits the surface of a dielectrical object it must first pass through the interface between the surrounding media and the object. As the refractive index of the surrounding media and the object normally is different some percentage of the incident light is reflected at the surface of the object. This reflection components is refereed to as surface reflection. Important it is to notice that surface reflection is similar to, but not identical to specular reflection. Specular reflection is related to the photometric angles which give maximum surface reflection. The colour of the reflected light is depending on Fresnel's law described by the refractive indices of the object and surrounding medium a function over wavelength and the polarization of the light. As the refractive index for many daily life objects is approximately constant over the visible spectrum the colour of the surface reflection will be similar to that of the incident light.

Taking the photo metrical angles, i *incident angle*, e *exit angle*, and g *phase angle* between the camera and observer into account the emitted light received by an observer (e.g. camera, human etc.) may be formulated as:

$$L(\lambda, i, e, g) = L_B(\lambda, i, e, g) + L_S(\lambda, i, e, g) \tag{4.6}$$

where $L(\lambda, i, e, g)$ is received light by the observer, $L_B(\lambda, i, e, g)$ is light emitted as body reflection, and $L_S(\lambda, i, e, g)$ is light emitted as surface reflection, respectively. The parameter λ represents wavelengths of the light spectrum. For simplicity the set of angles i, e, and g will in the following be represented by Θ.

As the spectrum of surface and body reflection from dielectrical objects are approximately unaffected by illumination geometry [27, 33], $L_B(\lambda, \Theta)$ and $L_S(\lambda, \Theta)$ may be separated into a geometrical scaling factor

m and a light spectrum c:

$$L(\lambda, \Theta) = m_b(\Theta)c_B(\lambda) + m_S(\Theta)c_S(\lambda) \tag{4.7}$$

Equation 4.7 describes the reflected light from a point on the object as an additive mixture of the two light spectra $c_B(\lambda)$ and $c_S(\lambda)$, scaled by the geometrical scaling factors $m_B(\Theta)$ and $m_S(\Theta)$, respectively. Later Tominaga [29] has categorized the model into three types valid for different materials.

Type I, *The Standard Dichromatic Reflection Model*

This model is based upon the *Neutral Interface Reflection* [19] assumption which assumes that the relative spectral composition of the surface reflection component is identical to that of the light source (i.e. the surface reflectance $\rho_S(\lambda)$ is constant). Substituting the two light spectra in equation 4.7 by their surface reflectances and the incident light as described by equation 4.3 gives:

$$L(\lambda, \Theta) = m_b(\Theta)\rho_B(\lambda)E(\lambda) + m_S(\Theta)\rho_S E(\lambda) \tag{4.8}$$

With ρ_S as a constant the last term gives the spectrum of the light source. Hence, it is possible to retrieve its spectrum a feature used in several colour constancy methods. We will return to this topic later.

The model has been found experimentally valid for materials having high contents of water or oil as these have a index of refraction approximately constant over the visible spectrum. Covered by the model are materials as plastics, paints, ceramics, vinyls, tiles, fruits, leaves, and woods [29, 25, 19].

Type II and III *Extended Dichromatic Reflection Models.*

Type II, unfortunately not all dielectrical materials follow the neutral interface assumption. Especially cloth including materials as silk, wool, polyester satin, cotton, velveteen but also coloured paper do not follow its assumption[29, 19]. Rather the reflectance of the surface $\rho(\lambda)$ has to be taken into account for these materials when modelling their reflection.

Type III, is an extension of the model also to include metals. Metals have quit different reflection properties than inhomogeneous materials as they have only surface reflection. Tominaga shows that reflection from metals in the specular direction tends to be whitened as the incident angles approaches the grazing angle 90°. Hence the model for metals may be formulated as an additive mixture of reflection at the specular angle (the major reflection component) plus reflection at the grazing angle. This gives a model that mathematically corresponds to Type I.

The dichromatic reflection model has now been introduced. In the next section its representation in the RGB colour cube and the rg-chromaticity plane will be presented.

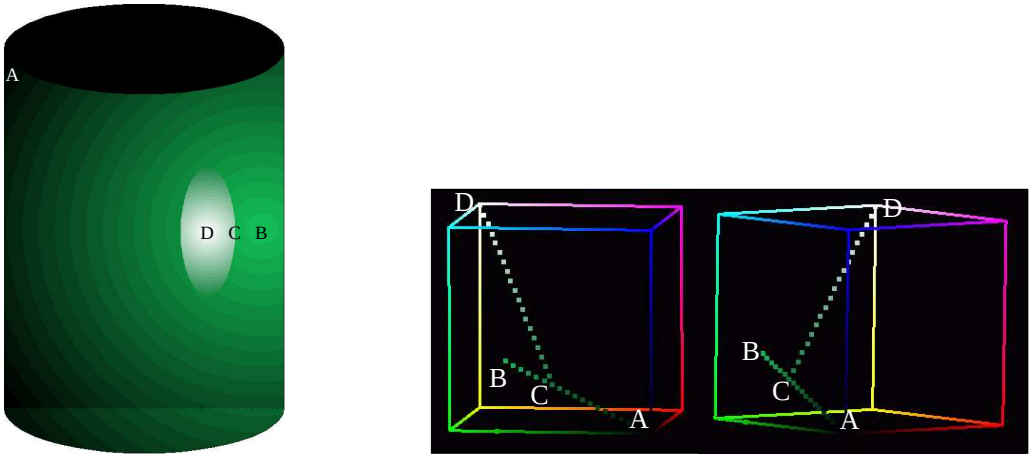
4.2.1 Camera Model

Transformation from the continuous light spectrum to the RGB-cube is done by spectral integration. In a camera with tristimuli characteristic specified by $\tau_f(\lambda)$ $(f = 3)$, the representation of the dichromatic reflection model becomes:

$$\begin{aligned} \mathbf{C}(x,y) &= m_B(\Theta) \int E(\lambda)\rho_B(\lambda)\tau_f(\lambda)d(\lambda) \\ &+ m_S(\Theta) \int E(\lambda)\rho_S(\lambda)\tau_f(\lambda)d(\lambda) \Rightarrow \\ \mathbf{C}(x,y) &= m_B(\Theta)\mathbf{C}_B(x,y) + m_S(\Theta)\mathbf{C}_S(x,y) \end{aligned} \tag{4.9}$$

where $\mathbf{C}(x, y)$ is a three dimensional colour vector for the pixel at location (x, y), from an additive mixture of $\mathbf{C}_B(x, y)$ and $\mathbf{C}_S(x, y)$ colour vectors of the body and surface reflection, respectively.

Equation 4.9 defines a plane spanned by the two vectors $\mathbf{C}_B$ and $\mathbf{C}_S$ the so called *dichromatic plane*, in which all pixel points from an object will be distributed. How they will be distributed on the plane



(a) (b)

Figure 4.4: *Reflection from a synthesized green dielectrical cylinder illuminated by a point light source. (a), green cylinder with highlight. (b), distribution of pixel points in the RGB cube showing body and surface reflection vectors. Included are points A, B, C, and D as illustrated in (a).*

will be determined by the geometry of the object. For the important category of convex objects they will cluster around the body and surface vector. The body vector originates in $(0, 0, 0)$ and it may be shown that the surface reflection vector will begin in the upper 50% of the body vector, the so called *50% heuristic*, as introduced by Klinker [16].

Figure 4.4(a) illustrates a synthesized green convex dielectrical cylinder illuminated by a point light source following the neutral interface assumption. The body vector will be at minimum intensity at point A and increase to point B where it reaches its maximum. At point C on the edge of the highlight region the surface reflection vector will begin and increase sharply in intensity until point D where it will be at maximum intensity and corresponding to the colour of the illumination. Notice that the maximum of the surface and body reflections are not at the same location on the object. Modelling the body reflection by Lambertian reflection and simplifying the specularity to appear when the surface normal bisects the angle between the light source and camera then the two different reflections will have maximum as illustrated in figure 4.5.

From eq 4.9 and as illustrated in figure 4.5 it is evident that it is necessary to have adequate models of the geometric scalings factors m_B and m_S to be able to model the pixel points distribution in the colour cube.

m_S, Geometric Scaling Factor for Surface Reflection

For didactic and historical reasons m_S, the geometrical scaling for surface reflection, will be introduced first. The presentation will be concentrated about the model proposed by Torrance and Sparrow in 1967 [32]. The "key" idea in this is to model the material surface as a collection of symmetric V groove cavities which facets normals are assumed Gaussian distributed with the variance σ describing the surface optical roughness. The facets' size is assumed larger than the wavelength of visible light, but too small to be seen as texture. Smooth surfaces will have a small value of σ whereas rougher surfaces will have a larger value. The concept of the V groove cavities and how they affects the reflected light is illustrated in figure 4.6. (b), illustrates the situation where some of the reflected is masked due to the V cavity whereas in (c) both the effect of masking and shadowing are included. From the figure it is evident that the degree of masking and shadowing will be depending on the photometric angles and the optical roughness.

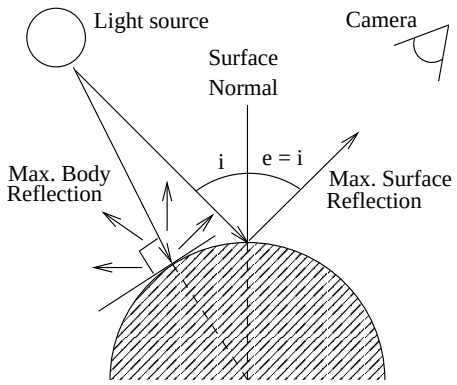


Figure 4.5: *Illustration of maximum body and surface reflection from a convex surface modelling $m_B(i, e, g)$ as Lambertian reflection and assuming maximum specular reflection when the surface normal bisects the angle between the camera and light source.*

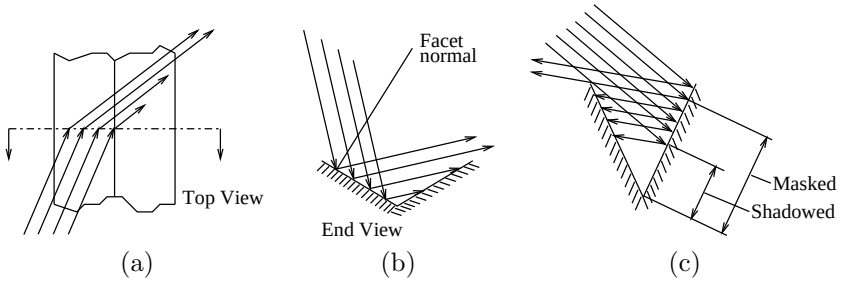


Figure 4.6: *Geometry of the V groove cavity. (a) top view of a cavity with direction of incident and reflected light. (b) cross section view of (a) with the reflected light beam masked. (c), simultaneous masking-shadowing. (Reproduced after [32])*

$$m_S(\Theta) = I_E \kappa \frac{FG(i, e, g)C}{cos(e)\sigma} e^{-\frac{\alpha^2}{\sigma^2}} \tag{4.10}$$

Mathematically the model may be expressed as in equation 4.10 where I_E is the intensity of the illumination, κ the gain of the camera. i, e, and g are the photometric angles and α is the off-specular angle which may be expressed by i, e, and g. C is a variable including the facet size. F is Fresnel coefficient which may be assumed to be about 1.5. G, is an attenuation factor depending upon the photometric angles and the V groove geometry as illustrated in figure 4.6. It is a complicated function divided into intervals specified by the photometric angles and will not be reproduced here.

One of the main features of the model is that it is capable of correctly specifying the effect that the maximum specular reflection for most common materials occur at an angle larger than the "perfect" specular angle, i.e. the angle specified by that the surface normal bisects i and e, as illustrated in figure 4.5.

$\mathbf{m_B}$, Geometrical Scaling Factor for Body Reflection

The first and traditionally used model for m_B is based upon that material surfaces appears equally bright from all viewing directions. This model was advanced by Lambert over 200 years ago and is now known

as Lambert's Law or model of reflection. Formally the model may be specified as:

$$m_B(i) = I_E \kappa A cos(i) \tag{4.11}$$

where I_E is the intensity of the illumination, κ the gain of the camera, and A the albedo of the object ranging from $0 \ldots 1$. As may be noted from eq 4.11 the magnitude of the reflection for a given illumination intensity and gain of the camera will only be depending of the incident angle i. Figure 4.7 (a) illustrates the principle of Lambert's reflectance i.e. the reflectance is constant for all viewing angles. Clearly, the model is simplicity itself and hence also used in many computer vision algorithms. However, a prerequisite for it to operate satisfactorily (within a error margin of $\pm$ 5%) is that i and e are both less than 50°, as demonstrated in [34].

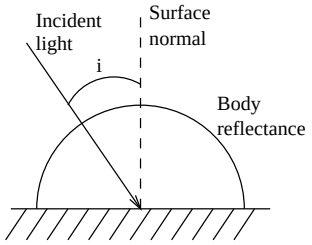


Figure 4.7: *Lambertian reflectance notice that the reflectance is independent of viewing angle.*

To account for the defectiveness of Lambert's model recently two newer models for m_B adequate for rough and smooth surfaces have been proposed by Oren and Nayar [23] and Wolff [35], respectively.

Model for rough surfaces

The model by Oren and Nayar is inspired by the V groove cavity representation of the surface roughness as in Torrrance and Sparrows' but assuming each facet surface to be Lambertian in reflectance. To illustrate the philosophy of the model consider figure 4.8 including a V groove cavity illuminated by a distant light source on the right side, figure 4.8 (a). Assuming the facet to be Lambertian gives that the left side will appear brighter than the right as it receives more light. In (b) and (c) the cavity is viewed by a distant observer from the left and right side, respectively. As the viewer moves from left to right the foreshortened cavity area of the bright area will increase and the dark decrease. Consequently, the total brightness, or radiance, of the cavity increases as the observer moves towards the direction of the light source.

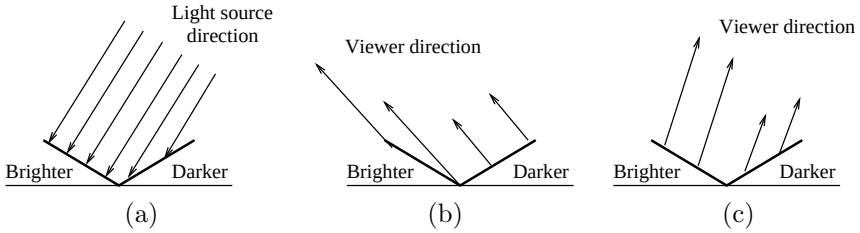


Figure 4.8: *Principle of Oren and Nayar model for diffuse reflection. The radiance of the V-cavity increases as the viewer moves towards the illumination direction. (Reproduced after [23])*

Note that the brightness of the surface will be dependent of the viewing angle and hence in contrast to Lambert's model which is independent of this. The above illustrates that rough Lambertian surfaces are inherently non Lambertian in reflectance.

Oren and Nayar derive three models for different distribution of the facets and V cavities taking masking, shadowing and first order inter reflection into account, i.e. light which bounce from one side of the cavity to the other side. As distribution are uni-directional single-slope with all facets having the same slope and the V cavities aligned in the same direction, isotropic single-slope with all facets having the same slope but their orientation uniformly distributed in the plane of the surface, and Gaussian slope-area where the facets slope-area is modelled by a Gaussian distribution with mean μ (typically set to zero) and variance σ and their orientation uniformly distributed in the plane of the surface, proposed.

The mathematical representation of the models are quit comprehensive and will not be presented in this context. Experiments shows that the model is able to quantify reflection from rough surfaces as wall plaster, sand, sand paper, foam, cloth, and wood shaving very accurately. Further implications of the model for machine vision algorithms as rendering of images, reflectances maps, photometric stereo, binocular stereo, and motion estimation are discussed.

An interesting feature using the model for rendering of a sphere is that with increasing roughness the sphere becomes more and more disk like. A phenomena well known from appearance of the moon which at full appears more as a disk than a ball which it should have if it strictly followed Lambert's law.

Model for smooth surfaces

The model proposed for smooth surfaces by Wolff is in contrast to Oren and Nayar not making any assumption about the surface roughness but instead modelling scattering of the incident light in the subsurface.

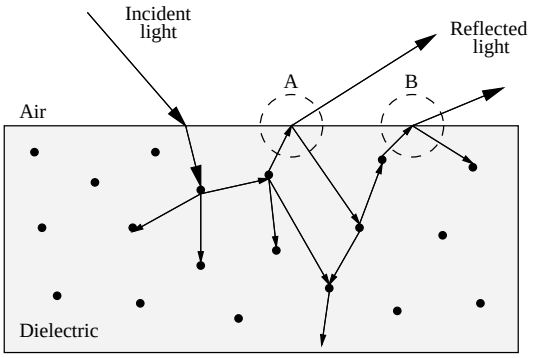


Figure 4.9: *Geometric refraction of incident light and subsurface scattered light within a dielectrical object with a smooth surface. At A, first reflection from dielectric into air and specular reflection back into the dielectrical. At B, second reflection from dielectrical into air and specular reflection back into the dielectrical.*

The model is based upon Chandrasekhar [4] theory of radiative transfer of multiple scattering of incident light upon stellar and planetary atmospheres. The problem of diffuse reflection and transmission from plane-parallel atmospheres in astrophysics has a number of similarities with the problem of diffuse reflection from inhomogeneous materials. Consider figure 4.9 including a dielectrical material consistent of a medium with a uniform index of refraction different from that of air with particles inhomogeneity embedded. Looking at the figure the material almost looks like an atmosphere with planets, which inspired Wolff to adapt Chandrasekhar theory.

The total reflection obtained by the model will due to: first reflection from light that penetrates into the material surface, scatters among subsurface particle inhomogeneities, and is then refracted back out into air. Notice that Fresnel coefficients is appearing both for light penetrating into the material (i.e from air to material) and for light exiting the material (i.e. from material to air). Second, at the material boundary some light is specular reflected back into the material where it again is scattered among subsurface particle inhomogeneities and then refracted into air where some light is reflected into

the material again, as illustrated at location A and B in figure 4.9. The process will theoretically continue to infinity and finally give the total reflection and the total diffuse albedo of the material.

Wolff experimentally used the model for modelling reflection of a smooth ceramic cup and a billiard ball. The latter was used for comparing rendering of the ball using the proposed model and Lambert's law. For both experiments it showed much better capabilities of quantifying the reflected light especially at grazing angles.

Pixel Points Distribution of Real Dielectrical Objects

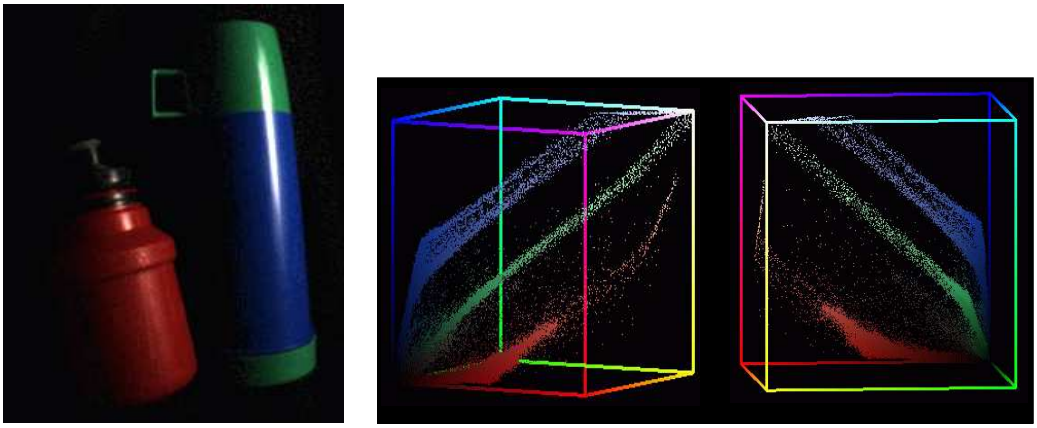


Figure 4.10: *Images of dielectrical objects. Left two plastic objects illuminated by a point light source. Middle and right, RGB-histograms of image.*

The distribution of the pixel points as illustrated in figure 4.4 for the synthesized cylinder is clearly very idealized. For real objects as illustrated in figure 4.10 the pixel points will be distributed around the body and surface reflection vectors. How much the pixel points will be dispersed will be depending of the quality of the camera and uniformity of the objects' colour. For all three colours red, green, and blue of the "ideal" objects the plastic bottle and vacuum jug in the figure it is evident that their pixel points are distributed according to the dichromatic reflection model.

Starting at the 'black' corner of the RGB-cube the three objects' pixel points are all distributed along a body vector having direction towards either the red, green, and blue corner of the cube, respectively. Next, somewhere in the upper 50% of the body vector the surface vector begins for all three objects with a direction towards the white corner of the cube. The distribution of the pixel points are forming what in the literature has been denoted as a "dogleg", "skew T", or "banana like cluster" distribution, but common for all three are that they are distributed in a plane spanned by the two vectorial directions. What also is evident for the objects is that they have most of their pixel points located along the body vector, of course the distribution along the two vectors will be depending on the photometric angles but in most "daily life" situations this concentration of the distribution may be expected, i.e. surface reflection is normally concentrated on very specific areas on objects where the photometric angles are specular or near specular.

Chromaticity Representation

As introduced a convenient and traditionally used transformation of the colour vector $\mathbf{C}$ is to transform it into its intensity independent chromaticities, by dividing it with its first norm. For the dichromatic reflection model this transformation becomes:

$$\mathbf{c}(x,y) = |\mathbf{C}(x,y)| = |m_B(\Theta)\mathbf{C}_B + m_S(\Theta)\mathbf{C}_S| \tag{4.12}$$

which may be expressed as:

$$\mathbf{c}(x,y) = \frac{1}{S_B + \frac{m_S(\Theta)}{m_B(\Theta)}S_S}\mathbf{C}_B + \frac{1}{S_S + \frac{m_B(\Theta)}{m_S(\Theta)}S_B}\mathbf{C}_S \tag{4.13}$$

where S_B and S_S are the sum of $\mathbf{C}_B$ and $\mathbf{C}_S$, respectively. $\mathbf{c}(x,y)$, is a three by one vector holding the chromaticities of pixel point (x,y). As the sum of $\mathbf{c}$ is equal to one:

$$\begin{aligned} sum(\mathbf{c}(x,y)) &= \frac{m_B(\Theta)S_B}{m_B(\Theta)S_B+m_S(\Theta)S_S} + \frac{m_S(\Theta)S_S}{m_B(\Theta)S_B+m_S(\Theta)S_S} \\ sum(\mathbf{c}(x,y)) &= \frac{m_B(\Theta)S_B+m_S(\Theta)S_S}{m_B(\Theta)S_B+m_S(\Theta)S_S} = 1 \end{aligned}$$

it sufficient to represent the chromaticities in two dimensions.

Looking closer at equation 4.13, if $m_S(\Theta) \approx 0$ (points A, B, and C in figure 4.4) the last term in equation 4.13 will approach zero and the denominator in the first term will simplify to S_B and hence $\mathbf{c}(x,y) = \frac{C_B}{S_B}$, corresponding to pure body reflection. On the other hand if $m_S(\Theta) \gg m_B(\Theta)$ (point D in figure 4.4) then the first term will approach zero and the denominator in the second term will simplify to S_S and hence $\mathbf{c}(x,y) = \frac{C_S}{S_S}$, corresponding to pure surface. These two "extreme" chromaticities for respectively pure body and surface reflection will be denoted as the *primary chromaticity* of the object and along the line segment connecting the two will the chromaticities of it be distributed.

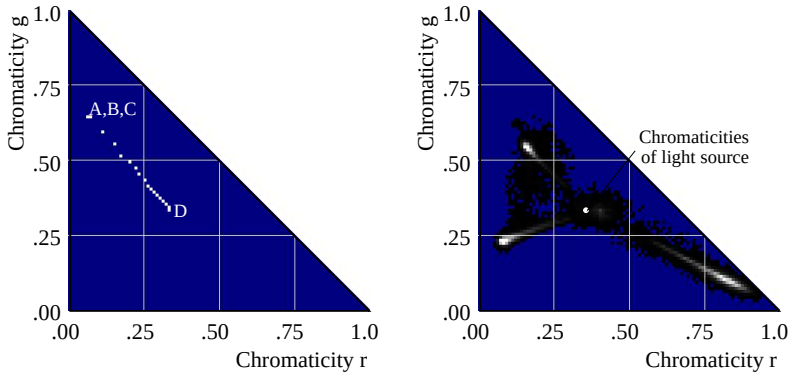


Figure 4.11: *rg-chromaticity plane. (a) Chromaticities of synthesis cylinder together with location of points A, B, C, and D in figure 4.4. (b) rg-chromaticity plane of plastic objects in figure 4.10. Included is location of the light source chromaticities.*

Figure 4.11 (a) illustrates the chromaticity distribution of the green synthesized cylinder with location of points A, B, C, and D included. From the figure it is evident that the chromaticities are distributed along a line segment connecting points A, B, and C with point D.

In figure 4.11(b) are the chromaticities of objects included in figure 4.10 together with location of the light source, illustrated. From it the linear distribution of the pixel points along the line segment connecting the primary body and surface chromaticity is clearly evident. Also, note that the three clusters beginning at the red, green and blue corner, respectively, all traverse towards the chromaticity of the light source as expected for objects with neutral interface reflection.

4.2.2 Spectral Variation of the Illumination

So far, the representation of the dichromatic reflection model has been constrained to include only one spectral composition of the illumination source $E(\lambda)$. However, images acquired in an unconstrained environment will be subjected to highly changeable illumination condition and hence the impact on the image formation must be investigated. The problem occurring is that if the spectral composition of the illumination changes then the colour of the objects in the scene will change as well. We as human beings have a mechanism called *colour constancy* that makes us able to adapt to spectral variation of illumination and perceive colour almost unaffected by it. Much harder it seems for computer vision system to adapt to such changes. Generally speaking the computer vision system will not detect that the colour of the object has changed due to illumination changes but rather recognize it as a new object in the scene.

Adaptation to spectral changes of the illumination may potentially be handled in two ways either by finding a representation of the colours responses invariant towards spectrally changes of illumination or by implementing a colour constancy method which transforms the colours in the image to a canonical illumination. The first approach is applicable especially when there are relatively few objects in the scene whereas if feasible the second clearly is the ultimately solution as it in general will make acquisition of images independent of illumination changes.

Modelling Illumination Changes

As this thesis will focus on the use of computer vision outdoor the presentation of illumination changes will be constrained to involve only daylight. Fortunately, Judd *et al* [14] in 1964 collected 622 spectral distributions of daylight (skylight, and sunshine plus skylight). For description of the various spectra these where related to the radiance from a complete radiator (i.e. a black body) by their *correlated colour temperature* CCT. The spectral radiance from a black body as function of absolute temperature K (degrees Kelvin), is described by Planck's formula:

$$\begin{aligned} & M(\lambda) = k_1 \lambda^{-5} (e^{\frac{k_2}{T\lambda}} - 1)^{-1} \\ \text{with} \quad & \\ & k_1 = 2\pi hc^2 = 3.7418 \times 10^{-16}\ Wm^2 \\ & k_2 = \frac{hc}{k} = 1.4388 \times 10^{-2}\ WK \end{aligned} \tag{4.14}$$

where c is velocity of light, h is Planck's constant, and k is Boltzmann constant. Relative power spectra of equation 4.14 for the temperature K denoted as colour temperature ranging from 4k to 25k degrees Kelvin, illustrated in figure 4.12(a). A high colour temperature corresponds to a bluish spectra i.e. the radiant power is concentrated at low wavelengths whereas a low colour temperature is corresponding to reddish light i.e. the radiant power is dominated by long wavelengths.

The 622 spectra where subjected to a principal component analysis and Judd *et. al* found that they could be approximated using the mean $\overline{E}(\lambda)$ and the first two principal components $V_{1,2}(\lambda)$ weighted by scalar multiples $M_{1,2}$ as:

$$E(\lambda) = \overline{E}(\lambda) + M_1 V_1(\lambda) + M_2 V_2(\lambda) \tag{4.15}$$

Judd *et. al* solved for the scalar multiples $M_{1,2}$ using any arbitrarily *xy* chromaticity coordinates derived by using the 1931 standard observer [36] as tristimuli function i.e. $\tau(\lambda)$ in equation 4.4. However, *xy* chromaticities of daylight spectra will not be randomly distributed indeed they clearly form a locus in the chromaticity plane. Hence, the *Commission Internationale De L'Éclairage* CIE [1] has derived a relationship between *xy* chromaticities and CCT by:

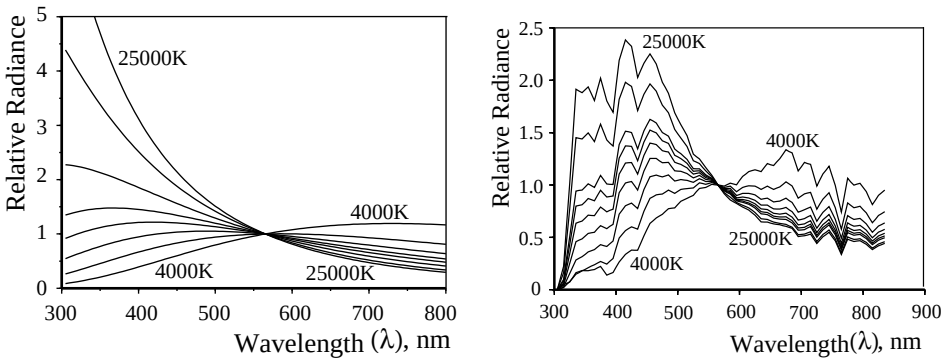


Figure 4.12: *Relative radiance normalized at 560nm. (a), black body according to equation 4.14 in steps showing 4k, 5k, 6k, 7k, 10k, 15k and 25k degrees Kelvin.. (b), phases of daylight with Correlated Colour Temperature varying from 4000K to 25000K, in steps showing 4k, 5k, 6k, 7k, 8k, 9k, 10k, 15k and 25k degrees Kelvin.*

$$
\begin{aligned}
x_D &= \begin{cases} -4.6070\frac{10^9}{T^3} + 2.9678\frac{10^6}{T^2} + 0.09911\frac{10^3}{T} + 0.244063 & \text{if } 4000<\text{T}<7000 \\ -2.0064\frac{10^9}{T^3} + 1.9018\frac{10^6}{T^2} + 0.24748\frac{10^3}{T} + 0.23704 & \text{if } 7000\leq\text{T}<25000 \end{cases} \\
y_D &= -3.000x_D^2 + 2.870x_D - 0.275 \\
M_1 &= \frac{-1.3515-1.7703x_D+5.9114y_D}{0.0241+0.2562x_D-0.7341y_D} \\
M_2 &= \frac{0.0300-31.4424x_D+30.0717y_D}{0.0241+0.2562x_D-0.7341y_D}
\end{aligned}
\tag{4.16}
$$

By the equations in 4.16 it is now possible to obtain the scalar multiples needed in equation 4.15. The mean value and principal component vectors are obtainable among others from CIE [1]. Important, it is to notice that the approximated daylight spectra is independent of the observer and hence they may be used in general for modelling of daylight spectra as a function of CCT. Figure 4.12(b) includes approximated daylight spectra from a CCT of 4000K to 25000K i.e. from reddish light at sunrise or sunset to a clear blue north sky. A "typical" CCT of daylight as a mixture between sky- and sunlight are the in range about 5700K.

Having a suitable model of the illumination it is now possible to model colour changes of an object with known reflectance and observer characteristics, respectively, $\rho(\lambda)$ and $\tau(\lambda)$ in equation 4.4. In figure 4.13 a reflectance characteristic of green vegetation (a) together with the tristimuli characteristic of JAI's M90 3CCD camera (b), are included.

To solve the modelling numerically the integration in equation 4.9 is replaced by summation so formation of the colour vector $\mathbf{C}$ at location (x, y) becomes:

$$
\begin{aligned}
\mathbf{C}_{f,CCT} &= m_B(\Theta)\sum_{\lambda=\lambda_a}^{\lambda_b} E(\lambda, CCT)\rho_B(\lambda)\tau_f(\lambda)\Delta(\lambda) \\
&+ m_S(\Theta)\sum_{\lambda=\lambda_a}^{\lambda_b} E(\lambda, CCT)\rho_S(\lambda)\tau_f(\lambda)\Delta(\lambda)
\end{aligned}
\tag{4.17}
$$

The wavelength range of the spectrum from λ_a to λ_b is divided into small intervals $\Delta\lambda$ of equal size with centered at wavelength λ. $\Delta\lambda$ may at its finest be divided into intervals of $\Delta\lambda = 1$nm but for practical purposes and with light sources having smooth spectra as daylight (i.e no fluorescent) a $\Delta\lambda = 10$nm will give a sufficient approximation to calculations made at 1 nm intervals [36].

