University of Kassel

Statistical evaluation of texture analysis from the biocrystallization method: Effect of image parameters to differentiate samples from different farming systems

Dissertation
submitted in fulfillment for the degree of Doctor of Agricultural Science
(Dr. agr)
Department of Organic Food Quality and Food Culture

by
M.Sc. (Food Technology)
Aumaporn Meelursarn

Supervisors:
1. Prof. Dr. Angelika Ploeger
2. Prof. Dr. Hans-Peter Piepho

Witzenhausen, December 2006
ERKLÄRUNG


Witzenhausen, 21.12.06
ACKNOWLEDGMENTS

I am very grateful to my supervisor Prof. Dr. Angelika Ploeger for her guidance and encouragement of my research efforts. In addition, I would like to thanks Prof. Dr. Hans-Peter Piepho, Prof. Dr. Peter von Fragstein and Prof. Dr. Ton Baars for his recommendation on my dissertation examination.

I would like to express my cordial deeply sincerely thanks to Dr. Nicolaas Busscher, who is very important for me and assists me until graduate, for his very kind counsels and encourage all of my research efforts. Dr. Johannes Kahl for any kind of guidance and sample information, Mrs. Gaby Mergardt for her responsibility to generate biocrystallization images.

I am grateful to Prof. Dr. Hans-Peter Piepho and Mr. Bertam Schäfer for their very kind recommendation in statistical model, statistical calculation via SAS program as well. Additional, Assoc. Prof. Dr. Jens Michael Carstensen for his useful instruction on texture analysis. Prof. Dr. Peter von Fragstein and Dipl. Ing. Agr. Michael Fleck for providing the samples.

I would like to thank Dipl. Oec. troph.(FH) Monika Roeger who assisted me in many occasions. And all department staffs, Triangle cooperation researchers as well for any kind support.

I would like to thank M. Eng. Phatcharasak Arlai for his great encouragement and motivation. Thanks also Thai and German friends for providing any support to me during study in Germany.

Thanks to Royal Thai government for the financial support during my study. Thanks to Asst. Prof. Dr. Niwat Klinngam and Nakhon Pathom Rajabhat University for all concerning to study in Germany.

Finally, I would like to deeply thank my family, especially my mum and brother, for their support and inspiration during this time.
Table of contents

1 Introduction .......................................................................................................................1
  1.1 Motivation ..................................................................................................................1
  1.2 The biocrystallization method ..................................................................................2
  1.3 The biocrystallization evaluation ..............................................................................2
    1.3.1 Visual evaluation .................................................................................................2
    1.3.2 Computerized image analysis .............................................................................3
  1.4 Objective of study ....................................................................................................3
  1.5 Outline of the thesis ..................................................................................................4

2 Method description .........................................................................................................5
  2.1 Biocrystallization method .......................................................................................5
    2.1.1 The biocrystallization phenomena ......................................................................5
    2.1.2 The biocrystallization method ............................................................................6
  2.2 The main factors of the biocrystallization method ......................................................9
  2.3 Application of biocrystallization method ..................................................................10
  2.4 Evaluation of biocrystallization method ..................................................................12
    2.4.1 The visual evaluation ..........................................................................................12
    2.4.2 The computerized image analysis .....................................................................13

3 Theory .............................................................................................................................14
  3.1 Repeated measurement statistical model; mixed effect model ...............................14
    3.1.1 Introduction .........................................................................................................14
    3.1.2 The principal concept of mixed effect model ......................................................15
    3.1.3 The nested and crossed effects in the experiment ..............................................17
    3.1.4 Setting up the mixed effect model .....................................................................20
3.2 Computerized image analysis

3.2.1 Introduction

3.2.2 Texture analysis

3.2.3 Texture statistics

3.2.3.1 First-order statistics

3.2.3.2 Second-order statistic

3.2.3.3 Gray level difference histogram (GLDH)

3.2.3.4 Gray level sum histogram (GLSH)

3.2.4 Histogram matching

3.2.4.1 Introduction

3.2.4.2 Image histogram

3.2.4.3 Principle of histogram matching

3.2.5 The color to gray level transformation

3.2.5.1 Introduction

3.2.5.2 RGB model

3.2.5.3 HIS model

3.2.6 Region of interest (ROI) and texture analysis

3.2.7 Texture analysis applications

4 Study methodology and result

4.1 Optimization of the statistical model

4.1.1 Introduction

4.1.2 Statistical model implementation

4.1.3 Statistical model verification

4.1.3.1 The comparison of ANOVA results between SAS and R program

4.1.4 The ACIA program investigation: old and new versions with refined statistical model
4.1.4 Summary of the optimization of the statistical model ........................................ 74