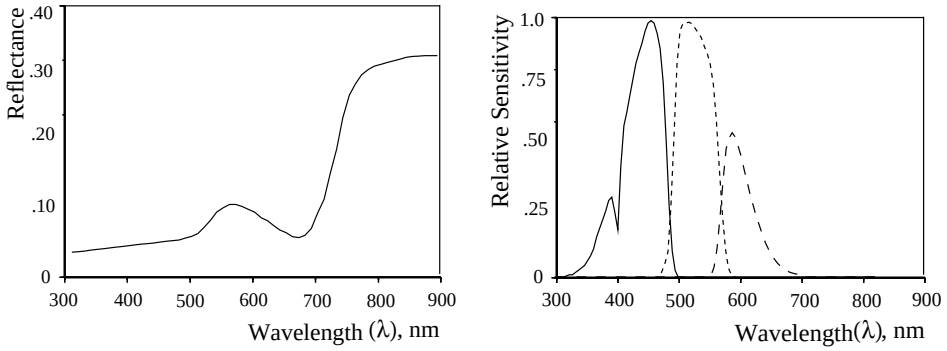


Figure 4.13: *The characteristics used for modelling. (a), reflectance characteristic of green vegetation. (b) tristimuli characteristic of JAI's M90 3CCD colour camera.*

Dividing the reflectance and camera characteristic into $\Delta\lambda = 10$nm a **Body** and **Surface** locus may be calculated for vegetation with daylight ranging in CCT from 4000K to 25000K. The body locus is obtained using the characteristics in figure 4.13. The surface locus is obtained assuming that vegetation follows the neutral interface assumption as found valid by [19, 29] in which case it will correspond to the daylight locus of the camera i.e. $\rho_S(\lambda)$ being constant.

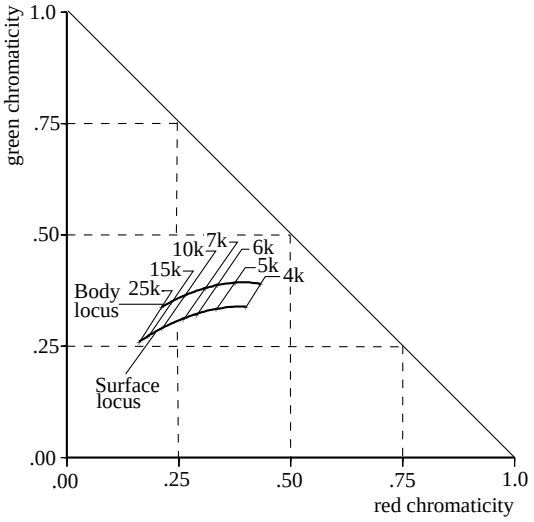


Figure 4.14: *Locus of body and surface reflection of vegetation with corresponding CCT at 4k, 5k, 6, 7k, 10k, 15, and 25k.*

Figure 4.14, illustrates how the two loci changes from reddish to bluish as the CCT of the daylight increases. Notice, how the CCT "squeezes" together as it increases along the two loci. This indicates that there is a hyperbolic relationship between CCT and the red chromaticity in correspondence with the relationship found for the standard observer in equation 4.16.

The figure also illustrates that in a "one object" situation it should be possible to solve for spectral changes of daylight by finding a direction in the chromaticity plane that approximate possible changes in the CCT.

4.3 Colour Constancy

The above section discussed how spectral changes of the illumination affects the location of the loci in the dichromatic plane. Also, introduced was that in a "one-object situation it should be possible to find a direction approximating possible changes. However, as previously introduced another approach to solve spectral changes of the illumination is to transform colours to a canonical illumination, so the colour of the object will be independent of the illumination; the so called colour constancy mechanism. The major advantage of a well functioning colour constancy method is that it is able to solve for varying illumination for scenes containing an arbitrary number of objects.

The problem of colour constancy may be formulated as finding a transformation $\mathbf{D}$ which transforms a colour vector $\mathbf{C}$ at location (x, y) to the colours of the location if it was illumination by a canonical illumination $\mathbf{C}^c$, as:

$$\mathbf{C}^c = \mathbf{C}\mathbf{D} \tag{4.18}$$

where $\mathbf{D}$, may be implemented as a full linear map. However, for narrow banded and orthogonal sensors (i.e. little overlap between sensitivities bands) the difference between a full linear map and a diagonal map will be almost non existing [7, 8, 2]. The prerequisite that the sensor has to be narrow banded or perfectly delta like to make the linear map work satisfactory is evident from:

$$\begin{array}{ll} \text{I:} & \frac{\int E(\lambda)\rho(\lambda)\tau_f(\lambda)d\lambda}{\int E(\lambda)\tau_f(\lambda)d\lambda} \\ \text{II:} & \frac{E(600)\rho(600)\tau_f(600)}{E(600)\tau_f(600)} \end{array} \tag{4.19}$$

Clearly, in I the $E(\lambda)$ does not cancel out whereas in II, the delta sensor $E(\lambda)$ cancel out such that $\rho(\lambda)$ may be isolated. Further, as $E(\lambda)$ in our project is regarded as continuous over (λ) the only way to make II feasible is to narrow the sensor's sensitivity bands, $\tau_f(\lambda)$. Fortunately, even if the sensors are broad banded or have a significant overlap it is usually possible to improve their independence by spectral sharpening of them [8]. Hence, the colour constancy problem simplifies to finding a diagonal map:

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} R \\ G \\ B \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \tag{4.20}$$

i.e. each channel is scaled independent, also known as von Kries transformation. The problem is now to find values of α, β, and γ.

Generally, colour constancy method may be divided into two categories which are statistically and physics based. The former tries to correlated statistics about colours in the image with statistical knowledge about surfaces and light sources. Many methods have been proposed and some tend to work well if there is a significantly diversity of the colours in the scene. The latter physics based category is founded on how physical processes as specularities and inter reflection manifest themselves in the image. The theory behind are more elegant, "versatile" and open the possibility to obtain colour constancy even when there are only two surfaces present in the image. However, they have so far been restricted to work only in laboratory conditions.

4.3.1 Statistically Based

In this section four methods will be introduced, two almost naive and very "simple" often used in comparative studies and two more recently introduced showing potential use even in realistic images.

The first naive method introduced is *grey world* mapping which simply scale each channel independently by the average of all RGB values in the image. The diagonal scaling factors are simple $R_{(x,y)}/R_{average}$, $G_{(x,y)}/G_{average}$, and $B_{(x,y)}/B_{average}$.

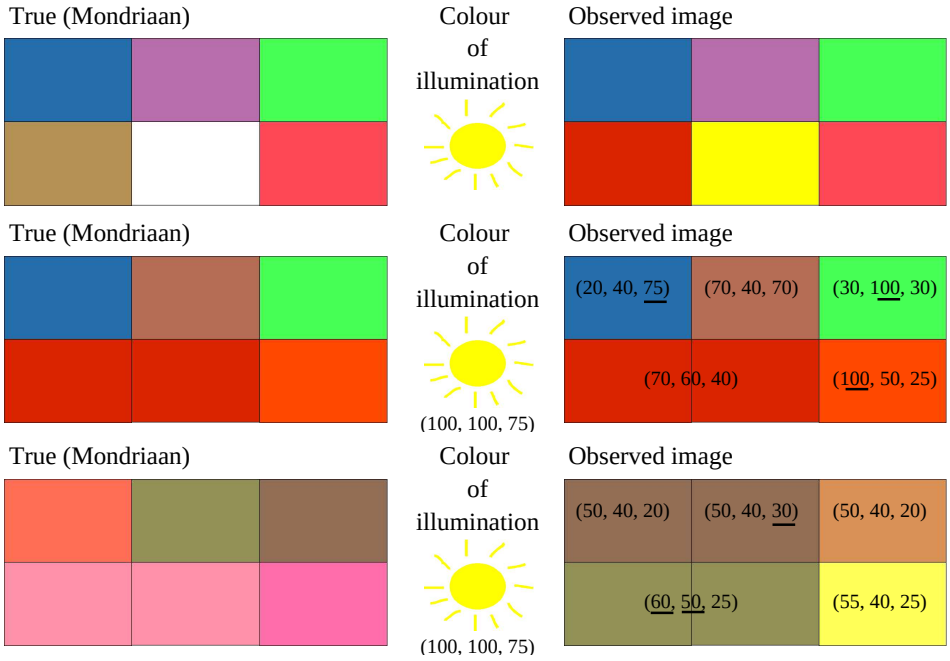


Figure 4.15: *Illustration of the Retinex method (reproduced after G. D. Finlayson talk: Resurrecting Retinex), see text for explanation.*

The other naive method is a version of one of the first colour constancy method *The Retinex method*, introduced by Land and McCann [17]. The version introduced is the infinite-path-without-reset case as described in [3] which scales each channel by the maximum response i.e. $R_{(x,y)}/R_{max}$, $G_{(x,y)}/G_{max}$, and $B_{(x,y)}/B_{max}$. Figure 4.15 illustrates the potential performance and problems of the retinex theory. In the first column the colours of a Mondriaan image viewed under the canonical illumination is illustrated. Mondriaan images are named after the Dutch painter Piet Mondriaan (1872-1944) who composed images consisting of patches of different uniform colours. The second and third columns illustrates the colour of the illumination and the appearance of the images as observed by the camera, respectively.

In the first row there is a white patch present which in the observed image will correspond to the colour of the illumination and hence its colour may be retrieved. The second row illustrates the case where there is no white patch present but where a correct estimate of the illumination may be obtained by the maximum in each channel. The last row, however, illustrates the major problem with the method, i.e. that the maximum in each channel not necessarily gives the colour of the illumination. Especially in images with saturated (clipped) colour bands this is a problem and gives that the method is working with poor performance [11]. Another problem with retinex is that it tries to uniquely characterize the illumination colour, which is not a realistic approach as a white wall viewed under a green light looks like a green wall under a white light.

This lead to the third method *Gamut Mapping*, introduced by Forsyth [10]. The method is based upon the philosophy that for a given observed image there is a gamut of plausible estimates of white.

First consider figure 4.16 which illustrates a series of colours with red and adapted red responses, using R_{max} for normalization. Using the constrain that von Kries adapted sensor responses are in the range from zero to one gives that all values above 80 for the red channel are plausible estimate of a potential

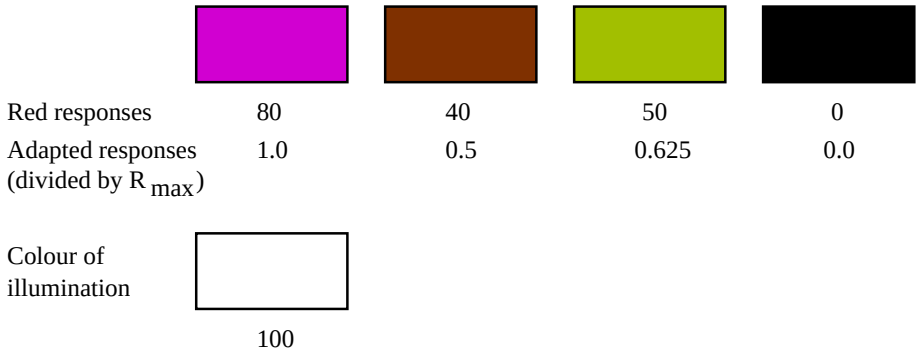


Figure 4.16: *Red and adapted red responses for a series of colours.*

white triplet i.e. the criteria has also to be satisfied for the green and blue channel. This gives that the red response of the four patches (left to right) are bound to be in the range $0 \ldots 1$, $0 \ldots 0.5$, $0 \ldots 0.625$, and finally $0 \ldots 0$, for any plausible illuminant.

The above illustrates the fundamental idea behind gamut mapping which includes illumination sources with infinite power. Clearly, in realistic condition such illumination sources may be regarded as not existing and hence the method may be further constrained.

Another, way of using the principle of gamut mapping is to use it in connection with a canonical gamut. The philosophy is that by acquiring reflectance spectra from hundreds of "daily life" objects it is possible to derive a canonical gamut giving a convex hull of sets of RGB values that would arise if the objects where illuminated by the canonical illuminant. Next, consider an RGB triplet illuminated by an light source such that it no longer lie within the canonical gamut. The set of diagonal transformations mapping the triplet back to the canonical gamut represents the set of possible unknown illuminations.

The sets of diagonal transformations may further be constrained by excluding mappings from the canonical to the unknown illuminant which gives RGB triplets that only can arise from very atypical illuminants [5]. As examples of illumination constrains are Finlayson [5] and Barnard [2] successfully restricting the possible illuminations to a convex hull given by the six daylight phases provided by Judd *et al.* [14], the CIE standard illuminants A, B, C, [36], a 2000K plankian radiator, and uniform white. Other ways of constraining the possible sets of transformations may be establish, as in [2] where varying illumination over the image is used.

To illustrate the method consider the following example. First the fact that if the illumination may be exactly modelled by a diagonal transformation applied to (R, G, B) then it will also be exact modelled by a diagonal transformation applied to $(R/B, G/B)$ [5]. Next, consider a database consisted of four reflectances with R/B, G/B chromaticities $(1, 1)$, $(3, 2)$, $(4, 4)$, and $(1, 2)$ which define the canonical gamut, as included in figure 4.17 (a). The canonical gamut define the convex hull in which R/B and G/B chromaticities from observed objects viewed under the canonical illuminant are expected to fall within.

Now two observation are made having $(R/B,\ G/B)$ chromaticities $(1, 3)$ (obs 1), and $(2, 4)$ (obs 2), included in figure 4.17 (a). Clearly, both observed colours fall outside the canonical gamut. The question is now what does it take to transform them to the canonical gamut ? For both observations there are potentially four known reflectances they may come from. Consider obs. 1, it takes a scaling of 1.0 and $1/3$ to transform $(1, 3)$ to the known reflectance $(1, 1)$, but transforming it to $(4, 4)$ requires a mapping of 4 and $4/3$. Altogether, there are four potential maps for each observation giving the two convex hulls as included in figure 4.17 (b). For each observation the convex hull spanned by the coefficients transforming them to one of the known reflectances encloses the possible mappings of one of the colours in the canonical gamut to the colour of obs. 1 or 2, respectively. Further, if obs. 1 and 2 are illuminated by the same

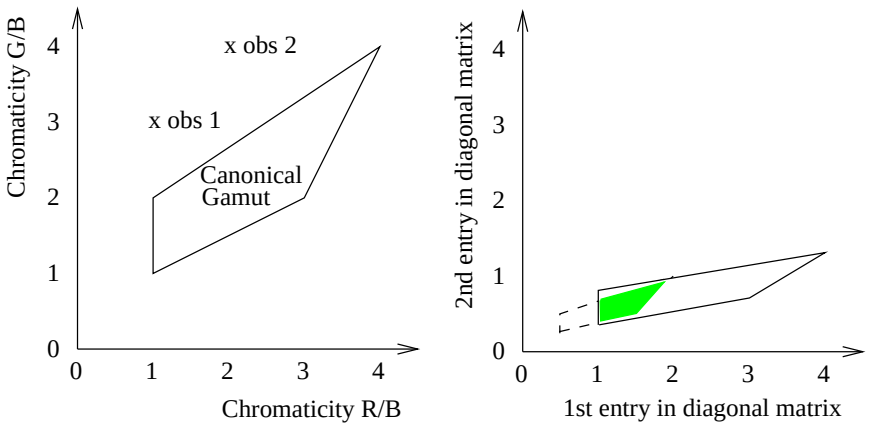


Figure 4.17: *Illustration of the principle of gamut mapping. (a), illustration of the convex hull (solid lines) defined by the reference reflectances together with location of observation 1 and 2. (b), illustration of the convex hull spanned by the entries in the diagonal matrix mapping observation 1 and 2 to the canonical gamut. Green area shows the intersection between the two hull sets and gives the sets of candidates for the transformation matrix.*

light source then the mapping has to account for both of them simultaneously, hence the potential map may be constrained to the union of the two convex hulls, included as the shaded area in figure 4.17 (b).

The convex hull of distinct $(R/B,\ G/B)$ in the image is called the image gamut which gives sets of mappings as candidates for the entries in the diagonal transformation matrix. Hence, some heuristic method has to be used to choose a final transformation. Several criteria may be used. In the 2D $(R/B,\ G/B)$ chromaticity plane the hull centroid have been used [11], whereas in 3D RGB-space [10] have shown that using the intersection set, which has maximum volume, gives a more colourful and better estimate than the centroid point.

Although gamut mapping is one of the colour constancy methods that have been able to adapt to realistic images it is still constrained by the fact that it has to be "trained". Furthermore, it has also a limited capacity to handle images with severe specularities [11] which so to speak "explodes" the canonical gamut.

The last statistically based method presented will be the *Comprehensive Colour Image Normalization* recent introduced by Finlayson *et. al* [6]. The method is based on that the reflected light of the objects in the scene may be described by Lambertian (diffuse) reflection. The method normalizes the image in an iterative process first due to lighting geometry and then illumination colour and so forth until the process converge. Finlayson *et. al* show that the process always converges and that the convergent image is unique. The normalization process is equivalent to first normalize RGB-triplets (i.e. calculate chromaticities) and then scale each R, G, and B channel (i.e. "gray world" normalization). These two steps are then repaeted to convergence. The method has been used on several image databases with an almost near perfect performance. However, it is new and needs more investigation especially into its ability to handle images with specularities.

4.3.2 Physically Based

As introduced in section 4.2 the dichromatic reflection model of Type I has the property that the reflected light is a additive mixture of body and surface reflection, where the colour of surface reflection is closely related to that of the light source. Further, having two objects with pronounced surface reflection gives two dichromatic planes in the colour cube with the surface reflection vector $\mathbf{C}_S$ directed towards the

colour of the illumination. Hence, the intersection of the planes will give the direction towards the colour of the illumination i.e. the gray line.

Perhaps easier it is to understand the principle by transforming the colour vector into its corresponding chromaticities as introduced in section 4.2.1. In this case the dichromatic plane is represented by a line in the chromaticity plane and the intersection between lines from two dichromatic objects will give the chromaticity of the light source. A principle also known as *chromaticity convergence* clearly illustrated in figure 4.11 (b) for the image of the plastic objects.

Using the method of intersecting dichromatic plane for estimation of the illumination colour have indeed been proposed by several authors [18, 31, 12] for three dimensional RGB images and more recently for multispectral images [28]. Though the method in theory is elegant and powerful as it is able to operate on images having highlights or more specifically is relying on the presence of highlights the method do not work well in practice. There are two reason for this. First the intersection calculation is very sensitive to image noise i.e for objects with approximately same colors their lines will be nearly co-linear so small orientation shifts will give large deviation in the intersection point and hence large estimate uncertainty. Secondly, one have to locate two specular surfaces. Whether this is possible will depend both on the characteristic of the objects material and the light source.

One way of getting around the problems of finding the illumination estimate by intersection between two planes may be by setting constraints on the possible colours of the illumination. Very recently Finlayson and Schaefer [9] showed that it is possible to model most "daily life" illuminants by a Plankian locus. They verified this by measuring 172 lights sources including daylight and fluorescent. Under this constraint it is now possible to find an estimate of the illuminant colour simple by the intersection between the fitted line and the model locus. Hence, it is possible to achieve an estimate even with only one specular surface in the image. The method is used on a "realistic" image of a plant viewed under three different light sources D65, TL84 and CIE A (blue, white, and yellow) and gives very good estimates of the actual illumination. However, the method is not complete physics based, but a combination of statistics about the illumination most likely to occur (though physically modelled) and a physics based reflection model.

In the preceding section a minor "state of the art" of colour constancy was given. It showed that there are several methods emerging, but they are all still quite restricted to specific condition for the image formation.

4.4 Summary

In the above a tutorial into the fundamentals of colour image formation, the dichromatic reflection model, modelling of daylight spectra, and a brief introduction to colour constancy has been presented.

In the following 4 chapters these theories and models will be used as a basis for development methods with objectives varying from assessment of illumination conditions, segmentation of images, to design of a dedicated sensor all operating on vegetation in varying daylight conditions.

A central problem of the following chapters will be adaptation to changing illumination conditions of daylight. Therefore a minor presentation of some of the existing colour constancy methods was included.

References

[1] Anonymous. Colorimetry. Technical Report 2. edition, Commision Internationale de L'Eclairage (CIE), 1986.

[2] K. Barnard, G. Finlayson, and B. Funt. Color constancy for scenes with varying illumination. *Computer Vision and Image Understanding*, 65(2):311–321, February 1997.

[3] D. H. Brainard and B. A. Wandell. Analysis of the retinex theory of color vision. *Journal of Optical Society of America A*, 3:1651–1661, Oct. 1986.

[4] S. Chandrasekhar. *Radiative Transfer*. Dover, New York, 1960.

[5] G. D. Finlayson. Color constancy in diagonal chromaticity space. In *Fith International Conference on Computer Vision*, pages 218–223, Los Alamitos, CA, 1995. IEEE Comput. Soc.

[6] G. D. Finlayson, B. Schiele, and J. L. Crowley. Comprehensive colour image normalization. In H. Buckhardt and B. Neumann, editors, *5th European Conference on Computer Vision*, volume 1, pages 475–490. ECCV, 1998.

[7] G.D. Finlayson, M.S. Drew, and D.V. Funt. Color constancy: Generalized diagonal transform suffice. *Journal of Optical Society of America A*, 11:3011–3020, 1994.

[8] G.D. Finlayson, M.S. Drew, and D.V. Funt. Spectral sharpening:sensor transformations for improved color constancy. *Journal of Optical Society of America A*, 11:1553–1563, 1994.

[9] Graham Finlayson and Gerald Schaefer. Single surface colour constancy. To appear in: 7th Color Imaging Conference, Scotts dale, Arizona, November 1999.

[10] D. Forsyth. A novel algorithm for color constancy. *Internatinal Journal of Computer Vision*, 5:5–36, 1990.

[11] B. Funt, K. Barnard, and L. Martin. Is machine vision good enough ? In H. Buckhardt and B. Neumann, editors, *5th European Conference on Computer Vision*, volume 1, pages 445–459. ECCV, 1998.

[12] G. Healey. Estimating spectral reflectance using highlights. *Image and Vision Computing*, 9(5):333–337, Oct. 1991.

[13] B.K.P. Horn. *Robot Vision*. MIT Press, Cambridge, 1986.

[14] D.B. Judd, D.L. MacAdam, and G.W. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54, 1964.

[15] G.J. Klinker. *A Physical Approach to Color Image Understanding*. A.K. Peters, Ltd., Wellesley, Massachusetts, 1993.

[16] G.J. Klinker, S.A. Shafer, and T. Kanada. A physical approach to color image understanding. *International Journal of Computer Vision*, 1(4):7–38, 1990.

[17] E. H. Land and J. J. McCann. Lightness and retinex theory. *Journal of Optical Society of America*, 61(1):1–11, 1971.

[18] H.-C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A*, 10(3):1694–1699, 1986.

[19] H.-C Lee, E.J. Breneman, and C.P. Schulte. Modeling light reflection for color computer vision. *IEEE trans. on Pattern Analysis and Machine Intelligence, PAMI*, 12(4):402–409, 1990.

[20] B. Maxwell. *Segmentation and Interpretation Using Multiple Physical Hypotheses og Image Formation*. PhD thesis, The Robitics Institute, Carnegie Mellon Unversity, Pittsburgh, Pennsylvania, 1996.

[21] C.L. Novak and Shafer S.A. *Encyclopedia of Artificial Intelligence*, chapter "Color Vision", pages 192–202. J. Wiley and Sons, 1992.

[22] C.L. Novak and S.A. Shafer. Method for estimating scene parameters from color histograms. *Journal of the Optical Society of America A*, 11(11):3020–3036, 1994.

[23] M. Oren and Nayer S. K. Generalization of the lambertian model and implications for machine vision. *International Journal of Computer Vision*, 14(3):227–251, April 1995.

[24] H. Palus. Counting of colored objects using highlights. In *Industrial Lasers and Inspection Polarisation and Color Techniques in Industrial Inspection*. EOS/SPIE, June 1999.

[25] F. Pla, F. Ferri, and M. Vicens. Colour segmentation based on a light reflection model to locate citrus for robotic harvesting. *Computers and Electronics in Agriculture*, 9:53–70, 1993.

[26] S.A. Shafer. Descibing light mixtures through linear algebra. *Journal of the Optical Society of America A*, 72(2):299–300, February 1982.

[27] S.A. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985.

[28] H. Stokman and Th. Gevers. Recovering the spectral distribution of the illumination from spectral data by highlight analysis. In *Industrial Lasers and Inspection Polarisation and Color Techniques in Industrial Inspection*. EOS/SPIE, June 1999.

[29] S. Tominaga. Dichromatic reflection models for a variety of materials. *COLOR research and application*, 19(4):277–285, 1994.

[30] S. Tominaga. Dichromatic reflection models for rendering object surfaces. *Journal of Imaging Science and Technology*, 40(6):549–555, 1996.

[31] S. Tominaga and Wandell B.A. Standard surface-reflactance model and illuminant estimation. *Journal of Optical Society of America A*, 6(4):576–584, April 1989.

[32] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of Optical Society of America*, 57:1105–1114, 1967.

[33] L.B. Wolf. Spectral and polarization stereo methods using a single light source. In *Proceedings of the First International Conference on Computer Vision (ICCV)*, pages 708–715, London, June 1987. IEEE.

[34] L.B. Wolf, S.K. Nayar, and M. Oren. Improved diffuse reflection models for computer vision. *International Journal of Computer Vision*, 30(1):55–77, 1998.

[35] L. B. Wolff. Diffuse-reflectance model for smooth dielectric surfaces. *Journal of Optical Society of America A*, 11(11):2956–2968, November 1994.

[36] G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, New York, 2 edition, 1982. page 28.

Chapter 5

Classifying Illumination Condition from Two Light Sources by Colour Histogram Assessment

Classifying Illumination Condition from Two Light Sources by Colour Histogram Assessment

H.J. Andersen and E. Granum

Abstract

We investigate and propose a method assessment of illumination condition covering two light sources. The method may be of some support for colour vision and multispectral analysis methods that rely on specific illumination condition. It is constrained to classifying illumination condition for dielectrical objects illuminated by two light sources. The reflected light is modeled by the dichromatic reflection model, which describes the light as a sum of its body and surface reflections. Further, reflected light from an object illuminated by two light sources may give from one to four primary reflections depending on the condition and it may be expressed as an additive mixture of these reflections. An additive mixture of two reflections expressed in chromaticities is limited to fall within the area enclosed by the chromaticities of the primary reflections of the light sources. So after finding the set of *primary chromaticities* enclosing the pixels points' chromaticities it is possible for one to assess the current illumination condition. Since the method operates on pixel points globally, it is independent of illumination geometry and hence may be used on irregular objects. Two experiments are performed. One uses regular objects in a well-control laboratory environment demonstrating that the pixel-points distribution is as expected. The second experiment demonstrates the method's potential use in support of spectroscopic analysis of vegetation through assessing the illumination condition of barley plants in an outdoor illumination condition.

Published in: Journal of Optical Society of America A., Vol 17, No 4, pages 667-676, April 2000

Extended with figure 5.1, 5.4, 5.8 and images (c) and (d) in figure 5.14.

5.1 Introduction

A problem that often limits or causes color and multispectral image analysis methods to fail is the inability to handle images acquired under varying illumination condition; in particular, outdoor application has always been limited by this variation. In this paper a method for assessment of the illumination condition covering two light sources with varying characteristics is introduced. It is introduced as a method that may support color image and multispectral analysis methods that rely on a specific illumination condition (traditionally, a diffuse condition with one light source). Clearly, having information about the current illumination condition for the scene of interest will make it possible to either change method to a more applicable one or to adjust the current method as part of an active vision system.

As an example, consider object recognition by spectral analysis. Generally this method relies on the fact that the reflected light from the object is either pure or dominated by body "diffuse" reflection. If the body reflection is divided by the spectra of the illumination, the reflectance characteristic of the object is retrieved and may be used for discriminating between objects or in analysis of them. However, object spectra acquired in realistic conditions often do not consist of pure body reflection; instead, the reflected light from them is a mixture of body and surface reflection (specular reflection), which may make spectra difficult to analyze.

Whether an object will give surface reflection depends on three conditions: the characteristics of the light source, the photometric angles, and the optical roughness of the object. The light sources has to produce light with a uniform direction (linear or spherical), the photometric angles have to be such that the angles of the incident light and emitting light are approximately equal (specular or near specular), and the material's optical roughness has to be so smooth that it gives pronounced surface reflection. Concentrating on objects give pronounced surface reflection reveals that there are two ways of avoiding surface reflection; either by having a light source that gives diffuse light (nondirectional light) or by placing the object relative to the camera and the light source so that surface reflection does not appear. In this study we will concentrate on how changes in the characteristic of the illumination, i.e. from diffuse to directional light from two light sources, will change the characteristics of the reflected light from dielectrical objects.

The method is based upon the dichromatic reflection model [13] , which in recent years has been used within colour vision for image segmentation [5], analysis of highlights [6], estimating scene properties [11], estimating colour of illuminant for colour constancy [17] and more recently for generation of computer graphics rendering [16]. In this paper we investigate its use to develop a basis for assessment of the illumination condition. This is done by analysis of the pixel points' location in the chromaticity plane for the object of interest. The analysis is constrained to dielectrical objects that have low optical roughness [18] and hence show pronounced surface reflection and illumination conditions from one or two light sources of varying quality. The method is in particular suitable for images acquired under daylight condition, which may consist of the two light sources direct sunlight and ambient skylight. A more comprehensive analysis and taxonomy of the formation of colour images may be found in [10].

In Subsection 5.2.1 we review the dichromatic refflection model and outline its representation in the chromaticity plane; this is followed in Subsection 5.2.2 by a discussion of the model's representation in the case of two illumination sources, which concludes with the first basis for the method to be developed. The reference to the correlated color temperature (CCT) of balckbody radiators for typical light sources is given in Subsection 5.2.3, and the distribution of the pixel points mapped onto the chromaticity plane is described in Subsection 5.2.4. The basis for development of a practical implementation of the method for assessment of the daylight condition is also established in Subsection 5.2.4. Two experiments are then presented in Section 5.3. One is performed in well-controlled conditions with regular objects, showing that the distribution of the chromaticities are as expected. The second uses the introduced method for assessment of daylight conditions for classification of 36 images of barley plants acquired in varying daylight conditions.

5.2 Modeling

5.2.1 The Dichromatic Reflection Model

The reflected light from and object $L(\Theta, \lambda)$ is a function of the wavelength λ and the photometric angles Θ, which include the illumination direction angle (i), the viewing angle (e), and the phase angle (g) (figure 5.1a). The dichromatic reflection model as formulated by [13], describes the reflected light from a nonhomogeneous dielectrical object as the sum of its body and surface reflections:

$$L(\Theta, \lambda) = m_B(\Theta)L_B(\lambda) + m_S(\Theta)L_S(\lambda) \tag{5.1}$$

where the terms $L_B(\lambda)$ and $L_S(\lambda)$ are the spectral power distribution of the surface and body reflection, respectively (figure 5.1b). The weights $m_B(\Theta)$ and $m_S(\Theta)$ are geometrical scaling for the body and surface reflection, respectively. The model assumes that the spectral power distribution and the geometrical scaling factors are separable. This condition is approximately satisfied by many materials having high contents of water or oil, because their index of refraction is nearly constant over the visible spectrum [8].

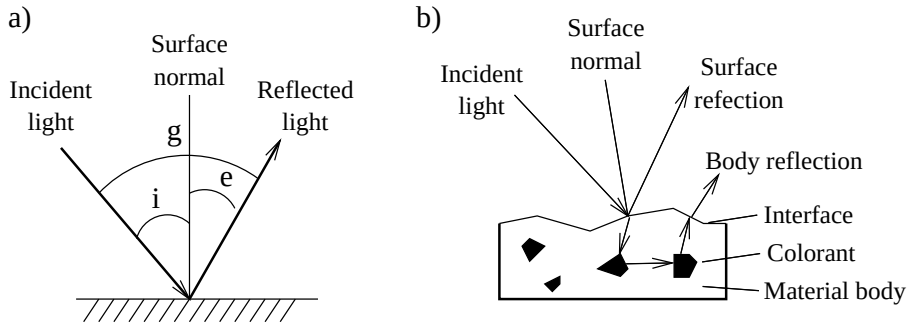


Figure 5.1: *a) Photometric angles. b) Reflection components from a non homogeneous opaque dielectrical object.*

The dichromatic reflection model as described in eq(1) covers the whole continuous light spectrum. This dimension of the spectrum is reduced in a camera reduced by the process of *spectral integration*:

$$C_f = \int L(\Theta, \lambda)\tau_f(\lambda)s(\lambda)d\lambda \tag{5.2}$$

where $L(\Theta, \lambda)$ is the incoming light in the camera, $\tau_f(\lambda)$ the spectral transmittance of filter f which limits the integration to a subband and $s(\lambda)$ the sensitivity of the camera.

In a standard RGB (red-green-blue) camera the colour spectrum is represented by three subbands, the three dimensional colour space, such that a pixel at location (x, y) in the image will be represented in colour space by a vector $C(x, y) = [C_{red}, C_{green}, C_{blue}]$, which for convenience may be refereed to as $[Red, Green, Blue]$. Here the dichromatic reflection model in colour space may be expressed by:

$$C(x, y) = m_B(\Theta)C_B + m_S(\Theta)C_S \tag{5.3}$$

which describes the so-called *dichromatice plane* in the Red, Green, and Blue space spanned by the body and surface reflection vectors, C_B and C_S, respectively, of the object. The object pixel points will be distributed in this plane.

The reflected light from the body and surface components $L_{(B,S)}(\Theta, \lambda)$ is formed by the incident light $E(\lambda)$ weighted with the reflectance $\rho(\lambda)$ of the body and surface, respectively, for every wavelength. Hence the dichromatic model may be expressed as:

$$C(x,y) = m_B(\Theta) \int E(\Theta,\lambda)\rho_B(\lambda)\tau_f(\lambda)s(\lambda)d\lambda + m_S(\Theta) \int E(\Theta,\lambda)\rho_S(\lambda)\tau_f(\lambda)s(\lambda)d\lambda \tag{5.4}$$

The standard dichromatic reflection model equation (5.1), also incorporates an assumption called the Neutral Interface Reflection (NIR) [8], in which it is assumed that the relative spectral composition of the surface reflection component is identical to that of the light source (i.e., the surface reflectance $\rho_S(\lambda)$ is constant). More recently, [15] this assumption has been categorized this as the dichromatic reflection model of type I, which has been shown experimentally valid for a variety of materials including, plastics, ceramics, tiles, fruits, leaves and other materials [8], [15], [12]. Alternatively, the standard model equation (5.1) may include a tranferfunction of the interface ((i.e., the surface reflectance $\rho_S(\lambda)$ is not constant) which [15] is categorized as type II and is experimentally found to be valid for silk, wool, satin, polyester, cotton velveteen and velour.

For analysis of the colour of the reflected light independent of the scale of intensity, it is convenient to transform the colour vector C into its corresponding chromaticities. This is done by taking the first norm of the colour vector, $r = Red/(Red + Green + Blue)$, $g = Green/(Red + Green + Blue)$, and $b = Blue/(Red + Green + Blue)$. Since $r + g + b = 1$, it is sufficient to consider the chromaticities in only two dimensions, e.g. the *rg-chromaticity plane*. Another useful property that is conveniently expressed by the chromaticities is that an additive mixture of two lights producing a third light will be a linear combination of the first two as [7]:

$$\begin{aligned} r_3 &= \tfrac{S_1}{S_1+S_2} r_1 + \tfrac{S_2}{S_1+S_2} r_2 \\ g_3 &= \tfrac{S_1}{S_1+S_2} g_1 + \tfrac{S_2}{S_1+S_2} g_2 \end{aligned} \tag{5.5}$$

where S_1 and S_2 are the sums of the colour vectors (i.e, $S = Red + Green + Blue$).

5.2.2 Pixel point distribution

As introduced in Subsection 5.2.1, a dichromatic plane will be represented in the chromaticity plane as a straight line segment connecting the chromaticities of the body and surface reflection vectors. The pixel points from the object will be distributed along this line. In this subsection we will see that having information about the pixel point distribution in the chromaticity plane makes it possible to assess the illumination condition. The analysis will be constrained to dielectrical objects that show pronounced surface reflection when illuminated by two light sources, E_1 and E_2.

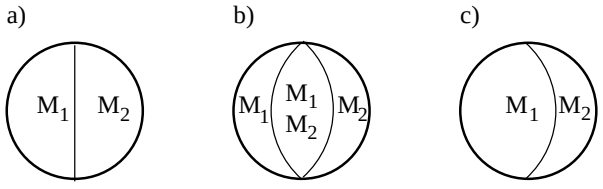


Figure 5.2: *Perspective projection of a sphere with reflections: M_1, due to illuminant E_1 and M_2, due to illuminant E_2 a) Reflection due to only one of the illuminations E_1 or E_2. b), reflections due to each illumination source E_1 and E_2, or E_1, E_2, a mixture of both. c), same as b) but with illimination source E_1 as a ambient light source.*

Two sources of illumination E_1 and E_2, both give two primary reflections $M_{1,B}$ and $M_{2,B}$ of the body and two primary reflections of the surface $M_{1,S}$ and $M_{2,S}$ for the object of interest. Figure 5.2 illustrates three possible conditions for the illumination of a sphere under perspective projection.

In the case in which all pixel points are illuminated by only one of the illumination sources E_1 or E_2, their chromaticities will be distributed along the line segments connecting their body and surface reflection chromaticities $b(E_1)$ to $s(E_1)$ and $b(E_2)$ to $s(E_2)$, respectively, as shown in figure 5.3 a) by thick solid lines.

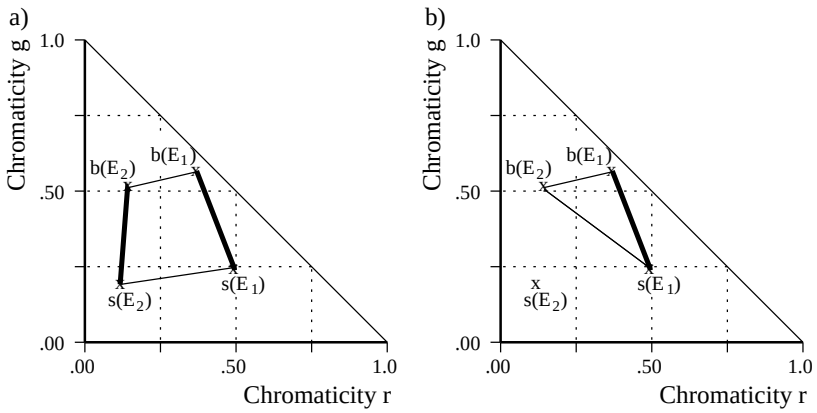


Figure 5.3: *Possible pixel point locations for a dielectrical object illuminated by light sources E_1 and E_2 having primary chromaticities $b(E_1)$, $b(E_2)$, $s(E_1)$, and $s(E_2)$. a) Primary chromaticities form a convex tetragon, b) only one light source produces surface reflection. Possible pixel-points locations are limited to areas within the interior line segments.*

However, realistic scenes seldom consist of pixel points illuminated by only one source. If there are pixel points reflecting light from both illumination sources E_1 or E_2 then these may be expressed as an additive combination of all four primary chromaticities, $b(E_1)$, $b(E_2)$, $s(E_1)$, and $s(E_2)$. Hence it is now possible to locate the pixel points' chromaticity distribution in the chromaticity plane within the area between the line segments where the possible additive mixtures of the four primary chromaticities may fall.

Table 5.1: *Pixel point distribution area due to reflections produced by illumination sources E. $b(E)$ and $s(E)$ are primary body and surface chromaticities; "x", indicates which primary chromaticities that will contribute to the additive mixture of the reflected light (i.e., enclosing the area of possible pixel chromaticities, figure 5.3 a) and b)).*

	Reflection due to illumination sources					
	E_1		E_2		$E{=}E_1{=}E_2$	
Pixel Point Distribution	$b(E_1)$	$s(E_1)$	$b(E_2)$	$s(E_2)$	$b(E)$	$s(E)$
Tetragon	x	x	x	x		
Triangle	x	x	x			
Triangle	x		x	x		
Line	x		x			
Line					x	x
Point					x	

So far, it has be assumed that the two illumination sources E_1 and E_2 will give surface reflection from the object. However, this may not necessarily be the case, and often one or even both of the sources may give only body reflections. Further, there is the possibility that the two sources also will have the same spectral

composition. Then the possible area of additive mixtures will be limited, owing to the number of omitted or similar primary chromaticities. Table 5.1, summarizes the pixel points' chromaticity distribution given by the different illumination conditions E_1 and E_2. The table is limited to the case in which there is a reflections due to a mixture of the two sources (figure 5.2, 2a and 2b).

Figure 5.3 b) illustrates the case in which surface reflection $S(E_2)$ is omitted. An import example of an illumination condition giving such a triangle distribution is daylight, which may consist of a mixture of skylight and sunlight. Here skylight is an ambient and hence diffuse light source that does not giv pronounced surface reflection compared with direct sunlight, which is likely to give pronounced surface reflection.

It is important to notice from figure 5.3 that all possible areas of additive mixtures are included in the area defined by the four primary chromaticities. Hence a strategy for finding the pixel points' chromaticity location could be to find this location and then decompose it into the underlying hypotheses.

In the above analysis it has been assumed that the primary chromaticities will form a convex tetragon, which will be the case in most daily life situations. However, for objects with very distinct colours the primary chromaticities may form a concave tetragon or even give dichromatic plane that cross the chromaticity plane. If they form a concave tetragon, then the area enclosed will be limited to the encircling triangle when both illumination sources give surface reflection.

5.2.3 Modeling Illumination Changes of Approximate Black Body Radiators

In Subsection 5.2.2 it was shown how the chromaticity of pixel points may be distributed under various illumination conditions. This subsection will focus on how illumination sources that may be described (approximated) by their CCT's will influence the location of the primary chromaticities and how they may change location. The CCT of a light source is related to the temperature of a blackbody that emits light of a similar spectral composition.

Looking more closely at eq(4) we may note that the formation of an image is depending on the following three sets of parameters 1) $E(\Theta, \lambda)$, the spectral composition of the illumination, 2) $\rho(\lambda)$, the reflectance of the object, and 3) $\tau(\lambda)$ and $s(\lambda)$, the characteristics of the camera.

For a given camera it is possible to obtain the tristimulis characteristics, and further assuming that the reflectance of the object of interest is known, then only the spectral composition of the illumination is left to potentially cause a change in the color of the reflected light.

The spectral composition of daylight and most general purpose illuminants (e.g. tungsten filemant, tubular, compact fluroscent, and all other kinds of gaseous discharge electrical lamps) may be characterized according to their *Correlated Colour Temperature* (CCT), as outline in [2]. Especially valid for daylight this characterization is as formulated in CIE's standard [1] on the basis of Judd *el al* [4], illustrated in figure 5.4.

An important property of radiation from a blackbody is that its spectral composition at increasing (color) temperature gradient moves from reddish to greenish and finally bluish light. Figure 5.4, illustrates this property for daylight when represented by its CCT.

A result of this graduated change in composition is that the locus formed by the primary body and surface chromaticities will traverse from red towards green and finally blue with increasing CCT; this is the reason that objects in most "daily life" illumination will form convex tetragons.

5.2.4 Finding the Pixel Point Distribution

Different CCT's of the illumination will make the line segment from dichromatic planes change position in chromaticity plane. Having suitable models of the spectral composition of the illumination at varying CCT makes it possible to derive a body and surface reflection locus for the object of interest. Figure 5.5,

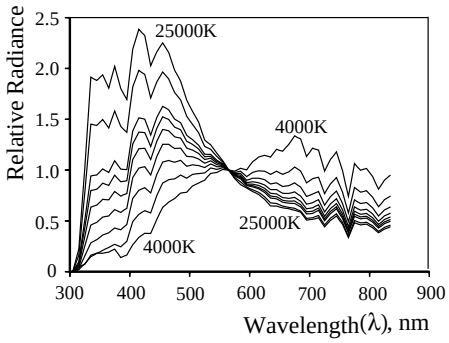


Figure 5.4: *Relative spectral radiance of daylight normalised at 560nm, as specified in [1]. Correlated Colour Temperature varying from 4000K to 25000K, in steps showing 4k, 5k, 6k, 7k, 8k, 9k, 10k, 15k and 25k degrees Kelvin.*

illustrates a body and surface reflection locus with corresponding CCT's of the illumination E.

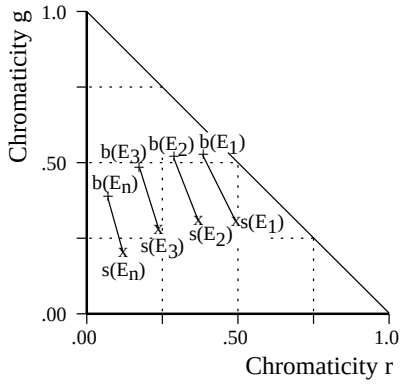


Figure 5.5: *Alternative positions in the chromaticity plane of the dichromatic plane due to varying CCT's of the illuminant* $E_1, \ldots, E_n$

One can now scan through possible tetragons defined by combining pairs of primary chromaticities ($b(E_1)$, $s(E_1)$ and $b(E_2)$, $s(E_2)$, etc.) and for each determine the number of enclosed pixel points. Further, the tetragon that best encloses the pixel-point distribution may be determined as the one that has the smallest difference in CCT of the primary chromaticities relative to the number of pixels. From this the illumination condition may be derived according to table 5.1.

The Daylight Heuristic

The analysis above assumes that the illumination and reflectance of the object may be modeled accurately without consideration to interreflection or second-order scattering. However, for outdoor scenes it is often not practically possible to achieve such well-defined environment for image formation. Instead the method for finding the pixel points can be "relaxed", so it is possible to adapt to the complexity of the scene.

First, as already mentioned, daylight consists of two illumination sources, the ambient skylight, which does not give pronounced surface reflection, and the direct sunlight, which is likely to give surface reflection. Hence we have only three loci, and the pixel distribution may be enclosed by a triangle, defined by the

following primary chromaticities: 1) $b_{skylight}$ due to the body reflection from the skylight and 2) $b_{sunlight}$ and $s_{sunlight}$ due to the body and surface reflection, respectively, from the direct sunlight.

Second, without loss af generality, the line segment connecting the two primary body chromaticities $b_{sunlight}$ and $b_{skylight}$ in the triangle may be ignored to compensate for the nonideal condition for image formation and to make the method simpler to implement. Hence the triangle becomes a wedge with its apex located at the chromaticity of the surface reflection ($s_{sunlight}$) and with the first leg directed according to the primary chromaticity of the body reflection ($b_{sunlight}$) at corresponding CCT and the second at the body reflection related to the CCT of the skylight ($b_{skylight}$).

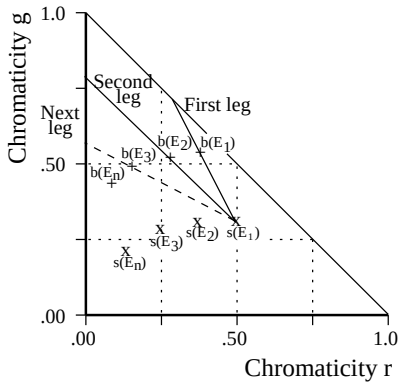


Figure 5.6: *Finding the location in the chromaticity plane of pixel-point distribution under the daylight heuristic. The location is found by spanning a wedge with its apex at the chromaticity of the surface reflection from the sunlight $s(E_1)$ and in the direction of the first leg after the body reflection at the corresponding CCT $b(E_1)$ and in the direction of the second leg at the CCT of the skylight $b(E_2)$. The dashed line illustrates the variable second leg in the scanning procedure running through a range of wedge angles from very wide to narrow. This is repeated for a range of positions of first leg and apex related to varying CCT's of the daylight, $E_1, \ldots, E_n$.*

Further, as sunlight generally will have a lower CCT than skylight, the pixel points may be located by placing the first leg of the triangle/wedge at corresponding CCT's and then vary the second leg for all body reflection at higher CCT's (figure 5.6).

Finally, the wedge that encloses a "sufficient" amount of pixel points and has the smallest angle between its legs may be determined for assessment of the illumination condition. The stop criteria for the search process will be related to the specific context in which the method is being used; examples will be given in the following experiments.

5.3 Experiments

In the following two experiments the potential of the proposed methodology for classifying illumination conditions will be demonstrated. Experiment A (exp. A) is performed with regular dielectrical objects in a laboratory setup with two well-known light sources. Experiment B (exp. B) demonstrates the use of the daylight heuristic for classifying the daylight condition of images of barley plants. This application is important for robust support of spectroscopic analysis for assessment of the growth conditions and classification of different plant species.

5.3.1 Exp. A. Material and Methods

This experiments was performed in a well-control laboratory experiment with two 58-W fluorescent light sources: a reddish Philips TLD 927 (2760K) and a bluish Philips TLD 965 (6900K), both with $R_a > .90$. Their chromaticities were obtained by reference images of a barium sulphate plate illuminated by either light source. A M90 3CCD camera (Jørgen Andersen Ingeniør firma, Denmark) was used with the gamma correction turned off. The camera was white balanced according to the barium sulphate plate illuminated by a Philips TLD 940 (3870K) fluorescent light source. Before calculation of the chromaticity values, the acquired images were adjusted for the black offset.

The test objects were placed against a black background of woolen material that had practically no reflection. They were illuminated by two fluorescent light sources, one on each side, with the camera placed between the light sources. A green plastic cup, a red enameled cup, a lemon, a blue plastic vacuum jug, a green ceramic pot and a sheet of magenta paper were used as test objects. First they were placed so that both light sources were giving both body and surface reflection. Next the illumination condition was changed so the reddish TLD 927 light source was giving only body reflection. Finally, the condition was changed so that both light sources were giving only body reflection.

According to the developed taxonomy, the distribution of the pixel points' chromaticities should change from being a mixture of four to three and finally two primary chromaticities, corresponding to distributions within a tetragon, and, finally, along a line segment.

5.3.2 Exp. A. Results and Discussion

In figure 5.7 the pixel points' chromaticity distribution is illustrated for the six test objects. The chromaticity of the two light sources is shown by white dots in all plots.

The first row shows the pixel points' chromaticity distribution for the green plastic cup and the red enameled cup. Both objects have a highly glossy surface and hence give pronounced surface reflection, which is clearly evident in column (a), where both give reflection clusters traversing from their primary body chromaticities towards the chromaticities of the light sources, as expected from objects with material properties that are assumed to follow the NIR assumption. In column (b) the reddish light source has been changed to give only body reflection; as expected, the reflection vaniishes as it traverses toward the chromaticities, of the light source, and the distribution of the pixel points changes to fall within a "triangle", which is most evident for the green plastic cup. In column (c) both light sources are changed to give only body reflection and consequently both reflection clusters traversing toward the chromaticity of the light sources disappear and the pixel points change to being distributed along a line segment. In figure 5.8 the three different illmination condition for the green plastic cup, is illustrated.

In the second row the blue plastic vacuum jug, a glossy regular uniform object, is shown together with the less uniform lemon. For the vacuum jug the distribution pattern of the pixel points is the same as for the plastic cup and enameled cup. More interesting is that its dichromatic planes from the two light sources cross over each other (this will be discussed below). For the lemon the distribution pattern is not as clear. It was especially difficult to get surface reflection from the reddish light source, because the red channel saturated before the green, and the blue reached values corresponding to the light's chromaticities. This might indicate that despite the objects' high oil contents, it does not follow the NIR assumption as completely as expected.

In the third row a green ceramic pot together with a sheet of magenta paper (formed as a cylinder) are included. The ceramic pot shows NIR properties, and its pixel-points distribution pattern is as expected. Further, because the surface is less glossy than that of the plastic and the enameled objects, the tetragon and triangle are more "filled out" with pixel points. For the magenta paper the situation is somewhat different. Owing to its structure, there is almost no surface reflection. This demonstrates the limitation of the proposed methodology for objects where the optical roughness is so high that it makes it impossible to assess the illumination conditions by their reflection clusters.

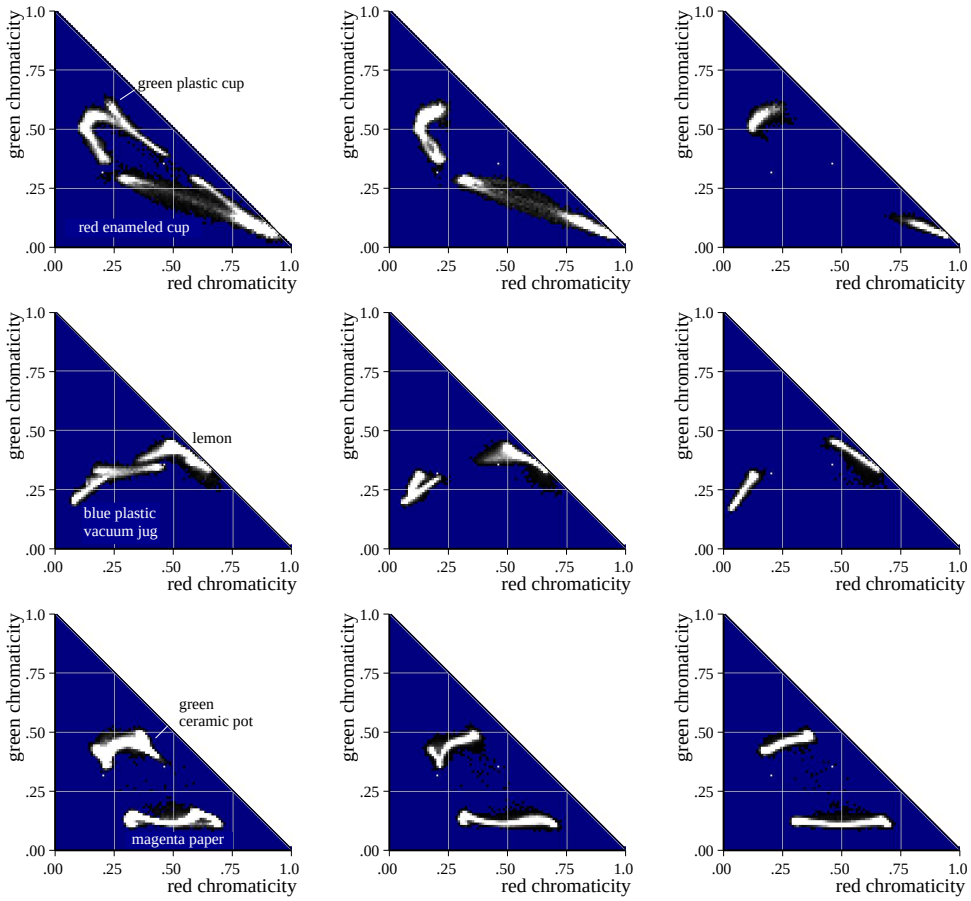
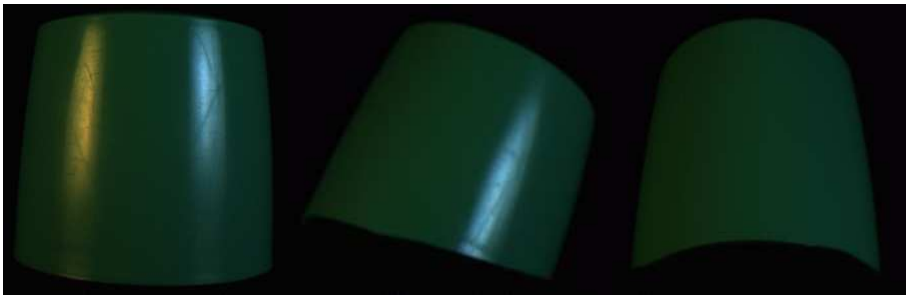


Figure 5.7: *Test objects' pixel-points chromaticity distribution. In column (a) both light sources give body and surface reflection. In column (b) the reddish light source give only body reflection. In column (c) both light sources gives only body reflection. The two white dots in all plots show the locations of the two light sources' chromaticities.*

Looking more closely at the distribution of the reflection clusters for the red enameled cup and the blue vacuum jug reveals that the primary chromaticities of these two objects almost form a concave tetragon and the dichromatic planes cross over each other. To illustrate this, we obtained the reflectance curves for the two objects, figure 5.9. Further, using the tristimulus characteristics of the camera and the light spectra as supplied by the manufacturers, the location of the primary chromaticities for the two objects were modeled, as shown in figure 5.10. The figure shows that the primary chromaticities of the red cup almost give a concave tetragon and that the blue vacuum jug has dichromatic planes that cross over each other. This demonstrates that for objects giving both body and surface reflection there is no guarantee that their primary chromaticities will form convex tetragons.

5.3.3 Exp. B. Material and Methods

A series of images of barley at the 10 to 12 leaf stage was acquired in the field at Taastrup (near Copenhagen) on June 12, 1997, at regular intervals from 6:30 a.m. to 7:20 p.m., with a image size of 637x472 pixels and a resolution of 1 mm/pixel. The weather condition was sunshine from an unclouded sky



(a) (b) (c)

Figure 5.8: *Illustration of the green plastic cup. (a), both light sources giving surface reflection. (b), one the "reddish" light source changed to give diffuse reflection. (c), both light sources giving diffuse reflection.*

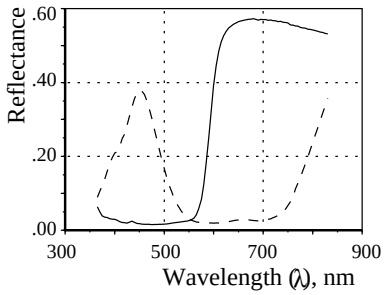


Figure 5.9: *Reflectance of the red enameled cup (solid curve) and the blue vacuum jug (dashed curve), obtained by a spectrometer designed and manufactured by VTT, Oulu, Finland.*

until 3:30 p.m. when the conditions changed to clouded. Images with the sun as primary illumination source and with a cardboard casting shadow over the image field (changing the primary illumination source to skylight) were acquired one immediately after the other. In this way, images of four classes were obtained:

A sunshine, unclouded

B sunshine, clouded

C skylight, clouded

D skylight, unclouded

In total, 36 images were acquired within a CCT range from 4950K to 18000K. The camera was white balanced at a CCT of 5700K for all images, the gain and shutter time was adjusted manually to best exploit the dynamic range of the camera, and the gamma correction was turned off. The CCT of the illumination was both assessed with a Minolta Colour Meter II and reference images of a standard barium sulphate plate. Before calculation of chromaticities, all images were adjusted for the black offset.

In this manner, images with clearly two illumination sources, direct sunshine and skylight (A) and images with generally one diffuse source (B, C, and D) but with differences in the CCT where B is expected to have high values, C medium, and D low were acquired. The analysis was performed on both presegmented and not-presegmented images. Segmentation of the presegmented images into a vegetation and background part was done by a comprehensive interactive method. To find the pixel points' location, we used the

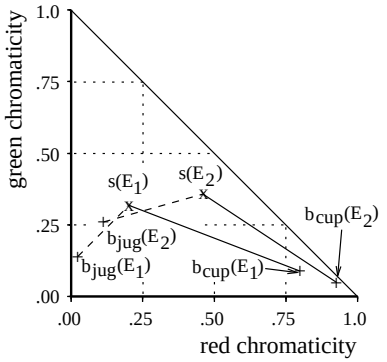


Figure 5.10: *Location of primary chromaticities for the red enameled cup and the blue vacuum jug. Solid and dashed lines connect primary chromaticities due to light sources E_1 (Philips TLD 965) and E_2 (Philips TLD 927), respectively.*

daylight heuristic with the criterion that it enclose at least 90% of the total amount of pixel points. For the not-presegmented images the wedge spanned by a CCT of 4000K and 25000K (i.e., from the lowest to the highest CCT) was used for to normalize all other wedges. For the presegmented images, the wedge enclosing 90% of the pixel points was used to describe the pixel points' dispersion.

For determining the location of the wedge, the reflectance of vegetation derived from the reflectance characteristic in [9] for maize under normal growth condition was used (figure 5.11), and this is in good agreement with characteristics for vegetation (typically crops and weeds), reported elsewhere in the literature [14, 3, 20, 19].

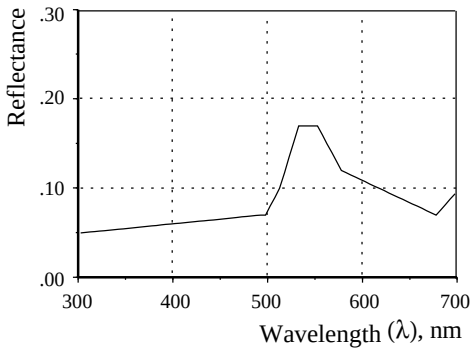


Figure 5.11: *Reflectance characteristic used for modeling the body reflection of vegetation (derived from [9])*

Since vegetation may be assumed to follow the NIR model [8, 15, 12], the surface reflection will have the same spectral composition as that of the illuminant; hence it will correspond to the daylight locus of the camera. The locus will be modeled according to equation 5.4 with $E(\Theta, \lambda)$ after the CIE standard [1] and according to the camera characteristics $\tau_f(\lambda)$ and $s(\lambda)$ after information provided by Sony for the DXC 930P 3CCD camera used by Sony.

The two modeled loci were divided into 18 pairs of primary chromaticities (b/s) with corresponding values of CCT, evenly distributed over the r chromaticity range from 4000 to 25000K. For description of pixel points enclosed by the wedge, the following two features were derived:

1. the generalized variance, $det(COV(r_{enclosed}, g_{enclosed}))$, of the pixel points enclosed;
2. the r chromaticities' mean value of $b(E_1)$ and $b(E_2)$

It is expected that the generalized variance should be able to discriminate class A from classes B, C, and D, while the mean value of the r chromaticity of $b(E_1)$ and $b(E_2)$ will describe the CCT of the illuminant and hence be able to discriminate classes B, C, and D (as these are expected to give point distribution at different CCT's of the illuminant (table 5.1)). In more general terms the generalized variance discriminates between one and two illumination sources, while the mean value discriminates between single illumination sources.

5.3.4 Exp. B. Results

Exp. B. Verification of camera model

For verification of the modeled range for the camera daylight loci, the modeled data according to information from Sony and the CIE daylight standard [1] were compared with the measurements of the barium sulphate plate. For modeling of the two types of data, the procedure in the CIE for the standard observer was used [1].

First, a hyperbolic model of the relation between CCT and the r chromaticity was fitted for the two types of data. From figure 5.12 it may be noted that there is good agreement between the two types until a CCT about 10000K, after which the experimental model have a tendency to over estimate the r chromaticity (10% at 18000K).

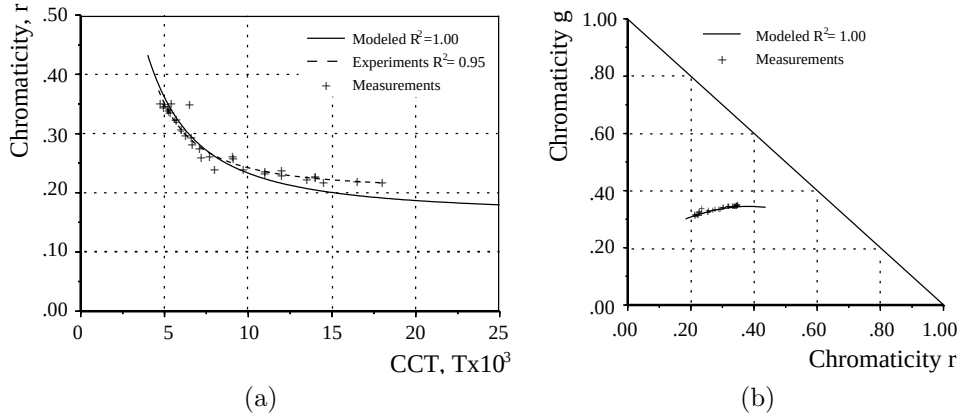


Figure 5.12: *Comparison of modeled and experimentally derived daylight locus for the camera. (a) Relationship between CCT and r chromaticity. (b) Relationship between r and g chromaticity; since the modeled and experimentally derived loci are overlying each other, only the modeled is included.*

Second, a parabolic model of the relation between the r and g chromaticity was fitted to the two types of data, figure 5.12. The modeled and experimentally derived models were compared and there was no significant difference between them (p¡0.05).

Exp. B. Classification of images

Using the derived cluster parameters for each of the 36 images, presegmented and not-presegmented, a linear discriminant analysis was performed to test the hypothesized class structure. The result of this

analysis shows that a class structure is clearly apparent for both sets; table 5.2. Taking the discriminants strictly 5 of 36 images (14%) and 6 of 36 (17%) for the presegmented and not-presegmented images, respectively, were misclassified. However, all "errors" are in the transition regions between neighboring classes where the preclassification is also uncertain.

The descriptors derived from the presegmented and not-presegmented images images are compared in figure 5.13 (a) and (b). As may be noticed, there is very good agreement between the values obtained for the two, $R^2 > 0.95$ for both relationships. Further, figure 5.13 clearly shows that the generalized variance discriminates class A from the others perfectly and that the mean value of the r chromaticity of $b(E_1)$ and $b(E_2)$ describes and classifies the other three classes, with D having the lowest content of red (i.e., highest CCT), A and B the highest and C in between, as expected. Figure 5.14 shows two images acquired immediately one after the other illustrating the change in reflection from the illumination conditions sunshine (A) to skylight unclouded (D).