4.2 Investigate the effect of image parameters on texture analysis of biocrystallization image: ROI, color transformation and histogram matching ............................................................ 75

4.2.1 Introduction ........................................................................................................ 75

4.2.2 Effect of region of interest (ROI) on texture analysis ..................................... 77

4.2.2.1 Introduction .................................................................................................. 77

4.2.2.2 ROI investigation ....................................................................................... 77

4.2.2.3 ROI result of wheat DOC ......................................................................... 78

4.2.2.4 ROI result of wheat market ...................................................................... 82

4.2.2.5 ROI result of carrot from University of Kassel .......................................... 86

4.2.2.6 ROI result of carrot market ...................................................................... 91

4.2.2.7 Summary of the effect of ROI on texture analysis ..................................... 96

4.2.3 Effect of color transformation on texture analysis ........................................... 97

4.2.3.1 Introduction ............................................................................................... 97

4.2.3.2 Color transformation investigation .............................................................. 98

4.2.3.3 Color transformation result of wheat DOC ............................................... 98

4.2.3.4 Color transformation result of wheat market ............................................ 101

4.2.3.5 Color transformation result of carrot from University of Kassel ............... 107

4.2.3.6 Color transformation result of carrot market ............................................ 113

4.2.3.7 Summary of effect of color transformation on texture analysis ................ 117

4.2.4 Effect of histogram matching on texture analysis ......................................... 118

4.2.4.1 Introduction ............................................................................................... 118

4.2.4.2 Histogram matching investigation .............................................................. 118

4.2.4.3 Histogram matching result of wheat DOC ............................................... 120

4.2.4.4 Histogram matching result of wheat market ............................................ 124
4.2.4.5 Histogram matching result of carrot from University of Kassel ........... 129
4.2.4.6 Histogram matching result of carrot market ................................... 136
4.2.4.7 Summary of effect of histogram matching on texture analysis ......... 144

4.3 The relation of texture parameters and visual evaluation ...................... 146

4.3.1 Carrot ......................................................................................... 146

4.3.1.1 Carrot from University of Kassel ........................................... 146

4.3.1.2 Carrot market ......................................................................... 156

4.3.2 Summary of the relation of texture parameters and visual evaluation .. 172

5 Discussion with conclusions and future work ........................................ 173

5.1 Optimization of the statistical model .................................................. 174

Summary of the optimization of the statistical model ................................. 175

5.2 Investigate the effect of image parameters on texture analysis of biocrystallization image: ROI, color transformation and histogram matching ................................................................. 176

Summary of the investigate the effect of image parameters on texture analysis of biocrystallization image: ROI, color transformation, histogram matching ................................................................. 184

5.3 The relation of texture parameters and visual evaluation ...................... 186

Summary of the relation of texture parameters and visual evaluation ........ 189

5.4 Future work ..................................................................................... 190

6 Abstract ............................................................................................. 191

Zusammenfassung .................................................................................... 193

7 References ........................................................................................... 195

8 Appendix .............................................................................................. 202

8.1 Appendix A: Sample information ...................................................... 202

A.1: Wheat DOC ................................................................................... 202

A.2: Wheat market ................................................................................ 203
A.3: Carrot from University of Kassel ................................................. 204
A.4: Carrot market ........................................................................... 205

8.2 Appendix B: Ime 16 cases of statistical model with repeated
measurement (crossed effects) ........................................................ 206

8.3 Appendix C: Main programs to scan and calculate
the biocrystallogram image ............................................................... 207
C.1: Applied crystallogram image analysis (ACIA) program ................. 207