Table 5.2: *Result of linear discriminant analysis for presegmented and not-presegmented images*

	Number of Images Classified into Class									
From	A		B		C		D		Total	
Class	Pre	Not	Pre	Not	Pre	Not	Pre	Not	Pre	Not
A	11	10	0	1	0	0	0	0	11	11
B	1	0	6	6	0	0	0	1	7	7
C	0	0	2	1	3	4	2	2	7	7
D	0	0	0	0	0	1	11	10	11	11
Total	12	10	8	8	3	5	13	13	36	36
Percent	33	27	22	22	8	14	36	36	100	100

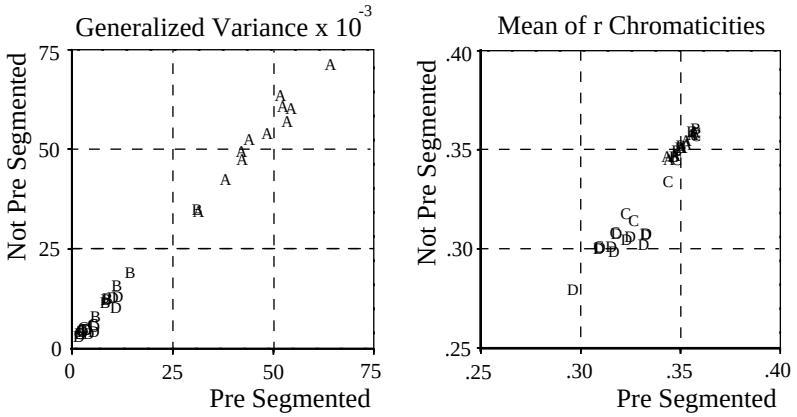


Figure 5.13: *Comparison of descriptors for the presegmented and not-presegmented images. (a) Relation between the generalized variance, $R^2 = .99$, (b) relation between the mean value of the r chromaticity of $b(E_1)$ and $b(E_2)$, $R^2 = .96$.*

5.4 Exp. B. Discussion

Theoretically, the chromaticities of all pixel points should be enclosed by the primary reflection chromaticities. However, because of nonuniform materials and border pixels between two objects, some

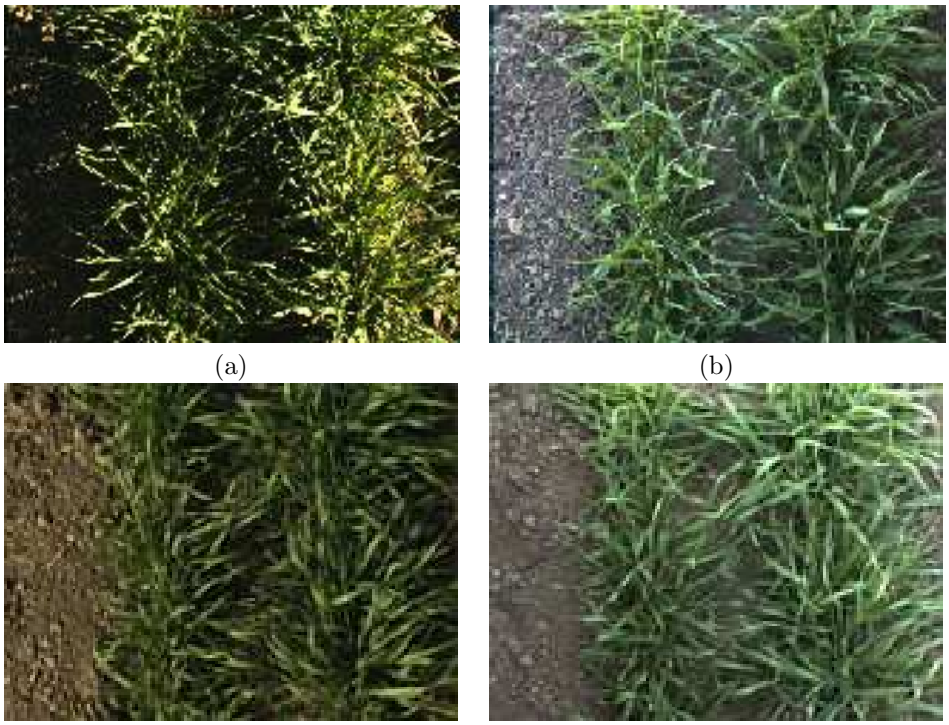


Figure 5.14: *Images of barley. Top row acquired at 6.50am with sunshine and unclouded skylight as illumination sources. (a), classified as class A, with a generalized variance of* $44.1x10^{-3}$*, a mean value of* r_{body} *of 0.37, and a CCT of 4750K. (b), acquired immediately after with a cardboard casting shadow over the image field, classified as class D, with a generalized variance of* 2.1^{-3}*, a mean value of* r_{body} *of 0.31, and a CCT of 14500K. Bottom row acquired at 4.45pm with clouded sunshine and skylight as illumination sources. (c), classified as class B with a generalized variance of* 11.0^{-3}*, a mean value of* r_{body} *of 0.35, and a CCT of 5350K. (d), acquired immediately after with a cardboard casting shadow over the image field, classified as class C with a generalized variance of* 3.5^{-3}*, a mean value of* r_{body} *of 0.32, and a CCT of 9100K*

chromaticities will not fall within the area. To compensate for this, it is necessary to introduce a criterion to determine how, for any pixel points that are regarded as sufficient to describe the pixel-point population for the object of interest, e.g., for a smooth uniform plastic object, the criterion would be different from that for vegetation. Hence the search method and stop criterion are context dependent, and achieving a general solution will be difficult and is beyond the scope of this paper. However, the taxonomy based on the physical modeling of the reflection will be the same.

In the analysis, a heuristic for the illumination condition daylight was establish. It was shown that the problem of finding the primary chromaticities that describe the objects' pixel-point population may be reduced to finding the location of a triangle or a wedge where one of its legs is constrained to be located at corresponding CCT for the surface and body reflection, respectively. It is important to note that this property is useful for "real time" implementation of the method, bcause it radically limits the search space. Also, the daylight heuristic ignores the rare possibility that the primary chromaticities will form a concave instead of a convex tetragon.

For the analysis to be performed, the tristimulis characteristics of the specific camera should be known.

However, this requirement does not seem to be very strict. The CCT and the "r" chromaticity for the experiments and for the model using information from Sony appears to be sufficiently in accord, although Sony provided a general product specification. Also, the reflectance characteristic of maize appears to be an adequate representation of barley for demonstrating the potential of the proposed method. This suggest that the smoothing effect of the spectral integration makes the method less sensitive to inaccuracy of the reflectance characteristics.

The results indicates that it is possible to classify not-presegmented images with the same accuracy as presegmented. Clearly, this result is closely related to the specific application, but it shows that with a fairly accurate reflectance characteristic of the object, it is possible to make assessment without presegmentation. The most important result is that class A is clearly discriminated from the other three classes, as this is the situation where the obtained reflectance characteristic may be a mixture of body and surface reflection.

5.5 Conclusion

Characteristics of illumination conditions for dielectrical objects may be obtained by considering their pixel points' chromaticities dispersion in the chromaticity plane. Through an analysis it was shown that by considering the set of primary chromaticities that enclose the distribution, it is possible to quantify the illumination condition for the image formation. The proposed method is especially suitable for characterization of images acquired under daylight conditions, and a heuristic for such conditions was introduced.

Experimental results were obtained in a well-control labotory environment demonstrating that demonstrated that the reflection clusters change with changing illumination condition, as expected. More important, the introduced method was also used on images of barley acquired in "realistic" outdoor illumination condition where it was able to assess the condition for the image formation.

Acknowledgment

This research was funded by the Danish Research Council under the Optimal Research Programme.

References

[1] Anonymous. Colorimetry. Technical Report 2. edition, Commision Internationale de L'Eclairage (CIE), Vienna, 1986.

[2] Anonymous. Method of measuring and specifying colour rendering properties of light sources. Technical Report CIE 13.3, Commision Internationale de L'Eclairage (CIE), Vienna, 1995.

[3] W.D. Billings and R.J. Morris. Reflection of visible and infrared radiation from leaves of different ecological groups. *American Journal of Botany*, 38:327–331, 1951.

[4] D.B. Judd, D.L. MacAdam, and G.W. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54, 1964.

[5] G.J. Klinker, S.A. Shafer, and T. Kanada. A physical approach to color image understanding. *International Journal of Computer Vision*, 1(4):7–38, 1990.

[6] H.-C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A*, 10(3):1694–1699, 1986.

[7] H.-C Lee. Illuminant color from shading. In M.H. Brill, editor, *Percceiving, Measuring and Using Color*, 1250, pages 236–244. SPIE, Aug. 1990.

[8] H.-C Lee, E.J. Breneman, and C.P. Schulte. Modeling light reflection for color computer vision. *IEEE trans. on Pattern Analysis and Machine Intelligence, PAMI*, 12(4):402–409, 1990.

[9] S.J. Maas and J.R. Dunlap. Reflectance, transmittance, and absorptance of light by normal, etiolated, and albino corn leaves. *Agron, J*, 81:105–110, 1989.

[10] B. A. Maxwell and S. A. Shafer. Physics-based segmentation of complex objects using multiple hypotheses of image formation. *Computer Vision and Image Understanding*, 65(2):265–295, February 1997.

[11] C.L. Novak and S.A. Shafer. Method for estimating scene parameters from color histograms. *Journal of the Optical Society of America A*, 11(11):3020–3036, 1994.

[12] F. Pla, F. Ferri, and M. Vicens. Colour segmentation based on a light reflection model to locate citrus for robotic harvesting. *Computers and Electronics in Agriculture*, 9:53–70, 1993.

[13] S.A. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985.

[14] C.A. Shull. A spectrophotometric study of plant reflection of light from leaf surfaces. *BOT. Gaz.*, 87:583–607, 1929.

[15] S. Tominaga. Dichromatic reflection models for a variety of materials. *COLOR research and application*, 19(4):277–285, 1994.

[16] S. Tominaga. Dichromatic reflection models for rendering object surfaces. *Journal of Imaging Science and Technology*, 40(6):549–555, 1996.

[17] S. Tominaga and B. A. Wandell. Standard surface-reflectance model and illuminant estimation. *Journal of the Optical Society of America A*, 6(4):576–584, April 1989.

[18] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of Optical Society of America*, 57:1105–1114, 1967.

[19] E.A. Walter-Shea and J.M Norman. *Photon-Vegetation Interactions*, chapter Leaf Optical Properties, pages 230–250. Springer-Verlag, 1991.

[20] J.T. Wooley. Reflectance and transmittance of light by leaves. *Plant Physiology*, 47:656–662, 1971.

Chapter 6

A model based, daylight and chroma adaptive segmentation method

A model based, daylight and chroma adaptive segmentation method

H.J. Andersen[a], E. Granum and C.M. Onyango[b]

a, Laboratory for Image Analysis, Aalborg University 9220 Aalborg, Denmark
b, Image Analysis and Control Group, Silsoe Research Institute,
Wrest Park, Silsoe, Bedford, UK, MK45 4HS

Abstract

An image segmentation method based on the dichromatic reflection model, is introduced. To adapt to changing illumination conditions the image formation process is modeled by the camera characteristics, the reflectance of the object of interest, and the CIE daylight standard. A priori, loci for the body and surface reflection for the object of interest is modeled according to changes of the illumination by CIE daylight standard. The two loci is approximated by two lines and the plane defined by these is used initially for segmentation. In the case of two objects, the image is segmented by the plane which is rotated about the surface locus to minimize Wilks λ.
The method is used for segmenting four images ranging in correlated colour temperature from 5200K to 11500K. To assess its performance the four images were manually segmented into three classes: vegetation, background, and an uncertain class. The method adapted to the changing light condition with total errors ranging from 3% to 12% and higher error rates being in the images with the largest uncertain group. The method was also compared with Bayes minimax criteria for finding the "best" rotation from which it deviated by only 0.8% on average. Published in: EOS/SPIE Conference on Polarisation and Colour Techniques in Industrial Inspection, Munich, pages 136-147, June, 1999.

Published in: EOS/SPIE Conference on Polarisation and Colour Techniques in Industrial Inspection, Munich, pages 136-147, June, 1999.

6.1 Introduction

A problem often limiting or causing colour image segmentation methods to fail is the lack of ability to handle images acquired under varying illumination conditions. Further, due to the absence of satisfactory colour constancy methods [5] it is important to address the problem of developing methods that may adapt to varying daylight conditions. In recent years, it has appeared that some "difficult" machine vision applications may be solved using reflection models. For analysis of colour images the dichromatic reflection model [15] has been used for image segmentation [8, 14, 12], analysis of highlights [9], estimating scene properties [13] and more recently for generation of computer graphics rendering [20].

In this paper the dichromatic reflection model and the neutral interface assumption are used to develop an image segmentation method with the following three features, its approximate invariant to changes in the *correlated colour temperature* (CCT) of daylight and it can separate objects with overlapping clusters in the "traditional" red and green chromaticity plane. For solving the problem of overlapping clusters a third dimension the intensity, is introduced. This makes it possible to separate clusters with different albedo and second order reflection due to e.g. transmitted-reflected light.

The method is constrained to segmentation of two objects in the intensity, red and green chromaticity space (rgI-space). For adaptation to the current CCT of the daylight is location of a plane spanned by the object of interest body and surface reflection loci in rgI-space a priori modeled. The plane will be approximate invariant to changes of the CCT and is contrained to pass through the surface reflection locus about which it may be rotated to find an optimal position for separation of the objects clusters. At every rotation Wilks λ is calculated and will be at its minimum when the reflection clusters are best separated.

The method is especially suited for segmentation where the reflectance characteristic for one of the objects is well known but the "background" may vary. It is used for segmentation of vegetation and soil acquired in changing daylight conditions a problem important for development of robust computer vision support within weed control [18, 17]. It is able to adapt to the current daylight condition and it gives segmentation results which will aid future development of computer vision support within the area.

The paper is outlined as follows. Section 2, first review the dichromatic reflection model followed by its representation in the rgI-space and introduction of suitable geometrical scalings factors. Next, spectral variation of the illumination together with derivation of the direction of a plane approximate invariant to such changes, are presented. Section 3, constrain the segmentation to two objects vegetation and soil. How their reflection clusters potentially may overlap, is illustrated, and Wilks λ as criteria for finding the position of the discrimination plane which best separates the clusters, is presented. Section 4, introduces the experimental data. The results are presented in section 5 and discussed and concluded in section 6 and 7.

6.2 Modelling

6.2.1 The Dichromatic Reflection Model

The reflected light from an object $L(\Theta, \lambda)$ is a function of wavelength λ and the photometric angles Θ which includes the illumination direction angle (i), the viewing angle (e), and the phase angle (g). The dichromatic reflection model as formulated by Shafer[15], describes the reflected light from a non homogeneous dielectrical object as the sum of its body and surface reflections:

$$L(\Theta, \lambda) = m_B(\Theta)L_B(\lambda) + m_S(\Theta)L_S(\lambda) \tag{6.1}$$

where the terms $L_B(\lambda)$ and $L_S(\lambda)$ are the spectral power distribution of the surface and body reflection and $m_B(\Theta)$ and $m_S(\Theta)$ geometrical scaling factors for the body and surface reflection, respectively. The

model assumes that the spectral power distribution and the geometrical scaling factors are separable. This condition is approximately satisfied by many materials having high contents of water or oil, because their index of refraction is nearly constant over the visible spectrum [10].

The dichromatic reflection model as described in eq 6.1 covers the whole of the continuous light spectrum. This dimension of the spectrum is in a camera reduced by the process of *spectral integration*:

$$C_f = \int L(\Theta, \lambda)\tau_f(\lambda)s(\lambda)d\lambda \tag{6.2}$$

where $L(\Theta, \lambda)$ is the incoming light in the camera, $\tau_f(\lambda)$ the spectral transmittance of filter f which limits the integration to a sub band and $s(\lambda)$ the sensitivity of the camera.

In a standard RGB-camera the colour spectrum is represented by three sub bands, the three dimensional colour space, such that a pixel at location (x, y) in the image will be represented in colour space by a vector $\mathbf{C}(x, y) = [C_{red}, C_{green}, C_{blue}]$, which for convenience may be refereed to as $[Red, Green, Blue]$. Hereby the dichromatic reflection model in colour space may be expressed by:

$$\mathbf{C}(x, y) = m_B(\Theta)\mathbf{C}_B + m_S(\Theta)\mathbf{C}_S \tag{6.3}$$

which describes the so called *dichromatic plane* in the Red, Green, and Blue space spanned by the body and surface reflection vectors, $\mathbf{C}_B$ and $\mathbf{C}_S$ of the object, respectively. Object pixel points will be distributed in this plane

The reflected light from the body and surface components $L_{(B,S)}(\Theta, \lambda)$ is formed by the incident light $E(\lambda)$ weighted with the reflectance $\rho(\lambda)$ of the body and surface, respectively, for every wavelength. Hence, the dichromatic model may be expressed as:

$$\mathbf{C}(x, y) = m_B(\Theta)\int E(\Theta, \lambda)\rho_B(\lambda)\tau_f(\lambda)s(\lambda)d\lambda + m_S(\Theta)\int E(\Theta, \lambda)\rho_S(\lambda)\tau_f(\lambda)s(\lambda)d\lambda \tag{6.4}$$

The standard dichromatic reflection model equation (6.1) also incorporates an assumption called the Neutral Interface Reflection (NIR) [10], which assumes that the relative spectral composition of the surface reflection component is identical to that of the light source (i.e. the surface reflectance $\rho_S(\lambda)$ is constant). More recently Tominaga, [19] has categorized this as the dichromatic reflection model of type I, which has been shown experimentally to be valid for a variety of materials including, plastics, ceramics, tiles, fruits, leaves and others [10, 19, 14]. Alternatively, the standard model equation (6.1) may include a transfer function of the interface ((i.e. the surface reflectance $\rho_S(\lambda)$ is not constant) which Tominaga[19] categorises as type II.

6.2.2 Representation of the Dichromatic Reflection Model in rgI-space

In the following a representation of the reflected light from an object by its red and green chromaticities together with the intensity will be introduced. This space will be denoted as the rgI-space.

For analysis of the colour of the reflected light independent of intensity, it is convenient to transform the colour vector $\mathbf{C}$ to its corresponding chromaticities which may be done by taking the first norm:

$$\mathbf{c}(x, y) = |\mathbf{C}(x, y)| = |m_B(\Theta)\mathbf{C}_B + m_S(\Theta)\mathbf{C}_S| \tag{6.5}$$

which may be expressed as:

$$\mathbf{c}(x, y) = \frac{1}{S_B + \frac{m_S(\Theta)}{m_B(\Theta)}S_S}\mathbf{C}_B + \frac{1}{S_S + \frac{m_B(\Theta)}{m_S(\Theta)}S_B}\mathbf{C}_S \tag{6.6}$$

where S_B and S_S is the sum of $\mathbf{C}_B$ and $\mathbf{C}_S$, respectively. As the sum of $\mathbf{c}(x,y) = 1$ it is sufficient to represent the chromaticity vector by only two components. In the following the red and green chromaticity will be used giving the *rg-chromaticity plane*. Further, if $m_S \to \infty$ or $m_B = 0$ then $\mathbf{c}(x,y) = \frac{1}{S_S}\mathbf{C}_S$ and if $m_B \to \infty$ or $m_B = 0$ then $\mathbf{c}(x,y) = \frac{1}{S_B}\mathbf{C}_B$. Hence, eq. 6.6 defines a line segment between the chromaticities of the reflection from an object due to only body or surface reflection along which its pixel points will be distributed. These chromaticities will be denoted as the *primary chromaticies* $\mathbf{c}_B$ and $\mathbf{c}_S$ of the object, as suggested in [1].

The intensity of the colour vector $\mathbf{C}(x,y)$ is defined by the Euclidean length:

$$I(x,y) = \sqrt{\mathbf{C}(x,y)'\mathbf{C}(x,y)} \tag{6.7}$$

By these two transformations the colour vector $\mathbf{C}$ may be transformed into its red and green chromaticity and intensity. To be able to model the chromaticities and intensity it is necessary to find suitable expressions for the geometrical scaling factors m_B and m_S, respectively.

The scaling factor for the body reflection m_B is modelled by Lambert's law which expresses the reflected light as a function of the incident angle i, as:

$$m_B(\Theta) = I_E g A cos(i) \tag{6.8}$$

where I_E is the intensity of the illumination, g the gain of the camera, and A the albedo of the object. Though, newer and more accurate body reflection models have been proposed by Wolf et al [24] Lamberts law will be use in this study because of its simplicity. Also, as discussed in [24] the model will be a sufficient approximation for many computer vision purposes (as an example it is accurate within a margin of $\pm$ 5% when i and e are simultaneously less than 50°).

The geometrical scaling factor for the surface reflectance m_S is modelled according to Torrance and Sparrow [21] with the same modification as in [13]:

$$m_S(\Theta) = I_E g \frac{FG(i,e,g)C}{cos(e)\sigma} e^{-\frac{\alpha^2}{\sigma^2}} \tag{6.9}$$

where I_E is the intensity of the illumination, g the gain of the camera. i, e, and g are the photometric angles and α is the off-specular angle which may be expressed by i, e, and g. C is a constant that includes the facet size (a variable in the original Torrance-Sparrow model). G is an attenuation factor that depends upon photometric angles and which comes into effect at grazing angles. It is a complicated function depending on the photometric angles and it will not be reproduced here, see [21]. F is Fresnel coefficient which over the visible spectrum may be assumed constant for materials with a high water or oil content [10] with a refractive index of 1.5, corresponding to about 4% Fresnel reflectance.

The facet size is assumed constant and larger than the wavelength of visible light, but too small to be seen as texture. The distribution of the facet normals is modelled as Gaussian with σ describing the surface roughness. Smooth surfaces will have a small value of σ whereas rougher surfaces will have a larger value.

By equation 6.6, 6.7, 6.8, and 6.9 it is now possible to model the distribution of the chromaticity and intensity at different values of the optical roughness σ and albedo A. In eq. 6.9 it may be noted that the numerator $cos(e)$ will approach 0 if $e \to \pm\frac{\pi}{2}$. According, to Torrance and Sparrow this will give a finite limit [21].

The modelling will be constrained to involve a uniform illumination source. The photometric angles is varied by changing the direction of the normal vectors' azimuth angle ϕ from $0 - 2\pi$ and the elevation angle β from $0 - \pi$. For simplicity the observer is kept perpendicular to the xy-plane and all constants, I_E and g in eq. 6.8 and 6.9 and F and C in eq. 6.9, are set to one. Figure 6.2 shows the result of such

modelling for 6 different combinations of σ and the albedo A. From the figure it is evident that the distribution of the pixel points is primarily determined by the optical roughness σ.

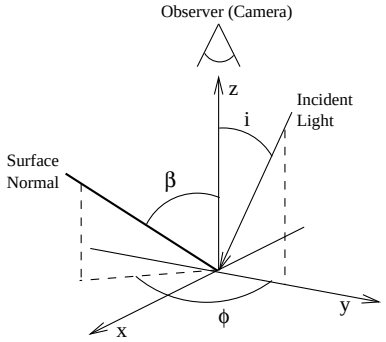


Figure 6.1: *Definition of azimuth and elevation angle ϕ and β, respectively, for the normal vector. Together with direction of the incident light and the observer perpendicular to the xy-plane.*

An important property illustrated also in figure 6.2 is that a line segment located at the primary chromaticity of the surface reflection at maximum intensity may approximately limit the dispersion of its pixels points towards the surface chromaticity. Included in figure 6.2 (c) are two line segments illustrating how the dispersion may be described. The line segments together with the surface reflection line and the g-chromaticity axis define a triangle. An important feature of this area is that pixel points having the same chromaticity as the object both at lower intensity will be located in this area. Such pixel points may, for example, be points illuminated by light first passing through an object that has the same transmittance and reflectance characteristic. As evident from (a) and (b) in figure 6.2 the area of the triangle will be limited by both σ and the albedo A.

6.2.3 Spectral Variation of the Illumination

In this section variation of the illuminant E will be introduced and its influence on the position of the dichromatic plane will be investigated. Further, a transformation of the rgI-points into a space having one direction approximately invariant to changes of the CCT, one in the direction which the chromaticity changes at changing CCT and one in the intensity direction, will be introduced.

So far, the analysis has used one spectrum of the illuminant E in eq. 6.4. In the following, changes of the illumination will be constrained to the illuminant source, daylight. The spectral composition of daylight may be characterized and approximated according to its *Correlated Colour Temperature* (CCT), as formulated in CIE's standard [2] on the basis of Judd *el al* [7].

Different CCT's of the daylight will make the line segments connecting the primary chromaticities $\mathbf{c}_B$ and $\mathbf{c}_S$ change location in the chromaticity plane. As shown in [1] reflection from a dielectrical object will form loci for its body and surface reflection, respectively. Further, assuming that the surface locus will be situated at maximum intensity for a given chromaticity, the rgI-values will be distributed within the rectangle defined by $(r_B, g_B, 0)$, $(r_S, g_S, 0)$, $(r_B, g_B, \max(I(r_S, g_S))$, and $(r_S, g_S, \max(I(r_S, g_S)$ with (r_B, g_B) and (r_S, g_S) at corresponding CCT. Figure 6.3, shows the position of the primary chromaticities ($\mathbf{c}_B$ and $\mathbf{c}_S$), which define the position of the dichromatic plane in the rg-plane and rgI-space, respectively. The points forming the body and surface reflection loci will be gathered in the matrices $\mathbf{\Lambda}$ and $\mathbf{\Delta}$, respectively.

From eq. 6.4 it is evident that for an object with constant reflectance the only source of variability of the colour vector $\mathbf{C}$ will be due to changes in spectral composition of the illuminant E. Hence, the normal vector to the plane defined by the two lines approximating the body and surface loci will be orientated

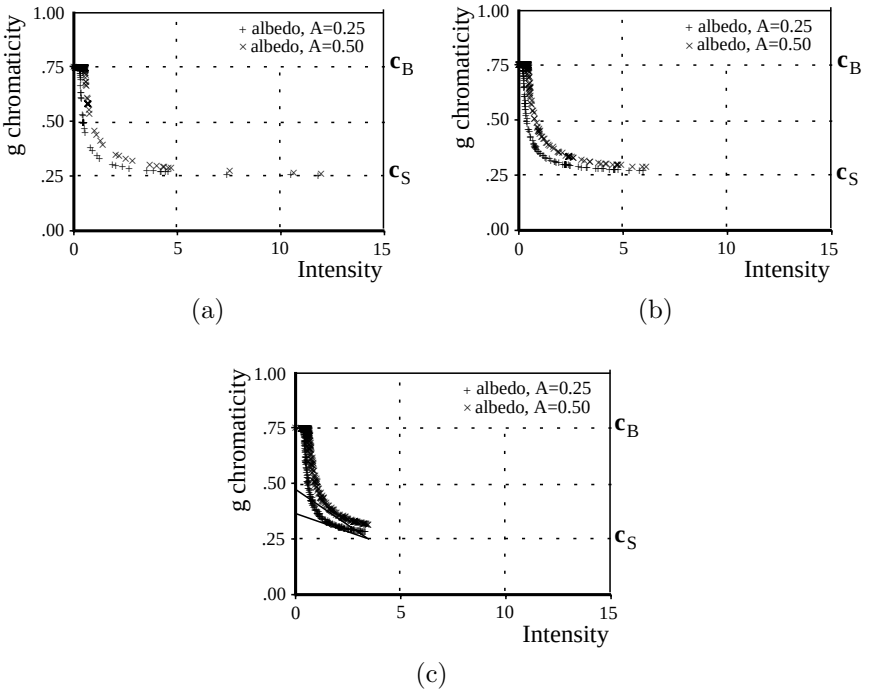


Figure 6.2: *Modelling of pixel distributions for a dielectrical object with chromaticity vectors* $c_B = [0.1, 0.75, 0.15]$ *and* $c_S = [0.1, 0.25, 0.65]$*, respectively. With* σ *equal* $5°$ *in a,* $10°$ *in b, and* $20°$ *in c. As the red chromaticity is not changing only the green is considered. Included in (c) are the line segments limiting the dispersion of the pixel points for the two values of the albedo.*

in a direction approximately invariant to changes in the CCT of the daylight. The normal vector may be found by first decomposing the covariance of $\mathbf{\Lambda}$ and $\mathbf{\Delta}$ into its eigenvectors, $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and choosing the eigenvector with the highest eigenvalue whereby the direction of it may be obtained by:

$$\mathbf{n} = \mathbf{e}_B \times \mathbf{e}_S \tag{6.10}$$

where $\mathbf{e}_B$ and $\mathbf{e}_S$ is the eigenvector of the surface and body loci with highest eigenvalues. $\mathbf{e}_B$, describes the direction in which the dichromatic plane will move for changing CCT, it will be chosen as the second direction. Finally, the third direction $\mathbf{a}$ is obtained from:

$$\mathbf{a} = \mathbf{e}_B \times \mathbf{n} \tag{6.11}$$

Further, $\mathbf{a}$ may be rotated about $\mathbf{e}_B$ so it becomes parallel to the intensity axis I and hence corresponds to the intensity. The rotation angle δ may be found by the following criteria:

$$\mathbf{a} \cdot I = 0 \Rightarrow cos(\delta) = \mathbf{a} \cdot I \Rightarrow \delta = acos(-\mathbf{a} \cdot I) \tag{6.12}$$

Hence, by projecting the pixel rgI-values onto $\mathbf{n}$ and $\mathbf{a}$ they may be represented in a plane approximately invariant to changes in CCT. The plane can be rotated so that $\mathbf{a}$ corresponds to the intensity I and $\mathbf{n}$ is a linear combination of the chromaticities.

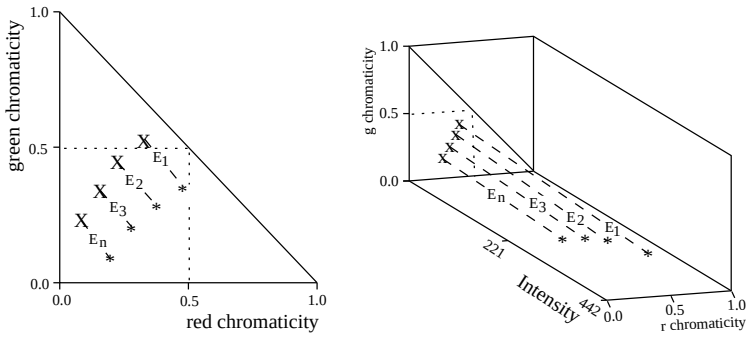


Figure 6.3: *Position of primary chromaticities and projection of dichromatic plane. (a), alternative positions of dichromatic plane in chromaticity plane due to varying CCT of the daylight $E_1 \dots E_n$. (b), alternative positions of the primary chromaticities of the body and surface loci in the rgI-space due to varying CCT's of the daylight $E_1 \dots E_n$. Points on surface locus are placed at maximum intensity for that particular chromaticity. "X", indicates location on body locus. "*", indicates location on surface locus. Dashed lines connect points with corresponding CCTs on the body and surface loci, respectively.*

6.3 Segmentation in a two object situation

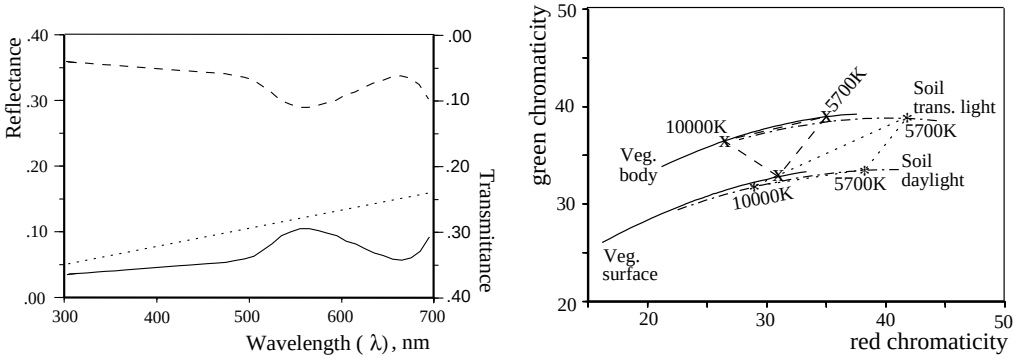
In the preceding section a description of a dichromatic plane in the rgI-space together with a projection of it onto a plane approximately invariant to changes of the CCT's, was introduced. In this section the obtained result will be used for development of a segmentation method based on a linear discriminant function.

6.3.1 Modelling reflection

For modelling of the reflection from natural scenes of vegetation and soil, four reflection clusters in differing illumination conditions, will be considered. Figure 5.14 shows four images of vegetation acquired under different illumination conditions. Images 5.14(a) and 5.14(b) are illuminated by diffuse skylight and will give two primary reflection clusters one for vegetation and one for soil. Images 5.14(c) and 5.14(d) are illuminated by uniform sunlight and diffuse skylight, potentially mixed with light transmitted through the vegetation. This will give four reflections clusters, vegetation and soil in sunlight and vegetation and soil in sky and transmitted light. Also, inter reflections within the vegetation may occur but this will not be considered as it will not cause any problems for the proposed method. For modelling of the reflections, suitable reflectance and transmittance characteristics for vegetation and soil will be obtained from the literature. The camera tristimuli characteristic is modelled according to information provided by JAI for their M90 3CCD colour camera. The camera was white balanced at 5000K.

The body reflectance ($\mathbf{C}_B$) of vegetation is derived from the reflectance characteristic in [11] for maize under normal growth condition, and this is in good agreement with characteristics for vegetation (typically crops and weeds), reported elsewhere in the literature [16, 4, 25, 23]. The transmittance of vegetation will have a characteristic very similar to the reflectance [25, 23], so in this study the same will be used for the two. The two characteristics are included in figure 6.4 (a). As vegetation may be assumed to follow the NIR model [10, 19, 14] the surface reflection will have the same spectral composition as that of the illuminant, hence it will correspond to the daylight locus of the camera.

The background will in this study be assumed to consist of soil. Reflectance characteristics of soil are are not as uniform as vegetation. Generally, they will vary from constant (grayish soil) to slightly increasing



(a) (b)

Figure 6.4: *(a), solid line, reflectance characteristic used for modelling the body reflection of vegetation (derived from [11]). Dashed line, transmittance characteristic of vegetation. Dotted line, characteristic for body reflection of soil. (b), reflection loci for vegetation (Veg.) and soil. Solid line, surface and body reflection loci for vegetation. Dashed and dotted line, reflection loci of soil due to illumination sources daylight and light transmitted through the leaves. All loci vary in CCT from 5000 to 25000 Kelvin from right to left. Triangles enclosed by dotted and dashed lines show areas for location of soil and plant pixel chromaticities with daylight having a CCT of 5700K in direct sunlit areas and 10000K in shaded areas, respectively.*

with wavelength (reddish soil) according to characteristics given in [6]. However for modelling, the "general" reflectance characteristic as introduced in [3], will be used (Figure 6.4 (a)). It is important to notice that the range of characteristics as given above will give soil loci situated on the opposite side of the daylight locus to the predominantly green reflection from the vegetation. Due to the structure of soils, no surface reflection will be considered from it.

Figure 6.4 (b) illustrates the location of the plant body and surface reflection loci together with the body reflection loci for soil in direct daylight due to reflection from light first being transmitted through the plant leaves. From figure 6.4 (b) it is evident that the body reflection for soil is very close to the surface reflection from the vegetation and further that pixel points situated in shade may, depending on the ratio between skylight and transmitted light, be mixed with the reflection from vegetation. Further, the projection of the pixel points onto **n**, may also cause the pixel points to be mixed. However, an important feature of the pixel points in shade is that their intensity will be low. Hence, they may be expected to be located in the area between the surface reflection loci and the pixel points of the vegetation, as introduced in section 6.2.2.

6.3.2 Finding the discriminant function

From the above it is now possible to find a discriminant function in the **na**-plane based on the physical modelling of the reflections. Under the neutral interface assumption the dichromatic plane will transverse from its primary body chromaticity towards its surface chromaticity. Hence, a linear discriminant function should be constrained to pass through the surface locus at the current CCT.

Figure 6.5, illustrates projection of an image onto the $\mathbf{na} - plane$. In the first row with non overlapping clusters a segmentation problem that may be solved in rg-chromaticity plane. In the second row with overlapping clusters and illustration of how the discriminant limit may adapt to the current conditions in the $\mathbf{na} - plane$. Note, that the situation with non overlapping may also be solve in the $\mathbf{na} - plane$.

Strictly, the discriminant function should pass through the projection of the surface locus at the corresponding CCT. However, as the majority of the pixels are expected to be due to body reflection this

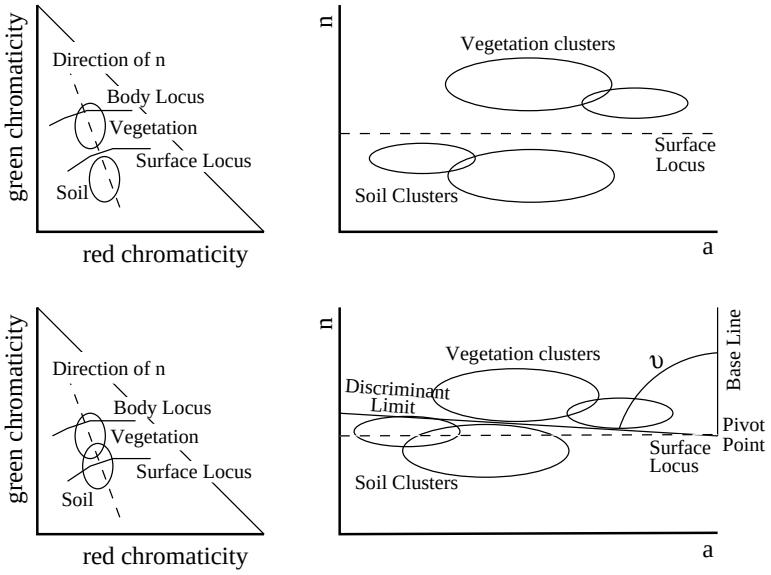


Figure 6.5: *Illustration of non overlapping (1. row) and overlapping (2. row) reflection clusters. Included is definition of rotation angle υ of the dicriminant limit. Large clusters illustrates body reflection, small potential surface and second order reflections, respectively.*

constraint will be relaxed so the function passes through the surface reflection locus at the white balance of the camera. This will also give the advantage that all pixel points will be on one side of the pivot point and hence the discriminant function will pass all points when rotated from 0° to 180°, as illustrated in figure 6.5.

The discriminant limit may now be rotated about the pivot point on the surface reflection locus and for every rotation Wilks λ may be calculated as:

$$\lambda = \frac{|W|}{|T|} \tag{6.13}$$

where $|W|$ and $|T|$ are the determinants of the pooled within variance and the total variance, respectively. The pooled within variance is obtained by pooling the variances on either side of the discriminant limit as:

$$W = (n_{[0;\upsilon]} - 1)W_{[0;\upsilon]} + (n_{]\upsilon;180]} - 1)W_{]\upsilon;180]} \tag{6.14}$$

Wilks λ will be at its minimum when the two clusters vegetation and background are best separated. Hence, it may be used for determining the optimal angle of the discriminant. When rotating the discriminant it is expected, from the modeling, that the vegetation cluster will be clockwise from the limit and hence the discriminant limit should be located in the interval from 0° to 90°.

Before calculation of Wilks λ the projection of the rgI-values onto $\mathbf{n}$, $\mathbf{e}_B$, and $\mathbf{a}$ where standardized. To avoid the influence of pixel points with a poor signal to noise ratio the standardisation and calculation of Wilks λ was based on pixels having an intensity above 25 and all three colour channels below 250.

6.4 Experimental Data

The proposed method will be used on the four images as illustrated in figure 6.6. Image (a), rape plants and seedlings acquired in skylight with a CCT of 11500K, (b) barley plants in skylight at a CCT of 10000K, (c) barley plants in direct sunlight at a CCT of 5700K, and (d) cauliflower in sunlight at a CCT of 5200K. All images were acquired during the summer of 1998 at Silsoe Research Institute, England, with M90 3CCD colour camera from JAI. The CCT was measured with a Minolta Colour Meter II.

In order to produce ground truth, the images were carefully manually classified into live vegetation and background. Like all ground truth methods, the human observer could not be certain about all decisions hence a third class "uncertain", was introduced. The proportion of the different classes is given in table 6.1.

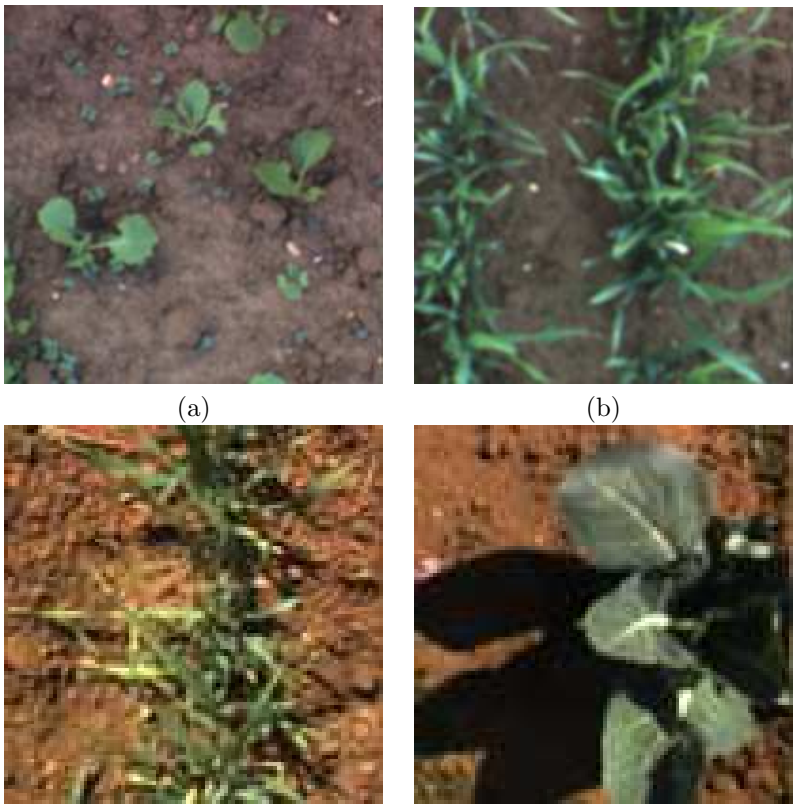


Figure 6.6: *Images used for evaluation of the proposed method.*

Table 6.1: *Proportion of classes vegetation, background, and uncertain obtained by human assessment for the images in figure 6.6.*

	Proportion in %		
Image	Vegetation	Background	Uncertain
a	8.7	87.2	4.1
b	26.2	51.8	22.1
c	19.6	71.3	9.2
d	42.7	54.5	2.8

6.5 Results

The largest eigenvalues of $\mathbf{\Lambda}$ and $\mathbf{\Delta}$ accounted for 99% and 99% of the variation, respectively, so the normal vector $\mathbf{n}$ may be regarded as invariant against changes of the CCT of the daylight.

For evaluation of the segmentation method it will be compared with Bayes minimax criteria for error receivers [22]. Varying υ from 0° to 180° with uncertain as vegetation and background, respectively, the Receiver Operating Characteristic (ROC) curve may be derived for each image. The ROC curve plots the true positive ratio (tpr) - (number of true positives)/(true positives + false negative) against the false positive ratio (fpr) - (number of false positives)/(true negative + false positive). Using the rotation angle υ as the curve parameter then it may be expected that the tpr-ratio will increase with increasing υ (i.e. the ROC curve will traverse towards (0.1) the optimal operation points) until υ becomes so large that the discriminant limit traverses into the background cluster after which the fpr-ratio also will begin to increase and hence the ROC curve will traverse towards (1,1). A criteria often used for selecting a "best" operating point on the ROC curve when the prior probabilities are unknown is the minimax criteria based on Bayes decision theory. The criteria minimizes the maximum possible risk assuming the priors are such that they make the performance as bad as possible. With costs of zero for making a correct decision and one for misclassification, the minimax criteria is equal to finding the curve parameter υ which satisfies $P_f = P_m$ [22], where P_f is the probability of a false alarm (i.e. the target is assumed present when it is not) and P_m is the probability of a miss (i.e. the target is assumed absent when it is present), respectively. The above criteria may be found on the ROC curve as the intersection between it and $1 - fpr$.

Figure 6.7 shows Wilks λ, the ROC-curves and the results of segmenting the image at the minimum of Wilks λ. Table 6.2 summarizes the values of υ for minimum of Wilks λ and the minimax criteria together with the error probabilities with the uncertain group as vegetation and background, respectively. Table 6.2, shows that Wilks λ has values at or below 90° for all images. Looking at the error rates of P_f and P_m for Wilks λ these are lower for the uncertain group as vegetation than with uncertain as background for P_f and contradictory for P_m. Further, Wilks λ divided the uncertain group into 73%, 63%, 65%, and 52% vegetation for images 6.6a-d. Hence, showing good agreement with the human segmentation. From table 6.1 and 6.2 it is also evident that the magnitude of the error rate follows the proportion of the uncertain class with image 6.6(a) and 6.6(d) having rates below 10% and image 6.6(b) and 6.6(c) slightly higher rates about 10%.

Also, comparing with Bayes minimax criteria gives that the rotation angle υ for the discriminant limit for images 6.6a-c are placed in between the angles obtained with the uncertain group as background and vegetation. For image 6.6(d) this is not the case, however, the angle only differs one degree for being at the limit of υ_2. In average the error rate P_{total} obtained by Wilks λ deviated 0.8% from the minimax criteria.

The segmented images using υ at the minimum of Wilks λ is illustrated in the third column of figure 6.7. Looking at images 6.6(a) and 6.6(b) the errors in the segmentation result may be regarded as being misclassification of border pixels for both and for 6.6(b) the areas in shade within the barley plants. For images 6.6(c) and 6.6(d) the main proportion of the errors may be regarded as being from the areas

in shade where the signal to noise ratio is poor. However, these errors give small misclassified areas clustered in shade regions which with use of further image analysis methods may be correctly classified into vegetation and background.

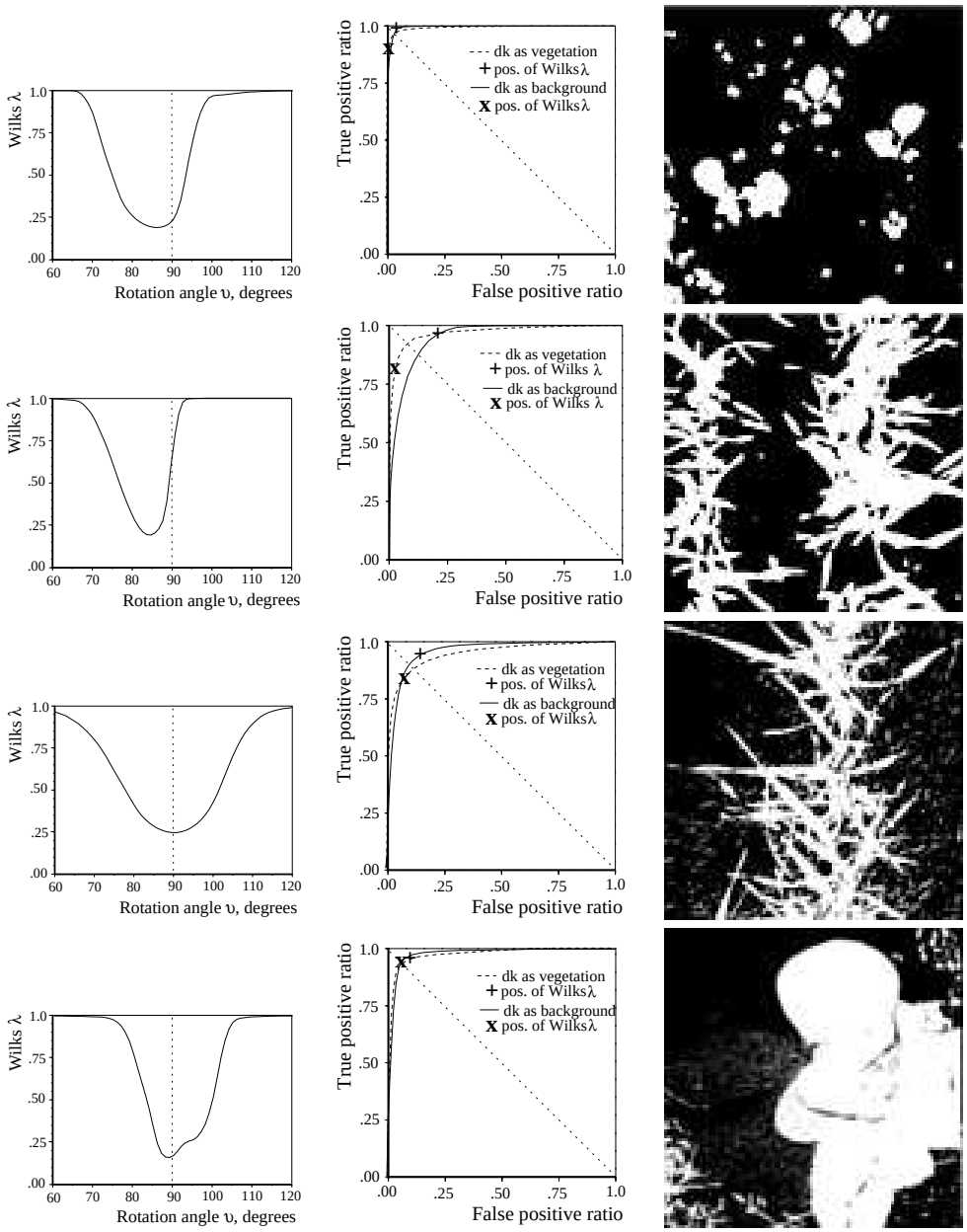


Figure 6.7: *First column, Wilks λ plottet against rotation angle υ, Second column, ROC-curves with υ as curve parameter with the uncertain group as vegetation or background, respectively. Labels $+$ and* $\mathbf{x}$ *indicates position of υ at minimum of Wilks λ. Third column, segmented images at minimum of Wilks λ.*

Table 6.2: *Probabilities at minimum of Wilks λ and Bayes minimax criteria, with uncertain group as background or vegetation, respectively. υ, rotation angle at minimum of Wilks λ. υ_1, rotation angle at minimum of Bayes minimax criteria with uncertain as background. υ_2 rotation angle at minimum of Bayes minimax criteria with uncertain as vegetation. pct, percentage of pixel points used for calculation of Wilks λ. P_f, probability of a false alarm i.e. vegetation pixel point classified as background. P_m, probability of a miss i.e. background pixel point classified as vegetation. P_{total}, the total probability of an error, $((P_f + P_m)/2)$.*

							Probabilities in %					
							uncertain as background			uncertain as vegetation		
Image	Method	υ	υ_1	υ_2	λ	pct	P_f	P_m	P_{total}	P_f	P_m	P_{total}
a	Wilks λ	86			0.19	99.9	4.1	1.7	3.0	0.9	10.3	5.6
	Minimax		85	89			3.5	3.1	3.3	2.8	3.9	3.3
b	Wilks λ	85			0.19	99.7	20.6	3.9	12.3	2.5	19.1	10.8
	Minimax		83	87			13.7	10.4	12.1	6.3	9.2	7.7
c	Wilks λ	90			0.24	98.7	14.9	6.0	10.5	8.6	15.4	11.6
	Minimax		87	92			10.3	10.1	10.2	12.4	11.8	12.1
d	Wilks λ	89			0.16	86.0	10.9	2.5	6.7	8.8	5.4	7.1
	Minimax		87	88			5.3	7.7	6.5	5.5	7.5	6.5

6.6 Discussion

Using the dichromatic reflection model and the neutral interface assumption a method for segmentation of images consistent of vegetation and soil illuminated by daylight, is introduced. The method relies on a modelling of the image formation to find a direction invariant to changes in CCT of the daylight. To perform this *a priori* analysis the tristimuli characteristics of the specific camera and the reflectance of the objects should be known. The information from JAI appears to be sufficient despite being a general product specification. Also, the reflectance characteristic of maize appears to be an adequate representative for modelling the reflectance of the vegetation used in the experiment.

From the results it may be concluded that the proposed method may adapt to images acquired at different CCT's of the daylight and location of reflection clusters. Wilks λ gives a distinct minimum even in images 6.6(c) and 6.6(d) that strictly consist of four reflection clusters. However, these clusters will be mixed and they will appear as one, under the constraint that the discriminant limit shall pass through the surface reflection locus clearly limits the possibility of separating them. Strictly, the discriminant limit should pass through the surface reflection locus at maximum intensity of the chromaticity at the current CCT. Looking at the error rates, the ROC curves and the segmented images, the relaxation of the constraint that the discriminant line must pass through the white balance of the camera, does not seem to give rise to increased errors (i.e. image 6.6(a) is the image for which the relaxation should have the largest impact).

The values of the rotation angle v are all at or below 90° which in general is in accordance with what is expected. The smallest value of v occurs for image 6.6(a) and 6.6(b), images having primary body reflection. For image 6.6(c) and 6.6(d) on the other hand smaller values of v might have been expected. Although, the shaded areas generally are classifying correct the impact of the constrain that the discriminant limit is passing through the white balance of the camera is an issue for further investigation, especially for images acquired in sunlit condition.

In this study the background has been constrained to consist of soil. However, in realistic scenes of vegetation there may be other objects as well (i.e. stones, plant residue, straw etc.). For further refinement of the method, the impact of such objects on it should be investigated.

6.7 Conclusion

A segmentation method based on the dichromatic reflection model and modelling of the image formation process, is introduced. The method is based on finding the minimum of Wilks λ based on the variances on either side of a discriminant line constrained to pass through a point on the surface reflection locus which is determined a priori. It is used for segmenting images of vegetation acquired in daylight ranging in correlated colour temperature (CCT) from 5200K to 11500K. To obtain "ground truth" the images used for evaluation were manually segmented into three classes vegetation, background and uncertain.

The results show that the method is able to adapt to the changing CCT of the daylight and position of reflection clusters. Total error rates obtained by the method ranged from 3% to 12% with the higher rates being in the images with the largest uncertain group. Also, the error rates were compared with the threshold obtained with Bayes minimax criteria for rotation of the discriminant limit where it deviated on average by 0.8%. Further, the method shows the desirable feature of dividing the uncertain class into vegetation and background. In general the results may be regarded as good and promising for future development of colour vision segmentation methods, particularly in the detection of specific plant structures in outdoor illumination conditions.

References

[1] H.J. Andersen and E Granum. Classifying illumination condition from two light sources by colour histogram assessment. *Submitted for: Journal of Optical Society of America A*, March 1999.

[2] Anonymous. Colorimetry. Technical Report 2. edition, Commision Internationale de L'Eclairage (CIE), 1986.

[3] H.R. Biller. Reduced input of herbicides by use of optoelectronic sensors. *Journal of Agricultural Engineering Research*, 71:357–362, 1998.

[4] W.D. Billings and R.J. Morris. Reflection of visible and infrared radiation from leaves of different ecological groups. *American Journal of Botany*, 38:327–331, 1951.

[5] B. Funt, K. Barnard, and L. Martin. Is machine vision good enough ? In H. Buckhardt and B. Neumann, editors, *5th European Confernce on Computer Vision*, volume 1, pages 445–459. ECCV, 1998.

[6] A.W. Hooper, G.O. Harries, and B. Ambler. A photoelectric sensor for distinguishing between plant material and soil. *Journal of Agricultural Engineering Research*, (21):145–155, 1976.

[7] D.B. Judd, D.L. MacAdam, and G.W. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54, 1964.

[8] G.J. Klinker, S.A. Shafer, and T. Kanada. A physical approach to color image understanding. *International Journal of Computer Vision*, 1(4):7–38, 1990.

[9] H.-C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A*, 10(3):1694–1699, 1986.

[10] H.-C Lee, E.J. Breneman, and C.P. Schulte. Modeling light reflection for color computer vision. *IEEE trans. on Pattern Analysis and Machine Intelligence, PAMI*, 12(4):402–409, 1990.

[11] S.J. Maas and J.R. Dunlap. Reflectance, transmittance, and absorptance of light by normal, etiolated, and albino corn leaves. *Agron, J*, 81:105–110, 1989.

[12] B. Maxwell. *Segmentation and Interpretation Using Multiple Physical Hypotheses og Image Formation.* PhD thesis, The Robitics Institute, Carnegie Mellon Unversity, Pittsburgh, Pennsylvania, 1996.

[13] C.L. Novak and S.A. Shafer. Method for estimating scene parameters from color histograms. *Journal of the Optical Society of America A*, 11(11):3020–3036, 1994.

[14] F. Pla, F. Ferri, and M. Vicens. Colour segmentation based on a light reflection model to locate citrus for robotic harvesting. *Computers and Electronics in Agriculture*, 9:53–70, 1993.

[15] S.A. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985.

[16] C.A. Shull. A spectrophotometric study of plant reflection of light from leaf surfaces. *BOT. Gaz.*, 87:583–607, 1929.

[17] B. Southall, J.A. Marchant, T. Hague, and B.F. Buxton. Model based tracking for navigation and segmentation. In *5th European Conference on Computer Vision Freiburg, Germany*, volume 1. ECCV, 1998.

[18] L. Tian. *Knowledge Based Machine Vision System for Outdoor Plant Identification.* PhD thesis, University of California, Davis, Dept. of Biological and Agricultural Engineering, 1995.

[19] S. Tominaga. Dichromatic reflection models for a variety of materials. *COLOR research and application*, 19(4):277–285, 1994.

[20] S. Tominaga. Dichromatic reflection models for rendering object surfaces. *Journal of Imaging Science and Technology*, 40(6):549–555, 1996.

[21] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of Optical Society of America*, 57:1105–1114, 1967.

[22] H.L. van Trees. *Detection, estimation, and modulation theory*, volume 1. John Wiley & Sons, 1968.

[23] E.A. Walter-Shea and J.M Norman. *Photon-Vegetation Interactions*, chapter Leaf Optical Properties, pages 230–250. Springer-Verlag, 1991.

[24] L.B. Wolf, S.K. Nayar, and M. Oren. Improved diffuse reflection models for computer vision. *International Journal of Computer Vision*, 30(1):55–77, 1998.

[25] J.T. Wooley. Reflectance and transmittance of light by leaves. *Plant Physiology*, 47:656–662, 1971.

Chapter 7

The Design and Operation of an Imaging Sensor for Detecting Vegetation

The Design and Operation of an Imaging Sensor for Detecting Vegetation

H.J. Andersen[a], C.M. Onyango[b] and J.A. Marchant[b]

a, Laboratory for Image Analysis, Aalborg University 9220 Aalborg, Denmark
b, Image Analysis and Control Group, Silsoe Research Institute,
Wrest Park, Silsoe, Bedford, UK, MK45 4HS

Abstract

There is a need to sense vegetation from ground-based vehicles so that plants can be treated in a selective way, thus saving on crop treatment measures. This paper introduces a sensor for detecting vegetation under natural illumination that uses three filters, red, green, and near infra-red (NIR), with a monochrome Charge Couple Camera (CCD) camera. The sensor design and the data handling are based on the physics of illumination, reflection from the vegetation, transmission through the filters, and interception at the CCD. In order to model the spectral characteristics of the daylight in the NIR we extend an existing standard using a black body model. We derive suitable filters, develop a methodology for balancing the sensitivity of each channel, and collect image data for a range of illumination conditions and two crop types. We present results showing that the sensor behaves as we predict. We also show that clusters form in a measurement space consisting of the red and NIR chromaticities in accordance with their expected position and shape. Presentation in this space gives a good separation of the vegetation and non vegetation clusters which will be suitable for physically based classification methods to be developed in future work.

Published in: International Journal of Imaging System and Technology,
Vol. 11, pages 144-151, 2000

7.1 Introduction

In recent decades, farming in the Western world has come increasingly relied on chemical crop protection. Lately, this heavy use of chemicals has provoked demands from consumers and environmentalists for reduction. A potential way of meeting these demands is to use precision techniques for various types of agricultural operations, in particular the application of a variety of chemicals so that they are placed where they will have an optimal effect with minimum quantity. Various authors [17, 20, 15] have proposed that optical sensing, which is used to detect vegetation from ground-based vehicles, could help to target crop treatments. Furthermore, if plants could be differentiated from weeds, chemicals could be directed only at the target species or eliminated altogether by precise mechanical methods.

Although it may be possible to differentiate plants from weeds on the basis of spectral content alone in some restricted circumstances [24], it is unlikely that there will be enough information for differentiation over a workable range of conditions. On the other hand, it may be possible to differentiate on the basis of appearance [8, 19, 5], this is after all how humans do it. In any method based on appearance, an image must be formed. Therefore, this work concentrates on CCD cameras as the basic sensing method. The idea of forming images of agricultural crops from a tractor-type vehicle, dividing the image into its various components, and treating accordingly is an attractive one. However, the analysis of images in natural environments is difficult and so the problems must not be underestimated. Because objects in the scene are non-ideal for analysis (size and shape variability, lack of sharp corners, and lack of straight edges) and their presentation to the camera is variable, leading to apparent size and shape changes [18], there are significant problems caused by natural and largely uncontrollable lighting. This latter effect manifests itself in shadows, variations in intensity, changes in the spectral content of the light, and other effects such as secondary reflection from plant foliage.

This paper deals with the problem of how to form an image that will eventually allow a good differentiation (image classification) between vegetation and other scene components such as dead plant material, soil, and stones. We will derive appropriate wavebands with which to observe our subjects and propose a measurement space in which to present the results. We will explain the formation of clusters in the space from physical principles. From this understanding, we propose physically based classification methods. Our basis is the physics of the optical processes involved, i.e., illumination, reflection, transmission through filters, and reception at the CCD sensor. We believe that this has advantages over, an approach based just on statistical classification of image pixels. Among these advantages are being able to design the sensing system for optimum (or at least good) performance, gaining an insight into the processes affecting lighting variability and proposing compensations, and dealing with situations where statistical significance using traditional classification methods might be difficult to achieve, e.g., where only a fraction of the scene is vegetation, the agronomically important situation of the early growth stage.

7.2 Physical basis for sensing

7.2.1 Illumination

Most general-purpose illuminants (e.g. tungsten filament lamps, tubular and compact fluorescent lamps and daylight) may be characterized according to their correlated colour temperature (CCT; [4]). The CCT is the temperature of a black body emitting light of an approximately similar spectral composition to that of the illuminant. The Commission Internationale de l'Eclairage (CIE) has a standard [2] that specifies the spectral composition of daylight at various CCTs, which is basis on a study by Judd et al. [11]. This standard (Fig. 7.1) only covers the 300- to 830nm band whereas vegetation reflects well into the near infra red (NIR; see below), a feature that will be used in the design of our sensor. A novel aspect of this work is that the CIE standard has been extended into the NIR by use of a black body model so that these wavebands can be investigated. The authors know of no experimental study, analogous to Judd et al.'s, that has covered this band.

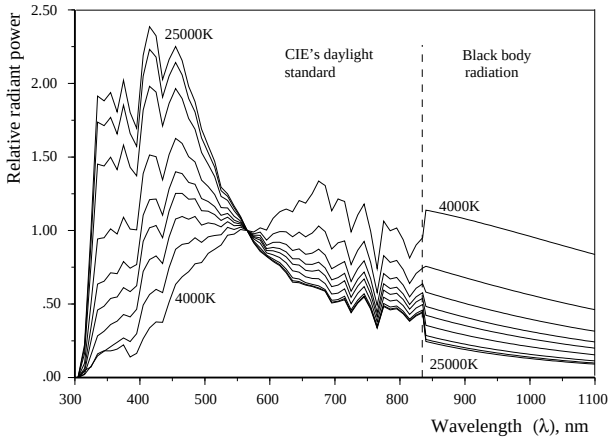


Figure 7.1: *Extension of CIE daylight standard with radiation from a black body.*

7.2.2 The dichromatic reflection model

For modeling of the reflected light from vegetation, the dichromatic reflection model [21] has been shown experimentally to be valid for leaves and vegetation [13],[23],[16]. The model states that the reflected light from an object $L(\Theta, \lambda)$ may be described as the sum of its body and surface reflections as:

$$L(\Theta, \lambda) = m_B(\Theta)L_B(\lambda) + m_S(\Theta)L_S(\lambda) \tag{7.1}$$

where the terms $L_B(\lambda)$ and $L_S(\lambda)$ are the spectral power distribution of the surface and body reflections with geometrical scaling factors $m_B(\Theta)$ and $m_S(\Theta)$, respectively. Θ includes the illumination direction angle (i), the viewing angle (e), and the phase angle (g). The model assumes that the power distributions and the scalings are separable, an assumption justified if the refractive index is constant over the range of wavelengths considered [13].

The standard dichromatic reflection model equation (7.1) also incorporates the neutral interface reflection assumption [13], which assumes that the spectral composition of the surface reflection component corresponds to that of the light source (i.e. the surface reflectance $\rho_S(\lambda)$ is constant). More recently, [23] has categorized this as the dichromatic reflection model type I, which has been shown experimentally to be valid for a variety of materials including, plastics, ceramics, tiles, fruits, leaves, and others [13, 23].

7.2.3 Reflection from the object

The reflected light from the body and surface components (i.e., the incoming light to the sensor) is formed by the incident light $E(\lambda)$ weighted for every wavelength with the body and surface reflectances $\rho_b(\lambda)$ and $\rho_s(\lambda)$ of the object. According to the neutral interface reflection model the surface reflectance is neutral, i.e, $\rho_s(\lambda)$ is a constant. Therefore the surface reflection will have the same colour as the illuminant.