8.4 Appendix D: The coding of texture parameters .................................. 208

8.5 Appendix E: 14 criteria of carrot reference pictures ......................... 209

8.6 Appendix F: 14 criteria of wheat reference pictures ......................... 237

8.7 Appendix G: The results of corn samples ....................................... 249
G.1: Biocrystallogram and histogram of corn samples ......................... 249
G.2: F and P value plotting versus histogram matching ....................... 249

8.8 Appendix H: F and P value of effect of ROI on biocrystallization image .................................................. 250
H.1: Wheat DOC ........................................................................... 250
H.2: Wheat market ......................................................................... 258
H.3: Carrot from University of Kassel .............................................. 262
H.4: Carrot market ......................................................................... 268

8.9 Appendix I: F and P value of effect of color transformation on
biocrystallization image ................................................................ 272
I.1: Wheat DOC ........................................................................... 272
I.2: Wheat market ......................................................................... 280
I.3: Carrot from University of Kassel .............................................. 284
I.4: Carrot market ......................................................................... 290

8.10 Appendix J: The F and P value of effect of histogram matching on
biocrystallization image ................................................................ 294
J.1: Wheat DOC ........................................................................... 294
J.2: Wheat market ......................................................................... 302
J.3: Carrot from University of Kassel .........................................................306

J.4: Carrot market..........................................................................................312

8.11 Appendix K: The definition of terms in this research.................................316

8.12 Appendix L: The ANOVA results of the extended random grouping design via lmer formula.................................................................317
### Table of figures

**Figure 2.1**: the biocrystallogram images. (a) Copper chloride solution. (b) Wheat sample. (c) Carrot sample.

**Figure 2.2**: the phenomena of biocrystallogram image, divide into 3 stages of image on the basis of increasing concentration of additives at the giving amount of CuCl$_2$. On the left side is 1-zonal structure; 2-zonal structure (middle) and 3-zonal structure on the right side.

**Figure 2.3**: the detail of biocrystallization method.

**Figure 2.4**: the details of biocrystallization method; starts from the laboratory until the evaluation method.

**Figure 2.5**: the concept to control the temperature and humidity in the chamber.

**Figure 2.6**: the crystallogram of healthy and non healthy subjects.

**Figure 3.1**: critical stages of statistical input in scientific investigations.

**Figure 3.2**: nested factor.

**Figure 3.3**: nested and crossed factor.

**Figure 3.4**: brick wall.

**Figure 3.5**: (a) different patterns with the same first order statistical value. (b) The same histogram of different patterns with the same first order statistical value.

**Figure 3.6**: the different pattern (a and b) with different GLCM.

**Figure 3.7**: 15 Brodazt textures and their GLCM.

**Figure 3.8**: the GLDH calculation methodology according to GLCM.

**Figure 3.9**: the GLSH calculation methodology according to GLCM.

**Figure 3.10**: 6x6 image and its histogram.

**Figure 3.11**: histogram equalization technique.

**Figure 3.12**: example of histogram equalization.

**Figure 3.13**: an illustration of the steps of histogram matching.

**Figure 3.14**: anatomy of the human eye.

**Figure 3.15**: rods and cones cell.
Figure 3.16: RGB cube

Figure 3.17: RGB color cube including HIS co-ordinate system

Figure 3.18: RGB color cube including HIS co-ordinate system

Figure 3.19: HIS hexcone

Figure 3.20: the GLCM (with Gaussian distribution matching) of different ROIs of sample H

Figure 3.21: the GLCM (with Gaussian distribution matching) of different ROIs of sample G

Figure 3.22: the texture analysis results of the comparison sample G and H by plotting F and P value over the different ROIs

Figure 4.1: the 1st day experimental diagram of carrot from University of Kassel in 2004 as a prototype experimental design, the 2nd is repeated in the same method

Figure 4.2: the ANOVA result from R program version 2.1.0

Figure 4.3: the ANOVA result from SAS program version 9.1

Figure 4.4: (a) the region of interest (ROI) of carrot from University of Kassel data at circle-50 and circle-80. (b) The texture analysis by plotting F and P value versus ROI, vary from circle-30 to circle-100

Figure 4.5: the methodology to image analysis. (a) The scanned biocrystallogram image. (b) The gray level occurrence matrix (GLCM) at 0, 1 for the texture analyzed calculation. (c) The ANOVA result by plotting F and P value of 15 texture parameters over the image parameters