The body reflectance $\rho_b(\lambda)$, derived from the reflectance characteristic in [14] for maize under normal growth conditions, was used (figure 7.2). This is in good agreement with characteristics for vegetation reported elsewhere in the literature [22, 7, 26, 25].

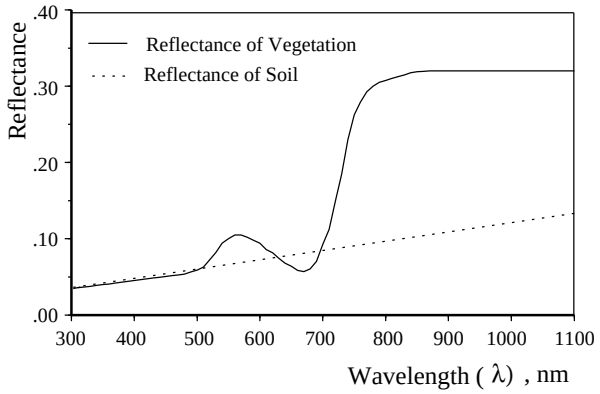


Figure 7.2: *Solid line, Reflectance characteristic used for modeling the body reflection of vegetation (derived primarily from [14]). Dotted line, characteristic of general reflectance of soil (derived primarily from [10])*

7.2.4 The camera

The model in equation 7.1 covers the entire continuous light spectrum. In a practical CCD camera, the incoming light passes through a set of filters. There are typically three: red, green, and blue. They peak in the appropriate area of the spectrum and have a defined shape and bandwidth. Although the filters in our dedicated sensor will not necessarily be red, green, and blue (in particular, blue will be replaced by NIR), we will restrict the work to three channels for engineering feasibility of an eventual practical device. The light at the sensor in each color channel is a convolution of the incoming spectrum with the filter characteristic of each channel. This process of spectral integration is represented by:

$$\mathbf{C}_f(x,y) = m_B(\Theta)\int E(\lambda)\rho_B(\lambda,x,y)\tau_f(\lambda)s(\lambda)d\lambda + m_S(\Theta)\int E(\lambda)\rho_S(\lambda,x,y)\tau_f(\lambda)s(\lambda)d\lambda \qquad (7.2)$$

where $\mathbf{C}_f$ is the output of channel f, $\tau_f(\lambda)$ is the spectral transmittance of filter f, and $s(\lambda)$ is the sensitivity of the camera.

If there are three filters in the camera, the continuous light spectrum will be reduced to a dimension of three. We use the term "color space", by analogy with a standard RGB camera. Hence, a pixel at location (x,y) will be represented by a colour vector $\mathbf{C}(x,y) = [C_1,\ C_2,\ C_3]$. The dichromatic reflection model in color space is expressed by:

$$\mathbf{C}(x,y) = m_B(\Theta)\mathbf{C}_B + m_S(\Theta)\mathbf{C}_S \qquad (7.3)$$

which describes a plane, the so-called *dichromatic plane*, spanned by the body and surface reflection vectors $\mathbf{C}_B$ and $\mathbf{C}_S$, in which the reflection clusters of the object will be situated.

For analysis of the color of the reflected light independent of the intensity, it is convenient to transform the color vector $\mathbf{C}$ to its corresponding chromaticities. This is done by taking the first norm of the color vector, $c1 = C_1/C_1 + C_2 + C_3$, $c2 = C_2/C_1 + C_2 + C_3$, and $c3 = C_3/C_1 + C_2 + C_3$. As $c1 + c2 + c3 = 1$, it is sufficient to consider the chromaticities in only two dimensions, the *c1c2 chromaticity plane*. Another useful property of the chromaticities is that an additive mixture of two lights to produce a third will be a linear combination of the first two as [12]:

$$
\begin{aligned}
c1_3 &= \frac{S_1}{S_1+S_2}c1_1 + \frac{S_2}{S_1+S_2}c1_2 \\
c2_3 &= \frac{S_1}{S_1+S_2}c2_1 + \frac{S_2}{S_1+S_2}c2_2
\end{aligned}
\tag{7.4}
$$

where S_1 and S_2 are the sum of the colour vectors, $S = C_1 + C_2 + C_3$ from the first and second light source, respectively. Equation 7.4 defines a line segment traversing from $(c1_1, c2_1)$ to $(c1_2, c2_2)$.

The dichromatic reflection model describes the reflected light from an object as a mixture of two light sources, the body and surface reflection component, respectively. Its corresponding chromaticities will be distributed along a straight line segment connecting the chromaticity of the body reflection (b) with the chromaticity of the surface reflection (s) (which under the neutral interface reflection assumption corresponds to the illumination source) in the chromaticity plane.

7.3 Sensor design

7.3.1 Choice of wave bands

In the previous section, the basis for design of a dedicated sensor was established. This section focuses on the wavebands that should be used in the sensor for detecting vegetation. As the sensor is constrained to be based on CCD technology, this will limit the wavelength range from about 300 to 1100nm.

The wavebands for each of the three filters must be chosen to allow the sensor to discriminate between vegetation and background. To do this, some modeling of the background has to be done but reflectance characteristics of the background are not as uniform as vegetation. The soil type may range from light sandy soil to black peat having very different reflectance. Further, the background may involve stones, plant residues, and manure, making a very complex situation. In the following, we try to model the background to help in finding optimal wavebands.

First, consider the background primarily to consist of soil. Reflectance characteristics of soil in general slightly increase with wavelength according to the characteristics given in [10] for sandy loam, clay and fen soil. Further, the reflectance of soil and vegetation is approximately at the same level in the blue region (300-500nm) of the spectrum. Figure 7.2 (from [10]) shows a general reflectance characteristic of soil. Accordingly, discrimination of soil from vegetation might be achieved by using the pronounced reflectance of vegetation in the NIR (700-1100nm) and the depression in the red(630-700nm). By placing wavebands on either side of 700nm, it should be possible to achieve a good contrast between the two. This approach is widely used within remote sensing and in the studies of [10, 9, 6]

However, as mentioned, the background does not solely consist of soil. Stones, plant residues, and manure may also be present. may be present. One source of misclassification due to the above criteria is that organic materials in the soil may also have high reflectance in the NIR and depression in the red due to the high content of chlorophyll (i.e., plant residues, peat soils, and manure). Hence, they will be likely to be misclassified. However, the color of these materials is generally less green than vegetation so the pronounced reflectance of vegetation in the green (500-630nm) may be used for discrimination.

As a result of the above physical considerations, the sensor should be based on wavebands in the green, red, and NIR. The blue is not involved for two reasons. First, it does not provide much discriminant power. Second, as the variation of the spectral composition of daylight as a function of CCTs is most pronounced in the blue region (figure 7.1) the sensor becomes less sensitive to daylight spectral changes.

The following standard Kodak Wratten filters are used as a basis for the sensor: green no. 58, red no. 25, and NIR no. 89b. To avoid NIR transmission through the green and red filters, an infra-red block filter type SWP-BL-685 from DELTA Light & Optics, Lyngby, Denmark is placed in front of them. The three filters are placed in a filter wheel mounted in front of a Pulnix PE2015 (Pulnix Europe Ltd., Hampshire, England) monochrome camera as illustrated in figure 7.3.

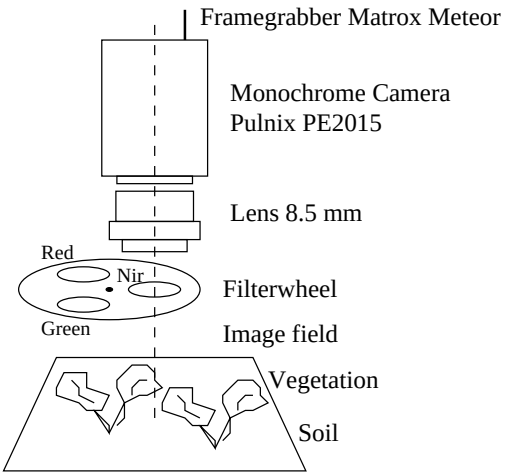


Figure 7.3: *Illustration of the experimental setup.*

From specifications provided by the manufacturers, it is now possible to model the three output channels of the sensor at a normal daylight (CCT = 5700K) condition by:

$$S_f(\lambda) = E_{5700}(\lambda)\tau_f(\lambda)s(\lambda) \tag{7.5}$$

where $E_{5700}(\lambda)$ is a reference daylight with a CCT of 5700K, $\tau_f(\lambda)$ is the transmittance of the three filters, and $s(\lambda)$ is the sensitivity of the Pulnix camera CCD.

7.3.2 Balancing between channels

Theoretical Balancing

The objective of balancing is to give approximately equal outputs from the three channels, making it possible to realize a large part of the dynamic range of the CCD sensor. If the scene consist of a single color, then balancing the system on a target of that color is appropriate.

For general use, in the absence of any prior knowledge on scene color, this done this is done with a white card or, in more scientific work, with a plate of barium sulphate. However, for mixed scene colors, it is difficult to devise a principle way of color balancing. The most constant part of our scenes, in terms of reflectivity, is the vegetation. Therefore, we take the pragmatic view (at least initially) that we will balance on the vegetation. Equation 7.5 gives the response of the unbalanced sensor with a reference daylight of 5700K with wavebands placed as above. For balancing, only body reflection from the vegetation is considered (first term in equation 7.2), which leads to equation 7.5 being weighted for every wavelength with the reflectance of vegetation, figure 7.2, and is included in equation 7.6 as $\rho_{vegetation}$. Hence, the body vector gets maximum dynamic range. The drawback of this approach is that under illumination giving pronounced surface reflection, the dynamic range of the sensor will be limited. So, if the sensor is to operate mainly under sunny conditions, it might be more sensible to balance it with a white reference. Neutral density glass filters from Delta Light & Optics were used for balancing and their attenuation found by the following criteria:

$$\begin{array}{lcl} ND_{green} \int \rho_{vegetation}(\lambda) S_{green}(\lambda) d\lambda & = & \\ ND_{red} \int \rho_{vegetation}(\lambda) S_{red}(\lambda) d\lambda & = & \\ ND_{NIR} \int \rho_{vegetation}(\lambda) S_{NIR}(\lambda) d\lambda & & \end{array} \tag{7.6}$$

where the NDs are the transmissions of the neutral density filters for the respective channels, assumed constant over the wavelength range of the filters. This theoretical balancing requires that the transmissions of each channel should be in the ratios (green:red:NIR) 1.000:0.965:0.385.

Practical Balancing

In the above the requirements for ideal balancing with vegetation as the object, modelled by information provided by the manufacturers, was established. However, realistic scenes of interest do not consist solely of vegetation and also the transmittance of filters may deviate from the manufacturers' data. Especially, the red channel may be expected to be dominated by the background due to the higher reflectance of soil in this waveband and hence saturated. To accommodate these problems and after some initial experimentation, it was decided to finalize the ratios at 1.000:0.500:0.500.

After the red and NIR channel filters were mounted in the filter wheel, with neutral density filters having an attenuation of 50%, the transmittance of each channel was measured with a Monolight spectrometer (Mono Instruments Ltd, England). Next, every filter was weighted for every wavelength with the sensitivity of the Pulnix camera to give the sensor tristimulus characteristic, figure 7.4. Importantly, we note that the potential sources of error in the camera are limited to the uncertainties of measuring the transmittance of the filters and the spectral sensitivity of the camera as provided by Pulnix. These two sources when compared with the rather crude modeling of the daylight may be regarded as minor.

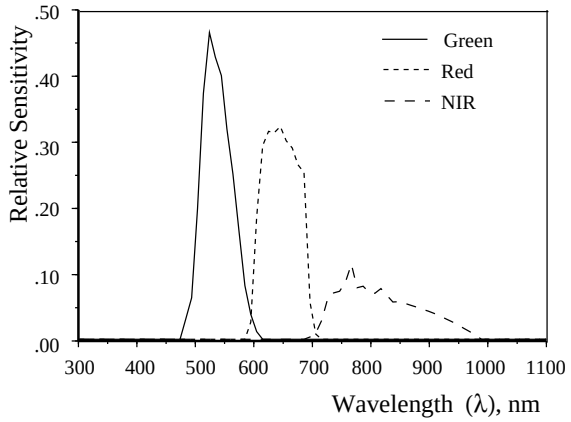


Figure 7.4: *Relative spectral sensitivity of the sensor used in the study. Transmittance of the filters was measured with the Monolight spectrometer and camera sensitivity by Pulnix.*

7.4 Evaluation

7.4.1 Modeling of Daylight

Before evaluating the sensor, the validity of using the black body radiation as an extension of the CIE daylight standard will be assessed. This extension is expected to serve only as an underlying model of the spectral composition of daylight into the NIR. It is expected that adjustments will be necessary.

As the sensor may be regarded as orthogonal, a diagonal matrix relating the modeled data to the measured may be obtained by:

$$||\mathbf{C}_{BaSO_4}|| = ||\mathbf{C}_{model}||\,\mathbf{\Lambda} + e \tag{7.7}$$

where $\mathbf{\Lambda}$ is a 3×3 diagonal matrix with coefficients α, β, and γ and the color vectors $\mathbf{C}_{BaSO_4}$ and $\mathbf{C}_{model}$ are normalized to unit vectors. The solution for $\mathbf{\Lambda}$ is obtained by least squares estimation to minimize the error e.

7.4.2 Modeling of Sensor

For evaluation of the sensor model, it is of interest to see if the vegetation clusters move with varying CCT as predicted. This will be done by calculating the CCT from the cluster positions and comparing it with the measured values.

For the first analysis, a series of images illuminated predominately by one light source (i.e., the sky or direct sunlight) at varying CCT are acquired. Under these conditions, the reflection clusters will be distributed either along the line segment connecting the chromaticities of the body and surface reflection (equation 7.4) (sunlight condition) or at the chromaticity of the body reflection (skylight condition). However, for both situations, the pixel points are expected to lie in the dichromatic plane spanned by the body and surface reflection vectors. The plane best describing the image pixels is found by the one giving the minimum average distance from every pixel point to the dichromatic plane at the given CCT, as:

$$\min \left(\frac{1}{m} \sum_{i=1}^{m} \frac{|\mathbf{n} \cdot \mathbf{C}_i|}{|\mathbf{n}|} \right) \tag{7.8}$$

where m is the number of pixel points, $\mathbf{n}$ is the normal vector $(\mathbf{C}_B \times \mathbf{C}_S)$ of the dichromatic plane spanned by body and surface reflection vectors at corresponding CCTs, and $\mathbf{C}_i$ is the color vector of pixel point i. Equation 7.8 may be expanded to:

$$\min \left(\frac{1}{|\mathbf{n}|} \left(\left| n_1 \frac{1}{m} \sum_{i=1}^{m} C_{1,i} + n_2 \frac{1}{m} \sum_{i=1}^{m} C_{2,i} + n_3 \frac{1}{m} \sum_{i=1}^{m} C_{3,i} \right| \right) \right) \tag{7.9}$$

where $\mathbf{n_i}$ are the components of the normal vector and $\mathbf{C}_{1,2,and3,i}$ of the color vector. As $\frac{1}{m} \sum_{i=1}^{m} \mathbf{C}_i$ corresponds to the mean of color $\mathbf{C}$ and $\frac{1}{|\mathbf{n}|}$ is constant for a given CCT of the body and surface vectors, the problem simplifies to finding the minimum of the scalar triple product:

$$\min |(\mathbf{C}_B \times \mathbf{C}_S) \cdot \mathbf{C}_{mean}| = \min \left\| \begin{array}{ccc} C_{B,1} & C_{S,1} & C_{mean,1} \\ C_{B,2} & C_{S,2} & C_{mean,2} \\ C_{B,3} & C_{S,3} & C_{mean,3} \end{array} \right\| \tag{7.10}$$

with $\mathbf{C}_B$ and $\mathbf{C}_S$ at corresponding CCTs.

7.5 Experimental Data

Images of barley and rape plants acquired during the summer of 1998 at Silsoe Research Institute were used for assessment of the sensor. For each image, the CCT was measured using a Minolta Colour Meter II, Ahrensburg, Germany. Images from both species were acquired in direct sunlight, skylight, and with a cardboard sheet casting a shadow over the image field for both situations. In total, 23 images (11 barley and 12 rape) varying in CCT from 5550K to 10000K were used for evaluation of the sensor model's ability to predict the reflection cluster movements with varying CCT.

To classify which pixels were vegetation, all images were pre-classified using a comprehensive statistical method. First, the images were transformed into their r, n chromaticities and a two- dimensional histogram was generated. In this histogram, a linear discriminant was trained by a training set from an

average linkage cluster analysis using the SAS system [3] with two clusters. To avoid the influence of outliers on the cluster analysis, the data were filtered first by thresholding the cumulative histogram at 5%, so bins with a lower frequency were excluded, and next by checking for neighbors in the r, n histogram in a 3×3 neighborhood where bins with less than four neighbors were excluded. After training the linear discriminant classifier, the whole image was classified. The final segmentation result was assessed by human assessment.

Reference images of a barium sulphate plate were acquired immediately after each of the acquired images and into November 1998. In this way, images were acquired from May to November, giving a range in CCT from 4850 to 17500K.

7.6 Results and Discussion

7.6.1 Sensor Behaviour

The coefficients in the diagonal matrix $\mathbf{\Lambda}$ (equation 7.7) were calculated as $\alpha = 0.95$, $\beta = 1.06$, and $\gamma = 1.26$ with standard deviations 0.0041, 0.0078, and 0.025, respectively. These were used to correct the modeling of the surface and plant body reflection loci. Figure 7.5 (a) illustrates the corrected and uncorrected loci together with a plot of the chromaticities of the barium sulphate measurements and projections of the dichromatic planes at CCTs of 4000, 5000, 7000, 10000, and 25000K onto it equation 7.4. Figure 7.5 (b) illustrates the relationship between the red chromaticity and the CCT for the modeled corrected surface locus and the barium sulphate measurements. The modeled locus systematically overestimates the red chromaticity. Unfortunately, it is not known whether this error is due to the color meter readings or to the sensor model. However, figure 7.5 shows that the error is minor and that it is reasonable to conclude that the corrected model follows the experimental barium sulphate measurements.

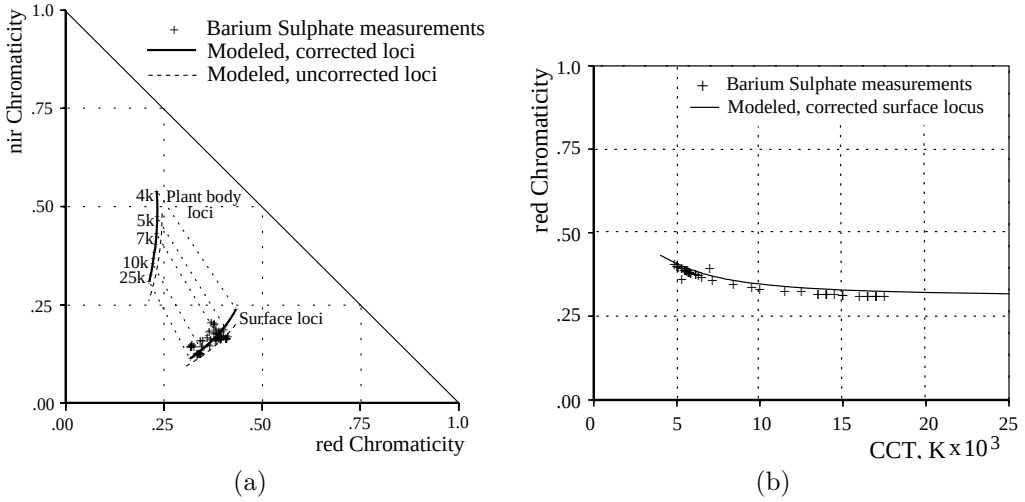


Figure 7.5: *(a) Corrected and uncorrected surface and plant body loci together with barium sulphate measurements. Dotted lines illustrate projection of dichromatic planes at CCT 4000, 5000, 7000, 10000, and 25000K. (b) Red chromaticity plotted against CCT for the modeled corrected surface locus and the barium sulphate measurements.*

The corrected values of the surface and plant body reflection vectors are now used for prediction of the CCTs of the 23 pre-classified images by finding the minimum of the scalar triple product according to equation 7.10. This procedure identifies which of the planes projected into the chromaticity diagram of

figure 7.5, and hence which CCT, best describes the vegetation data. In figure 7.6, the estimated values are plotted against the measured ones. There was no significant difference between the estimated and measured CCTs or between the two plant species (p ¡ 0.15).

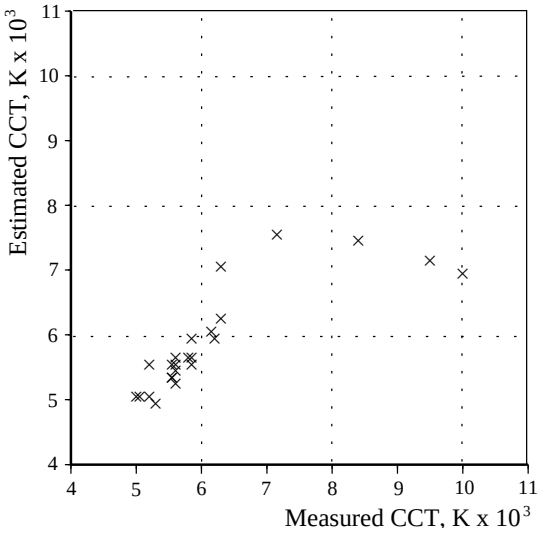


Figure 7.6: *Plot of estimated against measured CCT.*

Overall, the results are in good agreement with what is expected. Despite the rather crude extension of the illumination by radiation from a black body, the results suggest that for computer vision purposes it is sufficient when corrected by a coefficient for each channel. However, the result also indicates that an extension based on deviation of principal components and scaling factors of a larger dataset of spectral curves should be possible, as in [11].

When comparing the estimated CCT with the measured CCT, the model seems most adequate to estimate correct values at a low CCT. Unfortunately, it was only possible to get three images above 8000K, due to the rainy and misty summer of 1998. Although from these images the model seems to underestimate the CCT, collection of further images is necessary to draw any final conclusion. Also, one should be aware of the hyperbolic relationship between chromaticity and CCT, which squashes the chromaticities at high CCTs. Errors in estimation of the mean vector $\mathbf{C}_{mean}$ have a larger impact on the error of the estimated value in this region.

7.6.2 Cluster Positions and Shapes

Figure 7.7 shows four images with various characteristics that need to be handled by a method of classification. The blue channel is used to display the NIR component. All objects having pronounced NIR reflection appear bluish. The first row (sel1) shows an image of cauliflower and rape seedlings. Plant residues, which should not be classified as vegetation, include the two dead cauliflower leaves in the right part of the image and some two-day old cut grass. The second row (sel2) contains an image of barley seedlings together with straw and two stones in the upper left corner. The third row (sel3) shows a cauliflower plant with pronounced shading. The image in the bottom row (sel4) addresses the problems of a small population of vegetation and the three stones in the middle of the image.

Daylight is, in general, a mixture of two light sources, sunlight and skylight. It can be shown [1] that pixels from an object illuminated by such a mixture fall within a quadrilateral. This quadrilateral has corners, two on each of the body and surface loci at the CCTs of the two illuminants. If it is assumed [1] that the skylight is diffuse and produces no significant surface reflection, then the quadrilateral reduces

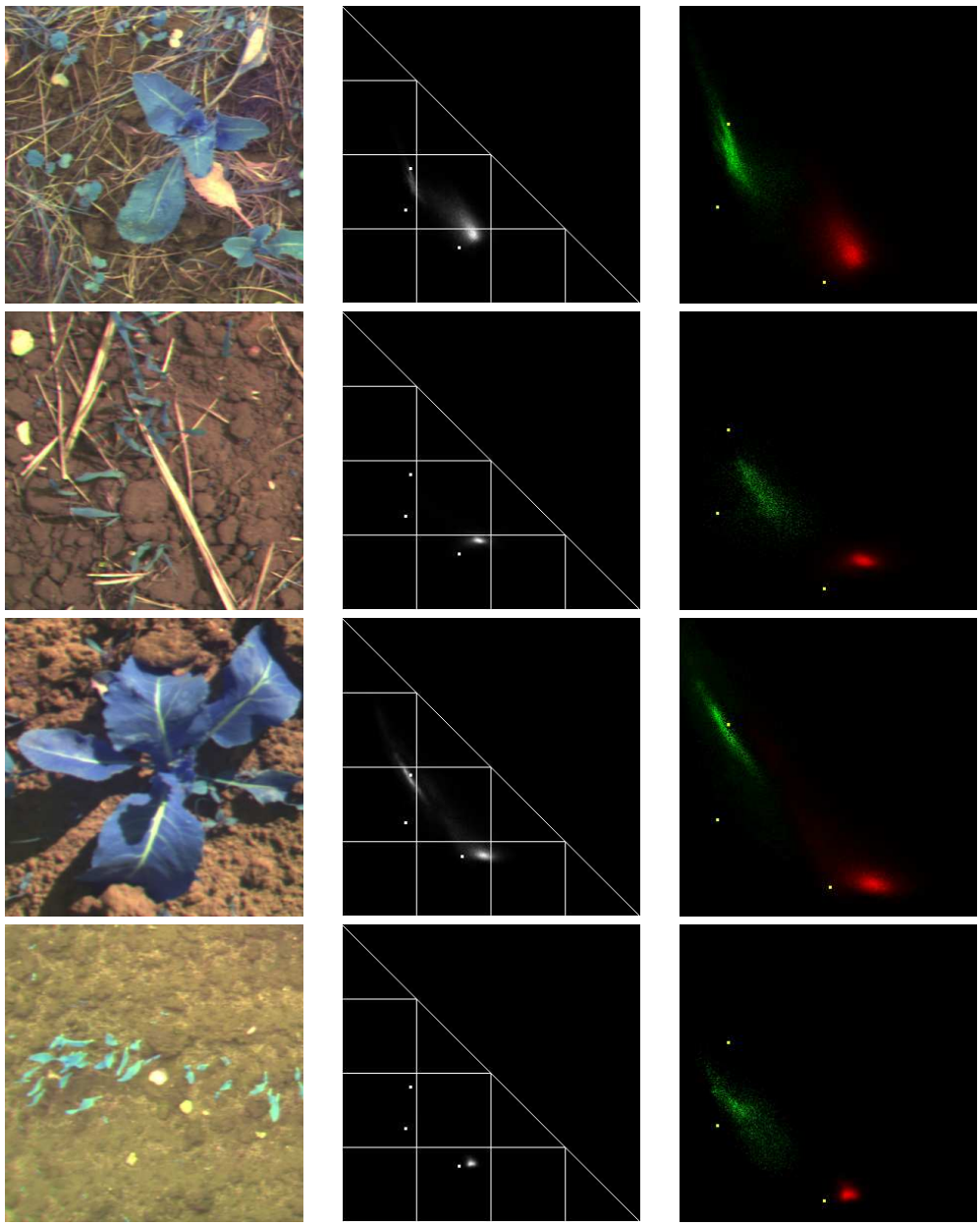


Figure 7.7: *Images used in comparative tests. From top to bottom: sel1, sel2, sel3, sel4. Left column: original images; middle column, histograms in red/NIR plane; right column, data from manual classifications in red/NIR plane. green shows vegetation, red shows background.*

to a triangle. One corner is on the surface locus at the CCT of the sunlight. The other two corners are on the body locus at the CCTs of the sunlight and the skylight. As the skylight CCT was not measured, we assume that the lower corner of the triangle on the body locus could extend as far as 25000K, which represents the maximum CCT considered possible.

Figure 7.7 shows (central column) histograms in the red/NIR chromaticity plane for all the pixels of the images in the left column. The axes and scales are the same as those in figure 7.5. The histogram is plotted such that the brightness corresponds to the number of pixels at the particular chromaticity, with the maximum value scaled to a brightness of 255. In some plots(e.g., sel3), two clusters can be clearly seen. In others (e.g., sel2), only one is visible. This latter effect is caused by the vegetation cluster having a very small population compared with the background - this would lead to difficulties if the space were partitioned using a purely statistical procedure. The bright dots are the three corners of the triangle mentioned above where the measured CCTs were 5600, 5600, 5100, and 5600K for the four images.

In order to produce ground truth, the four images were carefully classified into live vegetation and background by visualizing them as a color image with the NIR channel seen as blue. This was primarily done manually, by using a graphics tablet, assisted by the color classification tools in the Adobe Photoshop package. Areas suggested by the tools were accepted, rejected, or (more likely) modified by hand until the resulting classification was visually acceptable. Like all ground truth methods, the human observer could not be certain about all decisions, thus, as well as the two classes, we had a third "uncertain" class. The proportions of vegetation, background, and uncertain are given in Table 7.1.

Table 7.1: *Proportion of vegetation, background, and uncertain for the test images.*

	Proportion (%)		
Image	Vegetation	Background	uncertain
sel1	54.6	34.6	10.8
sel2	4.7	92.4	2.9
sel3	37.0	54.4	8.6
sel4	3.0	94.3	2.7

The right hand column shows the manually classified vegetation class (shown as green) and background class (red). Uncertain has not been plotted. The scale has been expanded somewhat for clarity and shows the chromaticity range 0.15 - 0.65. The spots defining the triangle corners can be used to register the diagrams. The brightness scaling here is done for the two classes seperately so that both are visible.

Figure 7.7 show that the two clusters in the central column are vegetation and background. There is no visible contamination between the components, even stones, dead vegetation, and shadows were present. The clusters are well separated in the diagrams, showing that the physically based choice of sensor wave bands has been successful. Soils are likely to have fairly uniform spectral responses (grey) or they increase slightly with wavelength (reddish). Therefore, the background cluster should be situated near to the daylight locus at the measured CCT or toward the red. This can be seen in all four diagrams. The vegetation cluster is generally contained within the predicted triangle, although the cluster in sel3 extends outside. In the absence of shadows, the triangle collapses to a line from the plant body locus to the daylight locus (upper right boundary of the triangle) and the vegetation should form a linear feature in the diagram. In fact, the shadow area is relatively small in all four images and it may be expected that the low intensity in shadows would lead to some variability when calculating chromaticities and so disperse entries in the histogram. Thus, shadowing should leave the linear nature of the expected features intact. This linearity is apparent in all four images, especially when there is enough vegetation to give a strong cluster (sel1 and sel3). Highlights add a significant proportion of the illuminant color to the pixels, whereas in diffuse conditions, the illuminant component is small. As there are comparatively few highlights, the expectation is that the linear feature would cluster nearer to the plant body locus than the daylight locus, which is also apparent in the four images.

There are, of course, some unexplained details. For instance, in sel2, the vegetation cluster is not centered around the plant body point; in sel4, the vegetation cluster is significantly into the triangle. On the other hand there are also some unmodeled effects like secondary reflection from the plants and the soil, and transmission through the plant leaves. However, the features are predicted well in a qualitative way and reasonably well quantitatively. A challenge for future work will be to use the predictions from physical

modeling (both qualitative and quantitative) in the best possible way to assist classification.

7.7 Conclusions

This work has introduced the design of a dedicated sensor for forming images of vegetation. The design is based on physical modeling of the image formation process involving modeling the reflection of vegetation, soil, and other artifacts in the background, the transmission of light through three filters, interception on the CCD, and modeling of the illumination. For modeling of the illumination, the CIE daylight standard was extrapolated into the NIR region. The results indicate that for computer vision purposes, this should be satisfactory. The analysis shows that the sensor should be based on three wave bands - green, red, and NIR. The modeled behavior of the sensor has been compared with 23 images ranging in CCT from 5550 to 10000K. The analysis shows that the sensor behaves as predicted. Finally, clusters in a measurement space of the red and NIR chromaticities for four selected images appear in accordance with their expected shapes and positions. Presentation in this space gives good separation of the vegetation and nonvegetation clusters, which will be suitable for physically based classification methods to be developed in future work.

7.8 Acknowledgments

The work of Hans Joergen Andersen was funded by the Danish Research Council under the Optimal research programe. The UK contribution was funded by the Biotechnology and Biological Sciences Research Council.

References

[1] H.J. Andersen and E Granum. Classifying the illumination condition from two light sources by colour histogram assessment. *Journal of the Optical Society of America A*, 17(4):667–676, April 2000.

[2] Anonymous. Colorimetry. Technical Report 2. edition, Commision Internationale de L'Eclairage (CIE), 1986.

[3] Anonymous. Sas/stat user's guide. Technical report, SAS Institute Inc. Cary, North Carolina, USA, 1994.

[4] Anonymous. Method of measuring and specifying colour rendering properties of light sources. Technical Report CIE. 13.3, Commision Internationale de L'Eclairage (CIE), 1995.

[5] J.V. Benlloch, A. Sanchez, S. Christensen, and M. Walger. Weed mapping in cereal crops using image analysis techniques. In *AgEng96*, volume 2, pages 1059–1060, 1996.

[6] H.R. Biller. Reduced input of herbicides by use of optoelectronic sensors. *Journal of Agricultural Engineering Research*, 71:357–362, 1998.

[7] W.D. Billings and R.J. Morris. Reflection of visible and infrared radiation from leaves of different ecological groups. *American Journal of Botany*, 38:327–331, 1951.

[8] R. Brivot and J.A. Marchant. Segmentation of plants and weeds using infrared images. *Proceedings of the Institution of Electrical engineers, Vision, Image and Signal processing*, 143(2):118–124, 1996.

[9] R.J. Hagger, C. J. Stent, and S. Isaac. A prototype hand-help patch sprayer for killing weeds, activated by spectral differences in crop/weed canopies. *Journal of Agricultural Engineering Research*, (28):449–358, 1983.

[10] A.W. Hooper, G.O. Harries, and B. Ambler. A photoelectric sensor for distinguishing between plant material and soil. *Journal of Agricultural Engineering Research*, (21):145–155, 1976.

[11] D.B. Judd, D.L. MacAdam, and G.W. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54, 1964.

[12] H.-C Lee. Illuminant color from shading. *In Proc. SPIE 1250: Percceiving, Measuring and Using Color*, pages 236–244, 1990.

[13] H.-C Lee, E.J. Breneman, and C.P. Schulte. Modeling light reflection for color computer vision. *IEEE trans. on Pattern Analysis and Machine Intelligence, PAMI*, 12(4):402–409, 1990.

[14] S.J. Maas and J.R. Dunlap. Reflectance, transmittance, and absorptance of light by normal, etiolated, and albino corn leaves. *Agron, J*, 81:105–110, 1989.

[15] J.A. Marchant and R. Brivot. Real-time tracking of plants rows using a Hough transform. *Real-Time Imaging*, 1:363–371, 1995.

[16] F. Pla, F. Ferri, and M. Vicens. Colour segmentation based on a light reflection model to locate citrus for robotic harvesting. *Computers and Electronics in Agriculture*, 9:53–70, 1993.

[17] F. Pla, J.A. Sanchiz, J.A. Marchant, and R. Brivot. Building perspective models to guide a row crop navigation vehicle. *Image and Vision Computing*, 15:465–473, 1997.

[18] D. Reynard, A. Wildenberg, A. Blake, and J.A. Marchant. Learning dynamics of complex motions from image sequences. In *Proceedings of the 4th European Conference on Computer Vision*, pages 357–368, 1996.

[19] J.M. Roger, E. Molto, G. Rabatel, and J. Blasco. Design of a robotized, non-chemical weed controller. In *Bio-Robotics'97*, pages 229–235, 1997.

[20] J.M. Sanchiz, F. Pla, J.A. Marchant, and R. Brivot. Structure from motion techniques applied to crop field mapping. *Image and Vision Computing*, 14:353–363, 1996.

[21] S.A. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985.

[22] C.A. Shull. A spectrophotometric study of plant reflection of light from leaf surfaces. *BOT. Gaz.*, 87:583–607, 1929.

[23] S. Tominaga. Dichromatic reflection models for a variety of materials. *COLOR research and application*, 19(4):277–285, 1994.

[24] E. Vrindts and J de Baerdemaerker. Optical discrimination of crop, weed and soil for on-line weed detection. In J.V. Stafford, editor, *Precision Agriculture '97*, volume 2, pages 537–544, 1997.

[25] E.A. Walter-Shea and J.M. Norman. *Photon-Vegetation Interactions*, chapter Leaf Optical Properties, pages 230–250. Springer-Verlag, 1991.

[26] J.T. Wooley. Reflectance and transmittance of light by leaves. *Plant Physiology*, 47:656–662, 1971.

Chapter 8

Evaluation of an Imaging Sensor for Detecting Vegetation Using Different Waveband Combinations

Evaluation of an Imaging Sensor for Detecting Vegetation Using Different Waveband Combinations

J.A. Marchant[a], H.J. Andersen[b] and C.M. Onyango[a]

a, Image Analysis and Control Group, Silsoe Research Institute,
Wrest Park, Silsoe, Bedford, UK, MK45 4HS
b, Laboratory for Image Analysis, Aalborg University 9220 Aalborg, Denmark

Abstract

This paper uses data collected from a previously reported imaging sensor to investigate the classification of vegetation from background. The sensor uses three wavebands:- red, green, and near infra-red (NIR). A classification method (the "alpha" method) is introduced which is based on an understanding of the imaging processes. This is compared with two ratio methods (red/NIR and red/green) and two single band methods (NIR and green intensity).
The Receiver Operating Characteristic Curve (ROC) method is used to evaluate the classifications on realistic test images. The alpha and ratio methods all perform reasonably well with the red/green ratio slightly the worst. The single band methods perform significantly less well with green intensity easily the worst. The alpha and ratio methods have "best" thresholds that correspond with detectable histogram features when there is a significant amount of vegetation in the image. The physical basis for the alpha method means that there is a detectable mode in the histogram that corresponds with the "best" threshold even when there is only a small amount of vegetation. The single band methods do not produce histograms which could easily be analysed and so their use should be confined to simple images.

Published in: Computer and Electronics in Agriculture. 32(2), 101-107, 2001

8.1 Introduction

Various authors, e.g. [8, 14, 23, 18, 4], have proposed that optical sensing used to detect vegetation from ground-based vehicles could help to target crop treatments. These proposals have been motivated by growing concerns (economic, environmental, and ethical) about the use of crop protection chemicals and the desire to reduce quantities applied. A previous paper [2] introduced a sensor (which we called the "dedicated sensor") that used three wavebands (red, green, near infra-red (NIR)) to detect vegetation. The sensor was based on an understanding of the physics of lighting, reflection by the vegetation, transmission through the bandpass filters, and reception at the sensor of a CCD camera. The previous work showed that the image data formed clusters representing vegetation and background respectively. The clusters had the shapes and positions predicted by the physical understanding in a two dimensional plane of red and NIR chromaticities.

In this paper we use the sensor to investigate a method of image classification that takes advantage of the physical understanding. We present a quantitative evaluation of the sensor against "ground truth" as established by manual image classifications. We also compare the performance with methods consisting of alternative combinations of subsets of the three wavebands.

8.2 The Dedicated Sensor

The dedicated sensor (described fully in [2]) consists of a monochrome CCD camera with a filter wheel fitted to the front that can position each filter in turn in front of the lens. The filters were chosen to give good separation of the vegetation and background components in colour space based on the physics of the optical processes involved. Attention was paid to balancing the optical pathways to give a good dynamic range in each channel. The spectral sensitivity (i.e. the characteristics of the three filters convolved with the CCD element of the sensor) is shown in figure 8.1

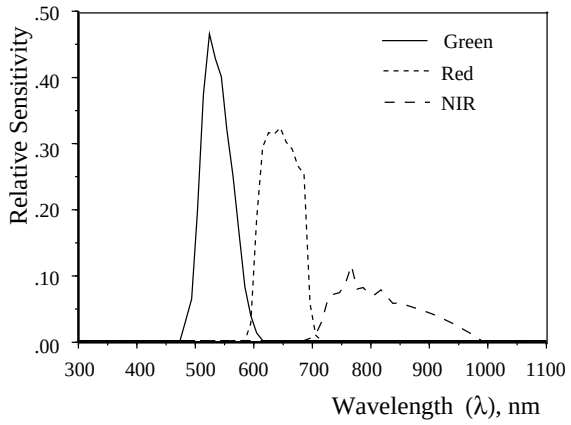


Figure 8.1: *Relative spectral sensitivities of the dedicated sensor used in the study.*

8.3 Analysis of Measurements

8.3.1 Measurement Space

A full description of the light reflected by an object as seen by a three-band colour camera is given by the colour vector C which, in the case of the dedicated sensor, has components from the red, green, and NIR

channels. For analysis of the colour of light independent of the intensity, previous work [2] transforms the colour vector C to its corresponding chromaticities. This is done by taking the first norm of the colour vector, $c1 = C_1/C_1 + C_2 + C_3$, $c2 = C_2/C_1 + C_2 + C_3$, and $c3 = C_3/C_1 + C_2 + C_3$. As $c1 + c2 + c3 = 1$, it is sufficient to consider the chromaticities in only two dimensions, in our case the *measurement space* is the *red/NIR-chromaticity plane*.

8.3.2 Body and Surface Loci

Our previous work uses the dichromatic reflection model [19] to represent reflected light from vegetation [13], [21], [16]. The model states that the reflected light from an object may be described as the sum of its body and surface reflections. The standard dichromatic reflection model also incorporates the Neutral Interface Reflection assumption [13], which assumes that the spectral composition of the surface reflection component corresponds to that of the light source. It has been found experimentally [11] that the spectral composition of daylight can be characterised by a single measurement,the correlated colour temperature (CCT). The relationship is embodied in a standard by the Commission Internationale de l'Eclairage (CIE) [3]. This was extended in our previous work into the NIR using a black body model and a correction based on measurements of reflection from a Barium Sulphate plate at different CCTs. Both the colour of the daylight and the vegetation body colour as seen through the sensor can be plotted on the red/NIR chromaticity plane as two loci as they vary with CCT, Figure 8.2 (a). The dichromatic reflection model predicts that the colours of points on an object illuminated from a single source will fall somewhere along a straight line that extends from a point on the object's body locus to the point at the same CCT on the surface locus.

8.3.3 Classification Method

Figure 8.2 (b) shows a histogram in the red/NIR chromaticity plane for the pixels of a typical image. The image, referred to as sel1, consists of about 60% live vegetation with the other 40% made up of soil and other non-vegetative material along with some dead vegetation. The actual image will be introduced later. The histogram is plotted such that the brightness corresponds to the number of pixels at the particular chromaticity with the maximum value scaled to a brightness of 255. In our previous work [2] we hand classify the image and show that the cluster towards the upper left is vegetation while that towards the lower right is background. Soils are likely to have fairly uniform spectral responses (grey) or slightly increasing with wavelength (reddish) and so the background cluster should be situated near to the daylight locus at the measured CCT or towards the red. This can be seen in Figure 8.2 (b) as the lower right cluster. As the background does not consist of a single coloured component the cluster will not be the linear feature predicted by the dichromatic reflection model - this explains its rather dispersed nature. On the other hand, the vegetation is much more uniformly coloured producing the linear feature along the line joining the body to the surface loci at the measured CCT.

We now introduce a method to transform from the two-dimensional to a one-dimensional histogram based on the physical understanding of the optical processes. Daylight is, in general, a mixture of two light sources:- sunlight and skylight. It can be shown [1] that pixels from an object illuminated by such a mixture fall within a quadrilateral. This quadrilateral has corners, two on each of the body and surface loci at the CCTs of the two illuminants. If it is assumed [1] that the skylight is diffuse and produces no significant surface reflection, then the quadrilateral reduces to a triangle. One corner is on the surface locus at the CCT of the sunlight. The other two corners are on the body locus at the CCTs of the sunlight and the skylight. As the skylight CCT was not measured, we assume that the lower corner of the triangle on the body locus could extend as far as 25000K, which represents the maximum CCT considered possible. Figure 8.2 (b) illustrates the baseline of the triangular region.

A one dimensional histogram can be obtained by summing pixel points having angle α defined as the angle between the pixel chromaticity and the baseline. To account for errors in modeling we allow α to extend

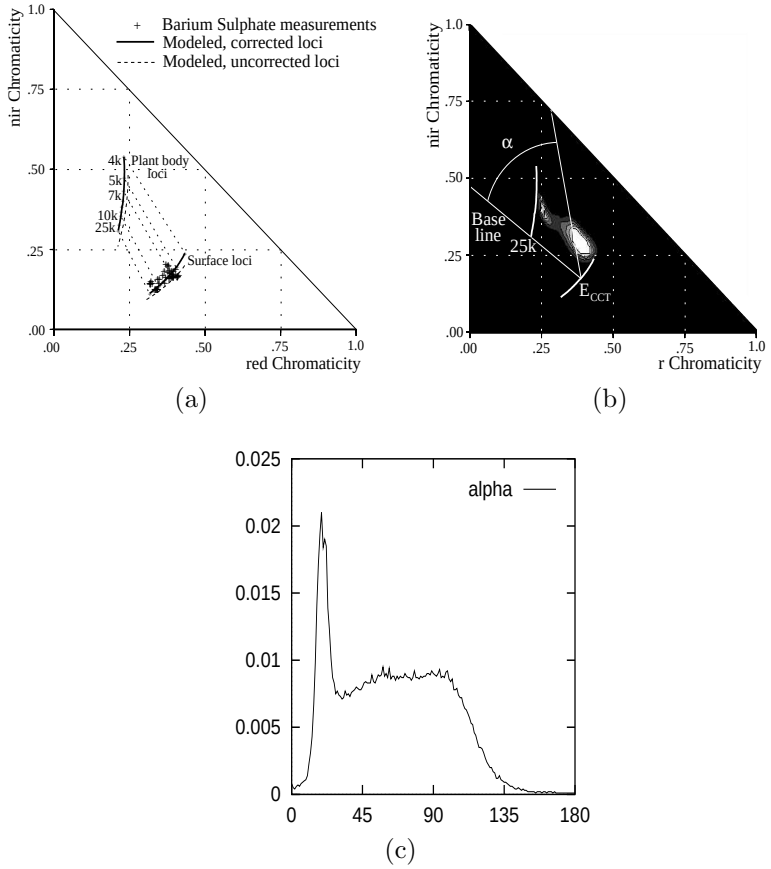


Figure 8.2: *(a) Plant body and surface loci. (b) Grey level plot of the red/NIR chromaticity histogram of a typical image. The contents of each histogram bin are represented by the brightness. E_{CCT}, is the chromaticity of illumination for the image. The base line is included as the line from E_{CCT} to the chromaticity on the plant body reflection loci at 25000K. (c) Histogram produced by the alpha method.*

past the region delineated strictly by the model, i.e. past the point on the body locus corresponding to E_{CCT}. This idea will be called "the alpha method". This is equivalent to rotating a line (or more precisely a very narrow wedge) about the measured CCT point on the surface locus and counting the pixels within the wedge. The included angle of the wedge determines the bin size of the histogram.

Although the analysis above predicts vegetation points anywhere within the triangle, the points tend to be more clustered than this. In general the amount of vegetation in shadow is relatively small - this is not surprising as plants have evolved to take advantage of daylight. This means that there is only one significant illuminant corresponding to the measured (unshaded) CCT. Therefore points should cluster along a line between corresponding points on the two loci and form a linear feature as shown. As the narrow wedge sweeps past the cluster, most of the vegetation pixels fall into a small number of bins and so the linear clustering gives rise to an amplification of the vegetation mode in the histogram. This is useful in situations when there is a small proportion of vegetation in the image, as will be illustrated later.

Figure 8.2 (c) shows the one dimensional histogram where the amplification can be clearly seen. The expected position of the vegetation mode, obtained by calculating α at the point on the plant body locus at the measured CCT, is 23° which agrees well with the data. Because the background cluster extends

over a comparatively large range of α, the corresponding mode in the histogram is much flatter.

8.4 Quantitative Measurement of Classification Potential

When classifying from a one dimensional histogram, the analyst has to answer two questions:-

1. How successfully can the image be divided into two classes?

2. Can a sensible threshold be chosen from the histogram?

In this work we consider each question separately and this section investigates the first.

A suitable method for investigation is the Receiver Operating Characteristic (ROC) curve [22]. This method uses the notion of "true positives" (in our case vegetation classified as vegetation); "true negatives" (background classified as background); "false positives" (background classified as vegetation); and "false negatives" (vegetation classified as background). To form an ROC the true positive ratio, tpr - defined as (number of true positives)/(true positives+false negatives) - is plotted against the false positive ratio, fpr - (number of false positives)/(true negatives+false positives). The parameter that is varied here is the threshold on α. With the threshold at one extreme nearly all the image will be classified as foreground, the number of false positives will be very high, and fpr will approach 1.0. Conversely the number of false negatives will be small and so tpr will also approach 1.0. With a threshold at the other extreme the opposite situation obtains and both fpr and tpr approach 0.0. The ideal combination (i.e. a perfect classification) is fpr=0.0 and tpr=1.0. If the algorithm is a good one, permitting the image data to be thresholded well, the curve should pass close to this point.

In situations where the performances of competing algorithms (or sensors) are to be compared, various indices derived from the ROC are available. If there is no basis at all on which to choose a threshold, a global feature of the ROC can be used such as the area under the curve [9]. Alternatively, ROCs could be compared by selecting the "best" operating point taking the optimistic view that a threshold for such a point could eventually be chosen. Methods based on Bayes risk analysis are often used in this context. If the prior probabilities of vegetation and background are known along with the costs of making the various classifications (mistakes and correct decisions), then the threshold that minimises the cost occurs at a point where the slope is a certain function of the costs and the priors [20]. In fact if the costs of correct classifications are 0.0 and the costs of mistakes are equal then the required slope is 1.0. If, on the other hand, the priors are unknown a criterion known as "minimax" is used where it is assumed the priors are such that they make the performance as bad as possible. The criterion minimises the maximum possible risk. The threshold in this case is the intersection of the ROC with a line passing through the point $(0.0, 1.0)$ having a slope which is a certain function of the costs [22]. With the costs assumed as above, the slope is -1.0. As the prior probabilities in a scene are unknown beforehand, we will assume the latter criterion where the best threshold is at the intersection of the ROC with the negative diagonal. In order to compare ROCs we will present the proportion of misclassified pixels at the point on the curve satisfying the minimax criterion.

Figure 8.3 shows four images with various characteristics that need to be handled in a method of classification. The blue display channel is used to illustrate the NIR component and so all objects having pronounced NIR reflection will appear bluish. (a) sel1 is an image of cauliflower together with rape seedlings. Plant residues, which should not be classified as vegetation, include the two dead cauliflower leaves in the right part of the image and some two-day old cut grass. (b) sel2 consists of barley seedlings together with straw and two stones in the upper left corner. (c) sel3 is a cauliflower plant with pronounced shading. Correct classification of the shaded areas is especially interesting. (d) sel4 addresses the problem of a small population of vegetation and classification of the three stones in the middle of the image.

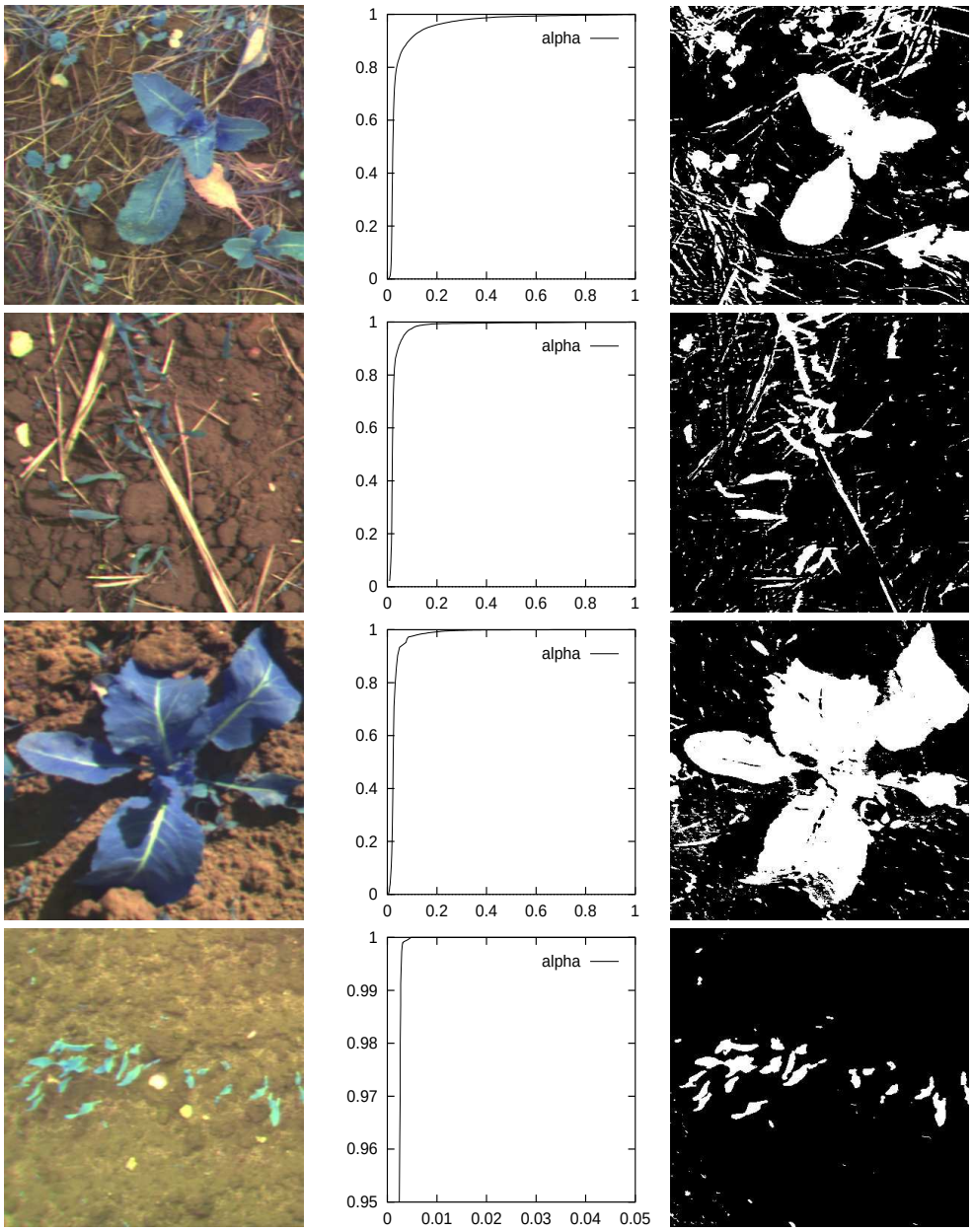


Figure 8.3: *Four test images, ROCs and classification results. The left column shows the test images, the centre column shows the ROCs, and the right column is the classification from the dedicated sensor. Note that the ROC for image sel4 has a different scale.*

In order to produce ground truth, the four images were carefully classified into live vegetation and background by visualising them as a colour image with the NIR channel seen as blue. This was primarily done manually, using a graphics tablet, but assisted by the colour classification tools in the Adobe Photoshop package. Areas suggested by the tools were accepted, rejected, or (more likely) modified

by hand until the resulting classification was visually acceptable. Like all ground truth methods, the human observer could not be certain about all decisions. As well as the two classes, we therefore had a third "uncertain" class. The proportions of "vegetation", "background", and "uncertain" are given in Table 8.1.

Table 8.1: *Proportion of "vegetation", "background", and "uncertain" for the test images.*

	Proportion in %		
Image	Vegetation	Background	Uncertain
sel1	54.6	34.6	10.8
sel2	4.7	92.4	2.9
sel3	37.0	54.4	8.6
sel4	3.0	94.3	2.7

When comparing automatic against human classifications, the "uncertain" category was ignored. The ROCs, which all have the expected shape, are shown in the middle column of Figure 8.3. The right hand column of Figure 8.3 shows classifications at thresholds corresponding to the minimax criterion. In all four images the live vegetation is generally classified correctly, but there are some background components, especially the edges of straw and dead grass and occasionally soil, classified as "vegetation". In image sel1, the three dead leaves in the right half are correctly classified as "background". The dead leaf in the left half is partially classified as "vegetation" and, in fact, there is some live tissue present. Most of the straw in image sel2 is classified correctly except for some of the edges. Edge effects could be caused by slight mis-registration of the colours caused by small shifts in the sensor position when rotating the filter wheel. The two stones on the left of sel2 are partly mis-classified. Image sel3 contains significant shadows, both on the soil and on leaves, which are generally well classified. Note that small plants in the shadow of the leaf at the left side (barely visible) are correctly classified. The stones in sel4 have been correctly classified and nearly all of the small areas on the left side of the image classified as "vegetation" are, on close inspection, really vegetation. The proportions of pixels misclassified at the thresholds corresponding to the minimax criterion are shown in Table 8.2 along with the thresholds used.

Table 8.2: *Proportion misclassified at threshold corresponding to minimax criterion for dedicated sensor. Figures in parentheses are the thresholds.*

Image	prop. mis. (t'hold)
sel1	0.092 (42)
sel2	0.061 (48)
sel3	0.056 (31)
sel4	0.003 (29)

8.5 Alternative Combinations of Channels

8.5.1 Ratio methods

As vegetation reflects more in the NIR than soil but less in the red it is appealing to classify on the basis of red/NIR ratio. [10, 8, 5] A similar argument can be applied to red/green ratio. Although the difference in reflectivity in green is not as marked as in NIR, red/green ratio has the advantage that both bands are available in conventional colour cameras. Both ratio methods can be depicted on a diagram of NIR *vs* red chromaticity. It is easy to show that contours of constant red/NIR ratio are straight lines radiating

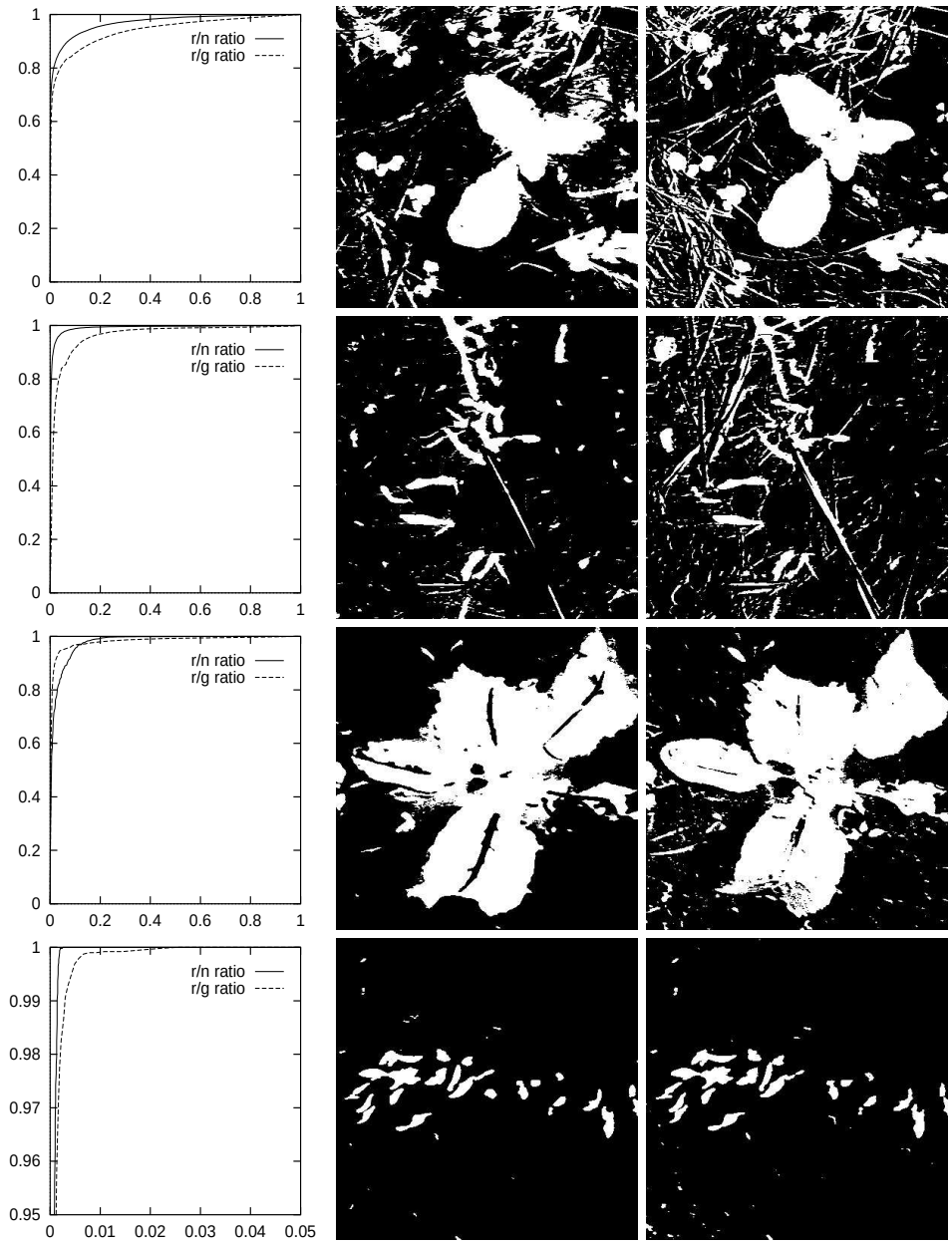


Figure 8.4: *ROCs and classification results for the four test images using two different ratio methods. The left column shows ROCs for red/NIR and red/green ratios, the centre and right columns show the classifications from the red/NIR and red/green methods respectively. Note that the ROC for image sel4 has a different scale*

from the origin, whereas contours of constant red/green ratio are straight lines radiating from the point (0.0,1.0).

The ROCs and classifications in Figure 8.4 (summary of performance in Table 8.3) show that the red/NIR ratio method has the potential of performing a good classification, often better than the alpha method. The reason for this is as follows. The dichromatic reflection model predicts (with one illuminant) that

pixels will be distributed along a line in a chromaticity diagram joining the plant body chromaticity to the illuminant chromaticity. Our previous work [2] shows distributions for the four test images and Figure 8.2(b) shows a typical example. In practice there will usually be relatively few highlights as they only appear when the illuminant is concentrated (e.g. on sharply curved surfaces) or when the surface is very smooth. As this is relatively rare, this means that the pixel distribution will be concentrated at the body locus end of the possible line. Thus, in many cases, a good direction and position for a linear discriminant would be at right angles to a line joining the vegetation and background clusters about halfway between them. This is approximately provided by the red/NIR ratio and it is this characteristic that allows the red/NIR ratio method potentially to give a good classification. Whether this potential can be realised will be discussed below. In cases that would give a significant proportion of highlights (e.g. if the scene were wet) the method would not be expected to perform so well as the highlight pixels would fill in the gap between the two clusters. It should also be noted that colour balancing of the sensor channels in a different place will change the positions of clusters and may change the suitability of the discriminant provided by the red/NIR ratio.

The performance, as measured by the percentage misclassified, Table 8.3, is usually worse with red/green ratio than with red/NIR ratio. This is not surprising as the directions of the discriminant provided by the red/green ratio (radiating from point 0.0,1.0) tend to merge the vegetation and background clusters rather than split them. An interesting exception to the relative performance of the red/green ratio is in the shadowed regions of sel3. While the shadowed areas of plants are correctly classified by all three methods, the shadowed soil is not well classified by the red/NIR ratio method. However, it may not be appropriate to draw conclusions from such low intensity regions as all three methods rely on ratios which will be subject to noise in these areas. Indeed Klinker [12] recommends disregarding all pixels below a certain threshold.

Table 8.3: *Proportion misclassified at threshold corresponding to minimax criterion for two ratio methods. Figures in parentheses are the thresholds.*

Image	red/NIR ratio	red/green ratio
sel1	0.088 (1.26)	0.128 (1.07)
sel2	0.036 (1.52)	0.092 (1.08)
sel3	0.080 (0.80)	0.045 (1.00)
sel4	0.002 (1.41)	0.004 (0.92)

8.5.2 Single band methods

Use of a single waveband is appealing because of simplicity. Using the NIR [15, 6, 17, 7] intensity is intuitive as vegetation has a considerably higher reflectivity than soil in the NIR. The same argument could be applied to green intensity but the reflectivity difference is not so great. Also, a sensor could be engineered using a monochrome camera and a suitable single filter. However, there are often objects in the background (such as the stones in sel2 and sel4) that also have a high reflectivity in these bands. The brighter background components can also have similar intensities to the darker parts of the vegetation (e.g. in sel3).