Figure 4.6: (a) the region of interest (ROI) of circle-50 (non blue area). (b) The region of interest of circle-80 and (c) circle-100

Figure 4.7: (a) F and P value of wheat DOC for overall assignments in 2003. (b) F and P value of wheat DOC for overall assignments in 2005

Figure 4.8: some of texture parameter over ROI. (a) Average diagonal moment value in 2003 and 2005. (b) Average energy value in 2003 and 2005: control-(A), mineral-(B), bio-dynamic-(C), organic-(D) and conventional sample(E) in 2003, control-(K), mineral-(H), bio-dynamic-(G), organic-(J) and Conventional sample(F) in 2005

Figure 4.9: the best ROI at circle-80 of wheat DOC (non blue area)

Figure 4.10: (a) F and P value of wheat market for overall assignments in 2004. (b) F and P value of wheat market for overall assignments in 2005
Figure 4.11: (a) F and P value of wheat market for the differentiation of general organic and conventional sample in 2004. (b) F and P value of wheat market for the differentiation of general organic and conventional sample in 2005.

Figure 4.12: (a) F and P value of wheat market for differentiation of varieties in 2004. (b) F and P value of wheat market for differentiation of varieties in 2005.

Figure 4.13: the best ROI at circle-70 of wheat market (non blue area).

Figure 4.14: (a) F and P value of different varieties of carrot from University of Kassel in 2004. (b) F and P value of different varieties of carrot from University of Kassel in 2005. (c) F and P value of different varieties of carrot from University of Kassel in 2006.

Figure 4.15: (a) the texture parameters value (diagonal moment and energy value) versus ROI changes of different varieties in 2004. (b) The texture parameters value (diagonal moment and energy value) versus ROI changes of different varieties in 2005. (c) The texture parameters value (energy and inertia value) versus ROI changes of different varieties in 2006.

Figure 4.16: (a) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2004. (b) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2005. (c) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2006.

Figure 4.17: the best ROI at circle-80 of carrot from University of Kassel samples (non blue area).

Figure 4.18: (a) F and P value of carrot market for overall assignments in 2004. (b) F and P value of carrot market for overall assignments in 2005. It is plotted on the Y axis versus the ROI on the X axis. Starting on the left with circle-40 and ending on circle-100.

Figure 4.19: (a) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2004. (b) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2005.

Figure 4.20: (a) F and P value of carrot market sample for differentiation of varieties in 2004. (b) F and P value of carrot market sample for differentiation of varieties in 2005.

Figure 4.21: the best ROI at circle-70 of carrot market (non blue area).

Figure 4.22: (a) RGB cube and (b) HIS hexcone model.

Figure 4.23: (a) F and P value of wheat DOC for the overall assignment in 2003. (b) F and P value of wheat DOC for the overall assignment in 2005.
Figure 4.24: (a) F and P value of wheat market for the overall assignment in 2004. (b) F and P value of wheat market for the overall assignment in 2005.

Figure 4.25: (a) F and P value of wheat market for differentiation of organic and conventional comparison in 2004. (b) F and P value of wheat market for differentiation of organic and conventional comparison in 2005.

Figure 4.26a: F and P value of wheat market for differentiation of varieties samples in 2004; Capo (ass 1), Bussard (ass 2) and Asketis (ass 3) in 2004.

Figure 4.26b: F and P value of wheat market for differentiation of varieties samples in 2005; Ludwig (ass 1), Bussard (ass 5) and Capo (ass 7).

Figure 4.27: (a) F and P value of different varieties of carrot from University of Kassel in 2004. (b) F and P value of different varieties of carrot from University of Kassel in 2005. (c) F and P value of different varieties of carrot from University of Kassel in 2006.

Figure 4.28a: F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2004.

Figure 4.28b: F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2005.

Figure 4.28c: F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2006.

Figure 4.29: (a) F and P value of carrot market for overall assignments in 2004. (b) F and P value of carrot market for overall assignments in 2005.

Figure 4.30: (a) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2004. (b) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2005.

Figure 4.31: (a) F and P value of carrot market sample for differentiation of varieties in 2004. (b) F and P value of carrot market sample for differentiation of varieties in 2005.