Thus significant mis-classifications can occur (Figure 8.5, middle and right column). In addition, the one dimensional measurement spaces do not generally have sufficient discriminating power to deal with dead vegetation (see the dead leaves in sel1). However, the shadows on soil are classified correctly (sel3) simply because they are darker. Unfortunately this is not true for the shadows on vegetation (sel3). Although the performance of the NIR intensity method is reasonable, the green intensity does not discriminate so well as shown by the ROCs in Figure 8.5 and the summaries in Table 8.4. This is to be expected due to the relative reflectivities of the two components (vegetation and soil) in the different wavebands. In fact

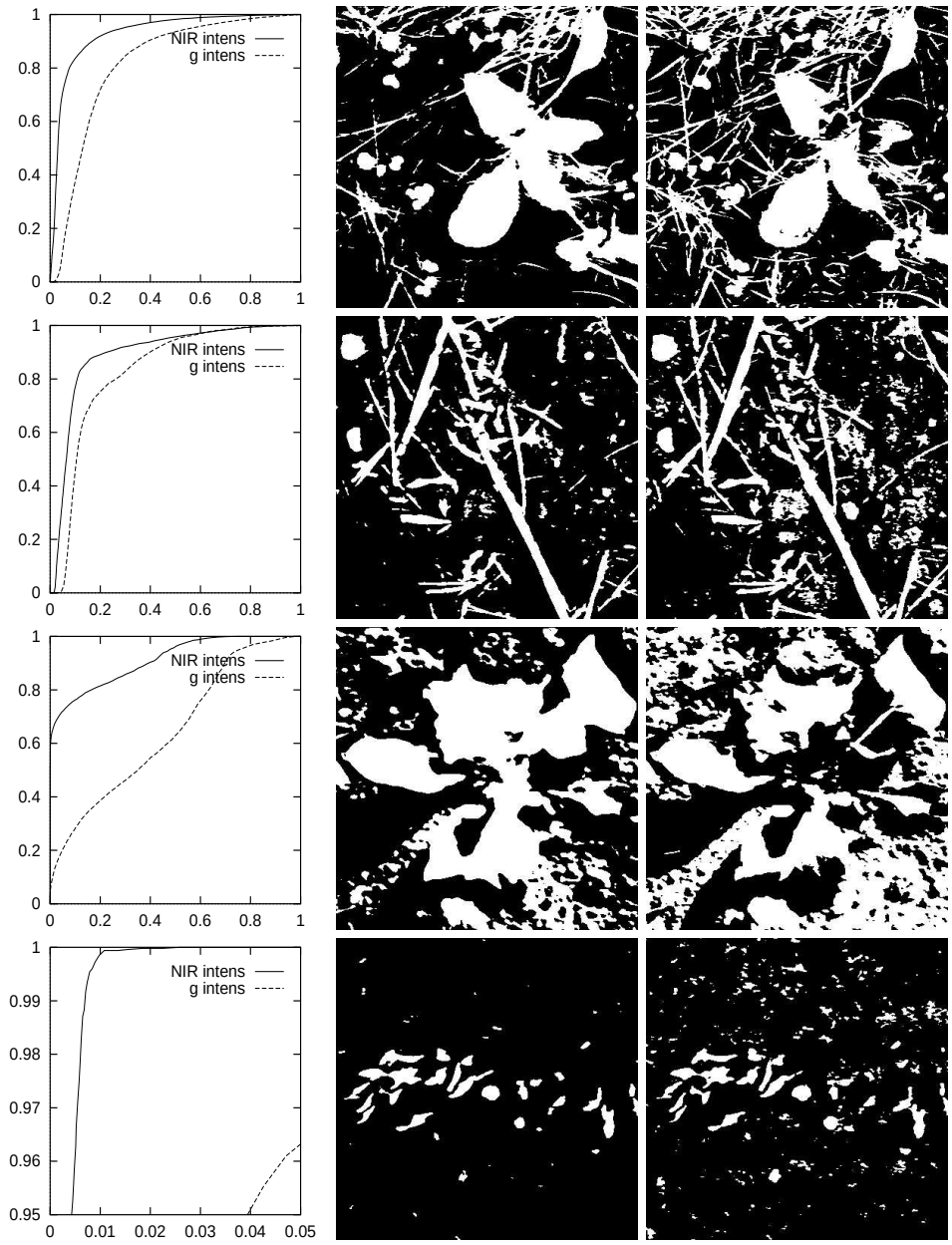


Figure 8.5: *ROCs and classification results for the four test images using two different single band methods. The left column shows ROCs for NIR and green intensities, the centre and right columns show the classifications from the NIR and green methods respectively. Note that the ROC for image sel4 has a different scale*

the classification of sel3 by green intensity is hardly better than random chance (which would be depicted by a diagonal line from (0.0,0.0) to (1.0,1.0) on the ROC [22]).

Table 8.4: *Proportion misclassified at threshold corresponding to minimax criterion for two single band methods. Figures in parentheses are the thresholds.*

Image	NIR intensity	green intensity
sel1	0.130 (70)	0.232 (78)
sel2	0.140 (58)	0.231 (70)
sel3	0.187 (59)	0.432 (63)
sel4	0.007 (87)	0.043 (139)

8.6 Choice of Thresholds

The previous two sections have discussed how well the various methods, either physically based or not using any physical understanding, could possibly classify images. This section addresses the problem of choosing the parameters of a discriminator, i.e., in our one-dimensional methods, the threshold.

Figure 8.6 and 8.7 shows the histograms generated for the four test images by the five methods. For the first three methods (fig. 8.6) the vegetation mode will be at the left end of the histogram while for the last two (fig. 8.7) it should be towards the right. The concentration effect produced by the alpha method is demonstrated clearly. In the cases (sel1 and sel3) when there is a reasonable proportion of vegetation there is a sharply peaked mode. From the measured CCTs the modes are expected at 23° for images sel1,2 and 4 and at 27° for image sel3. This agrees well with the data. Especially notable are the peaks in images sel2 and 4 where there is a very small proportion of vegetation (Table 8.1). We believe that it should be possible to find a good threshold automatically from the histograms, especially as the position of the vegetation mode can be predicted from the physical model, even with small amounts of vegetation present. In order to test this hypothesis, many methods of threshold selection would need to be tested along with many more images (with ground truth). This work is outside the scope of the present paper.

As the discriminant produced by the red/NIR ratio is an appropriate one for many images, there are similar peaks for this method for images sel1 and 3. However, the concentrating effect is no longer present. In fact the red/NIR discriminant sweeps the vegetation cluster in just about the optimum direction to spread rather than concentrate it. Thus the peaks of the vegetation modes for images sel2 and 4 are very small. This will cause significant problems for any automatic method.

Generally the best thresholds for alpha and red/NIR ratio (Tables 8.2 and 8.3) are close to characteristic features of their histograms (i.e. the tail of the vegetation mode or the trough between modes). However, an exception is in image sel3 where the shadow parts on the soil are contained partly within the right hand side of the vegetation mode continuing towards the trough between the two modes. Although the alpha method could classify the shadows correctly and better than the red/NIR ratio, it would be difficult to find the best threshold automatically for either method.

As the discriminant produced by the red/green ratio (a line radiating from the point (0.0,1.0) in the chromaticity diagram) has a component in the direction of the major axis of the vegetation cluster, there will be some concentrating effect of the vegetation mode. This can be detected with a low proportion of vegetation (sel2) but is not as great as with the alpha method. The predictability of the best thresholds for red/green ratio is not quite as good as with the other multi-band methods with the interesting exception of sel3. In this image the best threshold (which also classifies the shadows correctly) is at the trough between the first two modes. However, the robustness of using ratio methods at low intensities is questionable as discussed above.

The histograms for the two single band methods are far from easy to analyse, figure 8.7. Although Figure 8.5 shows that reasonable classifications are possible (more so with red/NIR ratio), the best thresholds do not generally correspond with histogram features. A possible exception is image sel1 where there is a slight trough. With small amounts of vegetation there are no real histogram features that correspond at all with the vegetation.

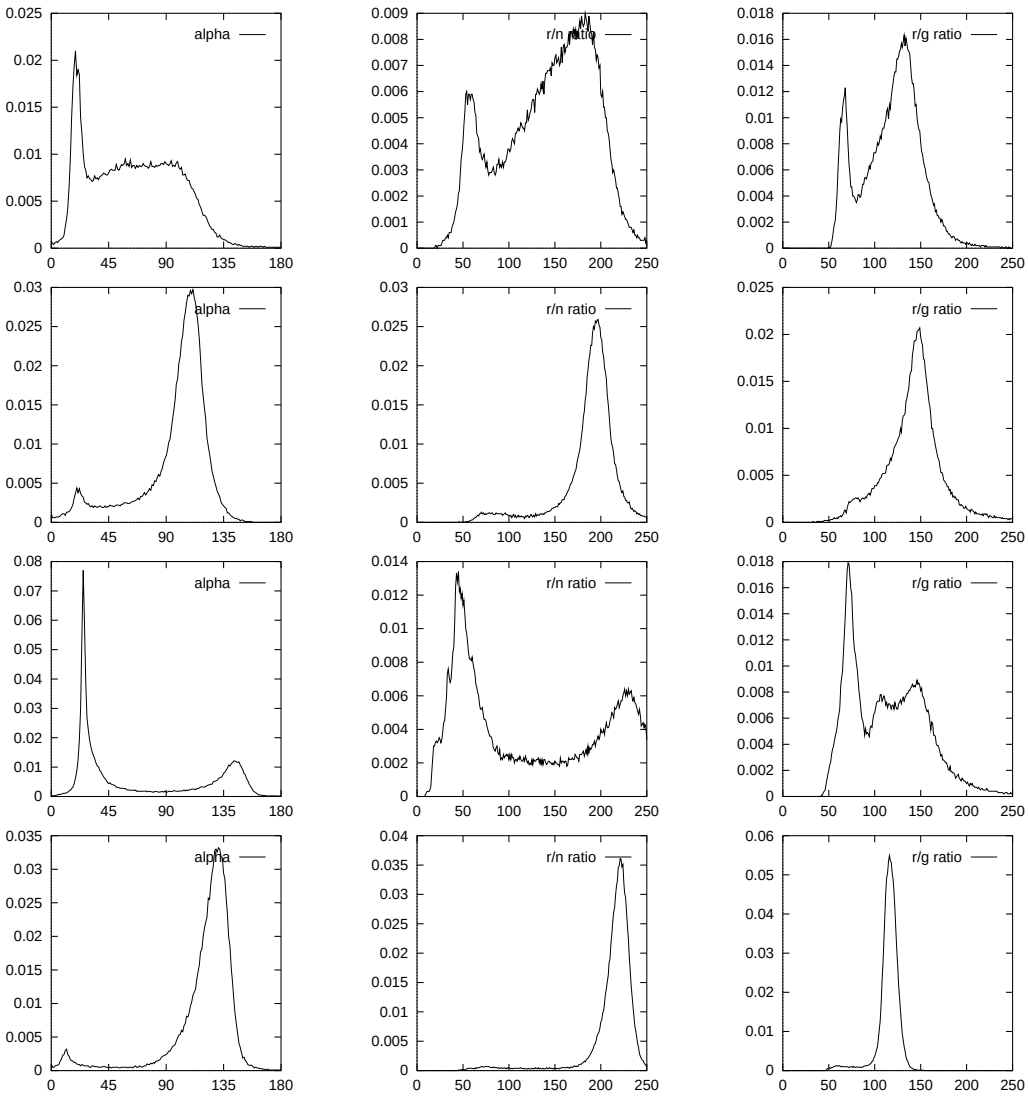


Figure 8.6: *Histograms for the four test images. The rows (top to bottom) correspond to images sel1, sel2, sel3, sel4 respectively. The columns (left to right) correspond to methods alpha, r/n ratio, r/g ratio, respectively. The horizontal axes are degrees (alpha), ratio×100 (ratios).*

8.7 Conclusions

A method for classifying vegetation from background using data from a previously reported imaging sensor is proposed. The sensor uses three bands, red, green, and near infra-red (NIR). An important aspect of the method (known as the "alpha method") is that it is based on an understanding of the physics of the image formation and capture.

The method is compared with two other multi-band methods (red/NIR and red/green ratio) and two single-band methods (NIR and green intensity) using four images previously classified by hand. The images all have characteristics that make them rather difficult to analyse.

The classification problem is addressed in two parts: How successful could classification be at its best,

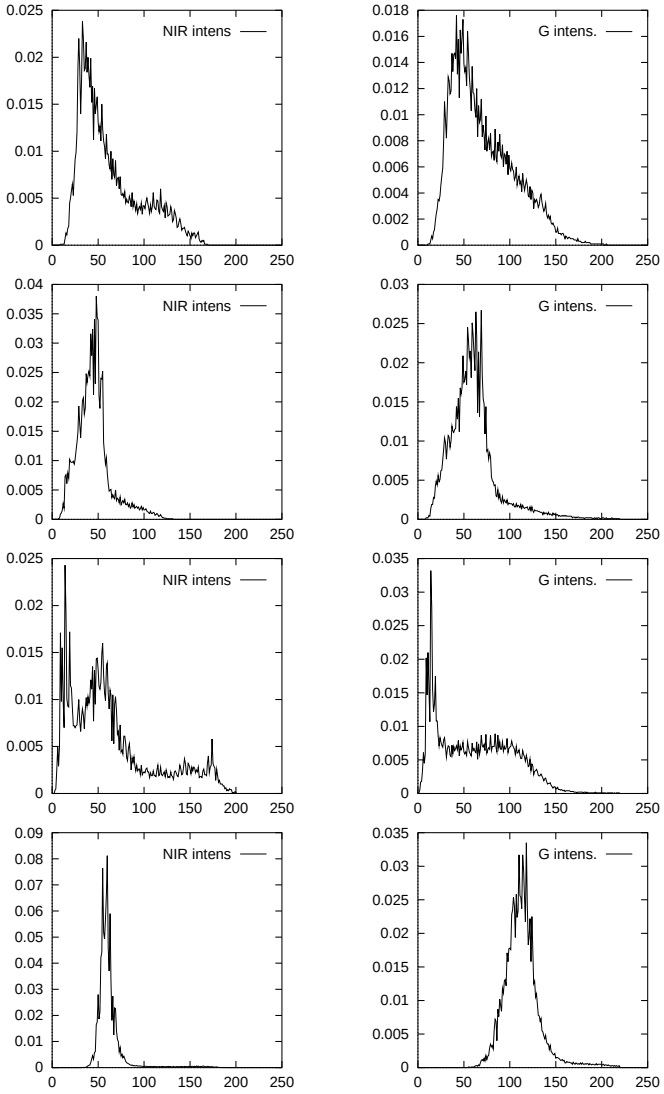


Figure 8.7: *Histograms for the four test images. The rows (top to bottom) correspond to images sel1, sel2, sel3, sel4 respectively. The columns (left to right) correspond to methods NIR intensity, green intensity respectively. The horizontal axes are digitized intensity value (single band)*

and whether it will be practically possible to choose a classification threshold. On the first part the three multi-band methods all perform reasonably well, alpha and red/NIR are similar overall with red/green slightly worse. NIR intensity also gives a reasonable classification except that bright parts of soil can be confused with darker parts of vegetation and dead vegetation is not classed as background. Green intensity performs markedly worse.

Although the dichromatic reflection model (the basis of the alpha method) predicts that the vegetation class will form a line in red/NIR chromaticity space, in practice it tends to cluster at one end. This explains the good performance of red/NIR ratio as the direction of the effective discriminant produced by the method is suitable for this clustering. The alpha method does however concentrate the vegetation cluster as its discriminant is aligned with the cluster's major axis and so performs well in situations where there is relatively little vegetation. In contrast the red/NIR discriminant is roughly at right angles to the

major axis and so the vegetation cluster is diffused. Thus, on the question of whether it will be possible to choose a threshold automatically, the alpha method appears to have advantages in images containing a small proportion of vegetation. In addition, it is possible to predict where the vegetation mode should be with the alpha method which will help in choosing a threshold. In the two single band methods the best thresholds do not correspond well with any characteristic features of the histograms and so it will be difficult to find any method of automatic analysis. The NIR intensity method performs better than green intensity but neither method may perform well unless the image is comparatively simple.

The relative benefits of each method must be judged in combination with their practicability. Single band methods require only a monochrome camera. All the multiband methods require more complex sensing arrangements, but the red/green ratio method can be implemented with a common (and therefore cheap) colour camera. In order to take advantage of the physical understanding, the alpha method requires a measurement of the CCT.

Of all the methods considered, the two that perform best are the alpha and the red/NIR ratio methods. There would seem to be scope in future work to combine the attractive features of both approaches, possibly by modeling the statistics of the distributions of pixels and combining this with the prior knowledge provided by the dichromatic reflection model.

8.8 Acknowledgments

The work of John Marchant and Christine Onyango was funded by the Biotechnology and Biological Sciences Research Council (UK). The work of Hans Joergen Andersen was funded by the Danish Research Council under the Optimal research programme.

References

[1] H.J. Andersen and E Granum. Classifying the illumination condition from two light sources by colour histogram assessment. *Journal of the Optical Society of America A*, 17(4):667–676, April 2000.

[2] H.J. Andersen, C.M. Onyango, and J.A. Marchant. The design and operation of an imaging sensor for detecting vegetation. *Accepted for publication: Int. J. Imaging Systems and Technology*, 2000.

[3] Anonymous. Colorimetry. Technical Report 2. edition, Commision Internationale de L'Eclairage (CIE), 1986.

[4] J.V. Benlloch, A. Sanchez, S. Christensen, and M. Walger. Weed mapping in cereal crops using image analysis techniques. In *AgEng96*, volume 2, pages 1059–1060, 1996.

[5] H.R. Biller. Reduced input of herbicides by use of optoelectronic sensors. *Journal of Agricultural Engineering Research*, 71:357–362, 1998.

[6] R. Brivot and J.A. Marchant. Segmentation of plants and weeds using infrared images. *Proceedings of the Institution of Electrical Engineers, Vision, Image and Signal processing*, 143(2):118–124, 1996.

[7] E. Franz, M.R. Gebhardt, and K.B. Unklesbay. Algorithms for extracting leaf boundary information from digital images af plant foliage. *Transaction of the ASAE*, 38(2):625–633, Mar 1995.

[8] R.J. Hagger, C. J. Stent, and S. Isaac. A prototype hand-help patch sprayer for killing weeds, activated by spectral differences in crop/weed canopies. *Journal of Agricultural Engineering Research*, (28):449–358, 1983.

[9] J.A. Hanley and B.J. McNeil. The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology*, 143:29–36, 1982.

[10] A.W. Hooper, G.O. Harries, and B. Ambler. A photoelectric sensor for distinguishing between plant material and soil. *Journal of Agricultural Engineering Research*, (21):145–155, 1976.

[11] D.B. Judd, D.L. MacAdam, and G.W. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54, 1964.

[12] G. J. Klinker. *A Physical Approach to Color Image Understanding*. A.K. Peters Ltd, 1993.

[13] H.-C Lee, E.J. Breneman, and C.P. Schulte. Modeling light reflection for color computer vision. *IEEE trans. on Pattern Analysis and Machine Intelligence, PAMI*, 12(4):402–409, 1990.

[14] J.A. Marchant and R. Brivot. Real-time tracking of plant rows using a Hough transform. *Real-Time Imaging*, 1:363–371, 1995.

[15] H.J. Olsen. Determination of row position in small-grain crops by analysis of video images. *Computer and electronics in agriculture*, 12(2):147–162, Mar 1995.

[16] F. Pla, F. Ferri, and M. Vicens. Colour segmentation based on a light reflection model to locate citrus for robotic harvesting. *Computers and Electronics in Agriculture*, 9:53–70, 1993.

[17] J.F. Reid and S.W. Searcy. An algorithm for separating guidance information from row crop images. *Transaction of the ASAE*, 31:1624–1632, 1988.

[18] J.M. Roger, E. Molto, G. Rabatel, and J. Blasco. Design of a robotized, non-chemical weed controller. In *Bio-Robotics'97*, pages 229–235, 1997.

[19] S.A. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985.

[20] J.A. Swets. ROC analysis applied to the evaluation of medical imaging techniques. *Investigative Radiology*, 14:109–121, 1979.

[21] S. Tominaga. Dichromatic reflection models for a variety of materials. *COLOR research and application*, 19(4):277–285, 1994.

[22] H.L. van Trees. *Detection, estimation, and modulation theory*, volume 1. John Wiley & Sons, 1968.

[23] E. Vrindts and J de Baerdemaerker. Optical discrimination of crop, weed and soil for on-line weed detection. In J.V. Stafford, editor, *Precision Agriculture '97*, volume 2, pages 537–544, 1997.

Chapter 9

Summary and Conclusion

The overall ambition of this thesis is to deliver results that may contribute to a reduction in the use of herbicides within outdoor weed control. Recent years concerns both from the consumers and the international society have pointed to the need for development of new agricultural practices and techniques that may lead to a reduction in use of herbicides. Within this framework the focus of the thesis has been to investigate the potential support of computer vision within outdoor weed control methods for reduced use of herbicides.

9.1 Summary

The thesis is divided into two main parts, *Computer Vision in Relation to Weed Control* and *Physics Based Analysis of Outdoor Images of Vegetation*. The first investigates the relation between computer vision and weed control. The second investigates and develops computer vision methods adaptable to varying daylight conditions. The subject is approached in two different ways first be the use of a "standard" RGB camera and second be development of a dedicated sensor for detecting vegetation.

9.1.1 Part I, Computer Vision in Relation to Weed Control

The Concept Accessibility
In this part an analysis of the different ways computer vision may be of support for weed control was found to be closely related to the *spatial resolution* and *spatial selectivety* required for the control treatment. By these two characteristics we divided the possible weed control treatments into the following five categories, global, site specific, site adaptive, and plant differential on row level or individual plant level.

The first two categories are solely relying on a priori knowledge about the need for weed control but differs in resolution. Global treatment is the "traditional practice" where a given field is regarded as one unit whereas site specific treatment divides the field into sites that may be treated individually a method that is emerging by support of Global Position Systems (GPS).

The next three are all characterized by the properties that they use real-time/on-site information to adapt to the current need for local weed control. All three treatments may potentially be supported by computer vision but with different objectives. The first category site adaptive needs information about the current weed pressure to adjust the treatment to current need. The last category plant differential treatment is characterized by the feature that it needs the spatial information about the location of the plants. The category is further divided into two, detection of row or individual plant structures respectively, depending on which level the culture plants position need to be determined. Detection of row structures needs less spatial resolution that detection of individual plant structures.

From the analysis of characterizing the objectives for support of computer vision we introduced a new concept *accessibility* for description and assessment of weed control treatments in a given context. Accessibility is defined by two linear discriminant functions one describing if it is possible to detect the culture plants in row structures and the other for detecting the individual plants. By this we end up with three main areas of accessibility, area 1) where neither row or plant structure is accessible, area 2) where the row structure is accessible, and area 3) where the plants may be located individually.

Looking at the three areas it can been seen that there is a correspondence between these and the objectives for the support of computer vision. Hence the concept may be used both for description of the weed control treatment and for assessment of the control treatment in a given context as all parameters influencing on the accessibility has to be in the area at which we want to operate. For instance if a parameter makes it impossible to locate the plants row structures it is impossible to perform plant differential treatment at row level without altering the parameter so it becomes possible to locate the culture plants in rows.

In this thesis we have developed the concept of accessibility and given one examples of how it may be used. However, if one really want to go into the depth of the potential use of the concept this will call for a major research project in it self.

Survey, Computer Vision Support Within Row Crops
In the other chapter of Part I a survey of the literature within the field was presented. The survey is organized according to the objectives for the support of computer vision, estimation of weed density, and detection of row and plant structures. In extension to these three categories defined by the accessible of the plant structure two categories for studies of more basic investigation was introduced: segmentation of living plant material from non-plant background, and identification of individual plants.

The survey shows that many computer vision methods and techniques have been investigated within the application of weed control but only few have taking their point of departure in the process "weed control". One exception from this general picture is though within detection of row structures. The research within this area have during the last years reached a level where implementation at farm level may be regarded as realistic.

From the survey we also concluded that one of the limited factor for almost any reported study where the constrained conditions at which the investigated methods where evaluated. Especially, adaptation to the varying conditions for the image formation was an area that needed more investigation for development of robust methods.

One of the major sources of variation for the image formation for images acquired under outdoor conditions may be recognized as being the illumination condition. Conditions that may change from direct sunshine to a clear blue sky as illumination source, a change that radically will alter the colours and conditions for the image formation.

As a consequence of this conclusion the focus of the project was concentrated on finding methods that will make it possible to implement computer vision robust under outdoor daylight conditions.

9.1.2 Part II, Physics Based Analysis of Outdoor Images of Vegetation

In Part II, we investigated two different approaches for support of weed control by computer vision adaptable towards changing daylight conditions.

We approached the problem through a comprehensive analysis of the physical conditions for images acquired under changing illumination conditions. For modeling of the reflected light from dielectrical objects we used the dichromatic reflection model together with CIE daylight standard, radiation from a black body, as weel as camera and reflectances characteristics. We showed that by taking advantage of this a priori model information it is possible to develop computer vision methods adaptable towards changing illumination and in particular changing daylight conditions.

Two Illumination Sources
As daylight may regarded as consisting of two illumination sources sun- and skylight, respectively, we initially investigated and proposed a method for assessment of illumination conditions of dielectrical objects illuminated by two light sources. We introduced the concept *primary* reflection, which is the "pure" body or surface reflection, respectively from a dielectrical object. As the reflected light from a given position of a dielectrical object may be expressed as an additive mixture of the primary reflections we where able to derive a taxonomy of possible distributions of the objects' chromaticities.

The taxonomy covered an object illuminated by two light sources which produced surface or body reflections from it. It categorize the distribution of the objects' pixel points chromaticities according to the number of primary chromaticities produced by the light sources. As the reflected light may be regarded as additive mixtures of the objects' primary reflections the distribution of the pixel points chromaticities will be limited within the area enclosed by the primary chromaticities.

In the extreme case where both illumination sources give body and surface reflections the chromaticities of the objects' pixel points will be distributions within the tetragon defined by its four primary reflections. In the other extreme case where both illumination sources have the same spectral composition and will only produce body reflection the chromaticities will be distributed around a single point. Between these two extremes the pixel points' will located within a triangle or along lines.

An important category of illumination condition was identified to be that of daylight. This may be considered as being consisting of a uniform and a diffuse light source. The uniform light source is due to the direct light which is able to produce surface and body reflections. The diffuse light source is the ambient skylight which only is able to produce body reflection. Hence the chromaticities of a dielectrical object illuminated by daylight will be distributed either within a triangle defined by the three primary chromaticities due to the surface and body reflection of the direct sunlight and the body reflection of the skylight or in absence of the direct sunlight in a single point due to the primary body reflection produced be the skylight.

The taxonomy was used in two experiments. The first assessed the reflection of six uniform objects illuminated by two known light sources in controlled laboratory conditions. The second assessed the illumination of realistic images in field condition of Barley plants monitored from early morning to late evening in direct sunshine, over cast, and shaded conditions.

The results showed that using information about an objects' pixel points distribution in the chromaticity plane it is possible to make assessment of the current illumination conditions. This information may in a active computer vision system for adjustment or change of the method to adapt to the current illumination conditions.

Adaptive Segmentation.
On basis of the study of illumination in a two illumination source situation we developed a method for segmentation of vegetation from background of soil, which is invariant towards spectral changes of daylight.

In addition to the chromaticities of the objects' we also introduced the intensity as a parameter. Hence a pixel point is represented by its r and g chromaticities and the intensity in a so called "rgI" space. Addition of intensity made the method possible to classify background areas in shade by the vegetation correctly. These areas will be biased towards the colour of the vegetation as the light is first transmitted through the plants. Instead they may be classified by the feature that they will have a lower intensity than the directly illuminated vegetation.

The method was used on four images captured at varying CCT of the daylight and with "hard" shadows on the soil surface from the plant leaves. To get ground truth the images where manually segmented into three classes vegetation, background, and uncertain. The result showed that the method was capable of adapting to varying CCT of the daylight and that inclusion of the intensity as parameter made it possible to correctly classify areas in shade.

Dedicated Sensor
In this study we looked at the benefits of designing a sensor dedicated for the problem of segmenting vegetation from background.

The philosophy behind the design was to take advantage of the reflectance characteristics of the vegetation and the most likely background materials for determination of the spectral sensitivity of the sensor. Naturally soil was the most important background material to investigate, however, other artifact as plant residues, stones, manure etc. was also included in design of the sensor.

Discrimination of soil from vegetation was achieved by using the pronounced reflectance of vegetation in the near infra red (NIR) (700-1100nm) and the depression in the red(630-700nm) due to the high content of chlorophyll. Hence by placing wave bands on either side of 700nm (i.e. in the red and NIR) a good contrast between for discrimination between the two objects was obtained. An approach that also is widely used within remote sensing.

For discrimination of organic materials also with high contents of chlorophyll as plant residues, peat soils, and manure a wave band in green (500-630nm) was included in the sensor, as these materials normally will appear less green than vegetation which have a pronounced reflectance in the green.

As a result of the analysis of the most likely materials to occur in a scene of vegetation the sensor was designed to have a wave bands in the green, red, and near infra red region. The performance of the sensor was assess in two different ways. First that the sensor response was as expected and secondly that it gave a good separation of the vegetation and background.

To be able to model the sensor behavior it was necessary to also model the characteristics of the daylight beyond CIE daylight standard into the near infra red region by the radiation from a black body. The results showed that modeling the illumination by a radiation from a black body may be regarded as sufficient for computer vision purposes.

For assessment of the sensor ability four images where manually segmented and their red and near infra red chromaticity plane where plotted. The cluster shapes clearly showed that the sensor provides a good basis for segmentation of vegetation even in scenes with freshly plant residues, stones and pronounced shading.

Evaluating Wave Bands Combinations
In the study we performed a more thoroughly investigation of the dedicated sensors capability for segmentation of vegetation from background. We compared two wave bands ratios red/NIR and red/green, two single green and NIR, and the physical based method "alpha" which is based on the location of the primary reflections of the vegetation.

The introduced alpha method was based on the modeling of the sensor. A drawback of method was that it needed information about the correlated colour temperature of the daylight. However, the method introduced in chapter 6 could easily be adjusted to the sensor and would due the well separated cluster structure provided by it perform as good as the alpha method or possibly better.

To compare the different methods the receiver operating characteristic curve was used to evaluate the classifications on realistic manually segmented images. The results of the analysis showed not surprisingly that both single bands method had a poor classification performance. The two band ratio and the alpha method all perform reasonable well with the red/green ratio slightly the worst.

Looking both at classification rate and the methods ability to produce a histogram with well defined detectable modes that corresponds with the "best" threshold the alpha method was superior to all the other methods. Even in images with a small amount of vegetation.

In general the investigation showed that utilizing the physics behind the image formation is a way forward for development of robust sensors and methods for segmentation of vegetation from background.

9.2 Conclusion

Colour Vision
In this thesis we have demonstrated that using a priori knowledge and modeling of the image formation may be of significant importance for development of computer vision methods and sensors for robust analysis of colour information captured under varying illumination conditions.

We have investigated and developed methods for:

- assessment of illumination conditions from image information
- segmentation of colour images invariant to spectral changes of the illumination
- design and use of a dedicated sensor

For assessment of the current illumination condition and segmentation of images invariant to spectral changes of the illumination, we have showed that modeling of the distributions of the colour reflections of objects' based on the intrinsic information about the physics behind the image formation may be used for development of computer vision methods.

For development of dedicated sensors we have demonstrated that using information about the reflectance characteristics of the objects' of interest in the scene may be used for design of the characteristics of the sensors' spectral sensitivity.

All these studies open general perspectives for development of new colour analysis methods and for outdoor computer vision methods in particular .

Application Outdoor Weed Control
Within the application of weed control we first through an analysis found the relation between computer vision and weed control and from this analysis introduced the concept "accessibility". This concept may be used as tool in further development of agricultural systems supported by on-site/real-time information as it has the ability both to describe the objectives for the weed control treatment and for assessment of parameters influencing on. Hence, it is able to serve as general basis for the control of the weed control treatment.

For the developed computer vision methods we demonstrated that they all are applicable within the field of outdoor weed control. Especially encouraging it is that the proposed methods are able to operate on images with vegetation on an early growth stage. A property which could be of vital importance for future plant growth systems with minimal use of herbicides.

The results of the thesis emphasis the possible benefit of using model of the physical process involved in image formation. Future work may advance computer vision further by more detailed studies of reflections from vegetation. If so, we may soon see computer vision systems as standard components in agricultural equipment.

Computer Vision
&
Media Technology Laboratory

THE COMPUTER VISION AND MEDIA TECHNOLOGY LABORATORY (CVMT) IS PART OF THE INSTITUTE OF HEALTH SCIENCE AND TECHNOLOGY. CVMT WAS FOUNDED IN 1984 AS THE LABORATORY OF IMAGE ANALYSIS (LIA). THE MAIN RESEARCH AREAS OF THE LABORATORY ARE COMPUTER VISION, MULTIMEDIA INTERFACES VIRTUAL REALITY, AUGMENTED REALITY SYSTEMS, AND AUTONOMOUS SYSTEMS AND AGENTS.

CVMT HAS ESTABLISH RESEARCH COOPERATION WITH MORE THAN 30 INSTITUTIONS IN 16 DIFFERENT COUNTRIES.

CVMT IS HEADED BY ITS FOUNDER PROFESSOR ERIK GRANUM.

CVMT

Computer Vision
& Media Technology Laboratory
Aalborg University
Niels Jernes Vej 14
DK-9220 Aalborg
Denmark

Telephone: +45 9635 8789
TeleFax: +45 9815 2444

E-Mail: info@cvmt.dk
URL: http://www.cvmt.dk