Figure 4.32: (a) histogram equalization. (b) Gaussian matching. (c) Triangle matching. (d) Parabolic matching and (e) hyperbolic matching.

Figure 4.33: (a) F and P value of wheat DOC for the overall assignment in 2003. (b) F and P value of wheat DOC for the overall assignment in 2005.

Figure 4.34: box plot of diagonal moment versus different histogram matching of Wheat DOC.

Figure 4.35: (a) biocrystallization image of wheat DOC. (b) The original histogram of wheat DOC. (c) The histogram of wheat DOC after match to Gaussian matching.
Figure 4.36: (a) F and P value of wheat market for the overall assignment in 2004. (b) F and P value of wheat market for the overall assignment in 2005.

Figure 4.37: (a) F and P value of wheat market for differentiation of organic and conventional comparison in 2004. (b) F and P value of wheat market for differentiation of organic and conventional comparison in 2005.

Figure 4.38a: F and P value of wheat market for differentiation of varieties samples in 2004; Capo (ass 1), Bussard (ass 2) and Asketis (ass 3).

Figure 4.38b: F and P value of wheat market for differentiation of varieties samples in 2005; Ludwig (ass 1), Bussard (ass 5) and Capo (ass 7).

Figure 4.39: (a) F and P value of different varieties of carrot from University of Kassel in 2004. (b) F and P value of different varieties of carrot from University of Kassel in 2005. (c) F and P value of different varieties of carrot from University of Kassel in 2006.

Figure 4.40: (a) the original histogram and biocrystallogram of Rodelika and (b) Rothild variety in 2004.

Figure 4.41: (a) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2004. (b) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2005. (c) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2006.

Figure 4.42a: the biocrystallogram image and original histogram of different nitrate levels in Rodelika in 2004.

Figure 4.42b: the biocrystallogram image and original histogram of different nitrate levels in Rothild in 2004.

Figure 4.43: (a) F and P value of carrot market for overall assignments in 2004. (b) F and P value of carrot market for overall assignments in 2005.

Figure 4.44: (a) the crystallogram and original histogram of paired assignment, e.g., assignment 9 (P1) and assignment 10 (P2) of carrot market in 2004. (b) Their F and P value result.

Figure 4.45: (a) the crystallogram and original histogram of paired assignment, e.g., assignment 1 (O1) and assignment 2 (O2) of carrot market in 2004. (b) Their F and P value result.

Figure 4.46: (a) the crystallogram and original histogram of paired assignment, e.g., assignment 5 (Q1) and assignment 6 (Q2) of carrot market in 2004. (b) Their F and P value result.

Figure 4.47: (a) F and P value of carrot market sample for differentiation of varieties in 2004. (b) F and P value of carrot market sample for differentiation of varieties in 2005.
Figure 4.48: the original histogram and some of crystallogram from different varieties; upper is Nerac variety and Narbonne in the lower. 142

Figure 4.49: (a) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2004. (b) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2005. 143

Figure 4.50: (a) F and P value of texture parameters versus ROI change of different varieties in 2004. (b) The biocrystallization image of different varieties. (c) Box plot of cluster prominence to differentiate the different varieties in 2004. 148

Figure 4.51: (a) F and P value of texture parameters versus ROI change of different varieties in 2005. (b) The biocrystallization image of different varieties. (c) Box plot of difference energy to differentiate the different varieties in 2005. 149

Figure 4.52: (a) F and P value of texture parameters versus ROI change of different varieties in 2006. (b) The biocrystallization image of different varieties. (c) Box plot of inertia to differentiate the different varieties in 2006. 150

Figure 4.53: (a) the biocrystallogram images, (b) F and P values of texture parameters versus ROI change, (c) box plot of cluster prominence and diagonal moment values of the different nitrate fertilizer levels of Rodelika variety. 152

Figure 4.54: (a) the biocrystallogram images, (b) F and P values of texture parameters versus ROI change, (c) box plot of cluster prominence and diagonal moment values of the different nitrate fertilizer levels of Rothild variety. 153

Figure 4.55a: the biocrystallogram images of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha). 154

Figure 4.55b: F and P values of texture parameters versus ROI change of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer level at 0 kg N/ha. 154

Figure 4.55c: box plot of cluster prominence of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha) in 2004. 155

Figure 4.55d: box plot of cluster prominence of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha) in 2005. 155

Figure 4.55e: box plot of inertia of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha) in 2006. 155

Figure 5.1a: region of interest (ROIs) of carrot from University of Kassel in 2004 are varying from circle-40(a1), circle-80(a2) and circle-100(a3). 177
Figure 5.1b: different GLCM which was calculated according to the ROIs changes in figure 5.1a, circle-40(b1), circle-80(b2) and circle-100(b3)………………..178

Figure 5.1c: F and P value plotting of texture parameters versus ROIs changes in figure 5.1a……………………………………………………….179

Figure 5.2: F and P value versus color transformation of wheat …………………181

Figure 5.3: F and P value versus color transformation of carrot …………………181

Figure 5.4: (a) F value of texture parameters versus histogram matching changes. (b) P value of texture parameters versus histogram matching changes. (c) Box plot of diagonal moment value versus histogram matching……….183

Figure 5.5: (a) the different biocrystallogram image and their GLCM according to the diagonal moment group. (b) F value of texture parameters versus ROI change…………………………………………………………………………… .187

Figure 5.6: (a) the different biocrystallogram image and their GLCM according to the energy group. (b) F value of texture parameters versus ROI change………188
Table of tables

Table2.1: the definition of terms in this research........................................8

Table3.1: table generation for histogram equalization of figure 3.12..............36

Table3.2: conversion from RGB color co-ordinate to HIS co-ordinate ............43

Table4.1: the overall basic details of a prototype data of carrot from University of Kassel in 2004 ............................................................50

Table4.2a: the overall basis data of carrot from University of Kassel in 2004 with complete experimental design........................................58

Table4.2b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with complete experimental design...............................59

Table4.2c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with complete experimental design (1, 2 are Rodelika variety)......................................................60

Table4.2d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with complete experimental design (3, 4 are Rothild variety)..................................................61

Table4.3a: the overall basis data of carrot from University of Kassel in 2004 with 2 days and 1 chamber experimental design................................62

Table4.3b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with 2 day and 1 chamber experimental design.......................63

Table4.3c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with 2 days and 1 chamber experimental design (1, 2 are Rodelika variety)..............................64

Table4.3d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with 2 days and 1 chamber experimental design (3, 4 are Rothild variety).................................65

Table4.4a: the overall basis data of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design..............................66

Table4.4b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design..............................67

Table4.4c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design (1, 2 are Rodelika variety)..............................68
Table 4.4d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design (3,4 are Rothild variety).

Table 4.5a: the overall basis data of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design.

Table 4.5b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design.

Table 4.5c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design (1,2 are Rodelika variety).

Table 4.5d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design (3,4 are Rothild variety).

Table 4.6: the significance for paired of wheat DOC assignments in 2003 and 2005.

Table 4.7: the significance for paired of wheat DOC assignments in 2003 and 2005.

Table 4.8: the summary result of histogram matching for wheat DOC in 2003 and 2005.

Table 4.9: the strongest effect of texture parameters to differentiate carrot from University of Kassel in 2004, 2005, and 2006.

Table 4.10: the strongest texture parameter to differentiate carrot market in 2004.

Table 4.11: the biocrystallogram image, difference in visual characteristic and box plot of texture parameters for carrot market in 2004.

Table 4.12: the strongest texture parameter to differentiate carrot market in 2005.

Table 4.13: the biocrystallogram image, difference in visual characteristic and box plot of texture parameters for carrot market in 2005.
Abbreviations

ACIA: Applied Crystallogram Image Analysis
ANOVA: Analysis of Variance
BMELV: Federal Ministry of Food, Agriculture and Consumer Protection
GLCM: Gray level co-occurrence matrices
GLDH: Gray level difference histogram
GLSH: Gray level sum histogram
HIS: Hue, Intensity and Saturation
HLS: Hue, Lightness and Saturation
HSV: Hue, Saturation and Value
lme: linear mixed effect model
RGB: Red, Green and Blue
ROI: Region of Interest
Chapter 1: Introduction

1.1 Motivation

Quality and safety of food are the important issues which receive increasing attention in the general public. The consumers are becoming more and more concerned about food quality, especially regarding how, when and where the foods are produced (Hoglund et al., 1999; Kahl et al., 2004; Alföldi et al., 2006). They perceive organic foods as being of better quality, healthier and with more nutritious benefits than foods that are produced using conventional methods (Lauridsen et al., 2005). It can be viewed that the growing interest and demand of organic foods are increasing worldwide and practiced in approximately 120 countries of the world. Currently, the countries with the biggest organic areas are Australia (12.1 million hectares), China (3.5 million hectares) Argentina (2.8 million hectares) and the European Union with more than 5.8 million hectares. The proportion of organically compared to conventionally managed land, however, is highest in Europe. Meanwhile, Latin America has the greatest total number of organic farms (Yussefi, 2006).

In order to sustain the organic market growth, one of the major orientations is quality control which will back the consumer’s perception of organic foods (Tauscher et al., 2003). Therefore, during recent years there has been a growing interest in the methods for food quality assessment, giving a more overall expression of crop and product quality. The biocrystallization method is one of these methods, applied not only for a discrimination between organic and conventional food, but also as complement to traditional methods of quality assessment, e.g., single compounds analysis (Meier-Ploeger and Vogtmann, 1991; Andersen and Busscher, 2006). The interpretation of images, which are generated from this method, should be accomplished and standardized as a validated tool for routine evaluation.

From the aspect of a product oriented definition of food, there are many differences between organic and conventional agriculture. At its simplest, organic food means food that is decreasing the external condition by limits placed on the use of chemical fertilizers, pesticides. Alternatively, crop rotation and the choice of optimal variety (Hoglund et al., 1999; Jones, 2003) are optimizing the health of soil and plants. Organic farming is a systemic farming. Therefore it needs a systemic approaches which reflects the benefit for soil, plant, animal and human. The biocrystallization method is one of that
approaches which can contribute to make visible possible benefits of the organic system.

1.2 The biocrystallization method

The biocrystallization method, also named “sensitive crystallization” or “copper chloride crystallization”, was originally introduced by E. Pfeiffer in the 1930ies. It has been developed from the viewpoint that living organisms do not just exist as substances, but have structuring and organizing properties. These properties control the form and function of an organism. The principle of biocrystallization is based on the crystallographic phenomenon that when crystallizing aqueous solutions of dihydrate CuCl$_2$ with additives of organic solutions, originating, e.g., from crop samples, biocrystallograms are generated with reproducible crystal patterns (Kleber & Steinike-Hartung, 1959). Its output is a crystal pattern on the glass plates from which variables (numbers) can be calculated by using image analysis. From this consequence, there are several scientists investigating and applying this method to differentiate organic and conventional food samples in agriculture or medical diagnosis (Cocude, 1998). Herewith, the effect of different stages of freshness and degradation on carrot quality was studied by Le Gia, 1995 and Le Gia et al., 1996. As well as the effect of liquid-N-freezing on carrot biocrystallogram, mineral N fertilization and light intensity on the pictomorphological properties of barley were worked by Andersen et al., 2001; Andersen and Busscher, 2001 and Andersen et al., 1999. In the section of medical diagnostic, it was studied as a diagnostic tool such as on the human health aspect via the human blood from healthy and unhealthy patients (Shibata et al., 1998; Piva et al., 1994 and Piva, 1998).

1.3 The biocrystallization evaluation

Concerning the evaluation of the biocrystallogram image, there are mainly 2 methods, i.e., visual evaluation and computerized image analysis. The visual assessment method is evaluating and interpreting the pictures visually by trained humans according to ISO standards (Kretschmer, 2003 and Zalecka, 2006). Meanwhile the computerized image analysis is a method that applied texture analysis to differentiate the images by using the knowledge from other applications, e.g., remote sensing, medical diagnostic and digital image processing etc.

1.3.1 Visual evaluation

The visual evaluation is an evaluation deduced from ISO – Norms (Kretschmer, 2003 and Zalecka, 2006) which is based on ranking/scoring of discrete scales of various morphological features in the macro- and microscopic crystal structure. One feature for visual evaluation is the coordination of the crystal structure, i.e., the degree of which the individual crystals are randomly or orderly distributed over the glass plate. And it is then including the degree to which the overall structure is characterized by sample-specific features (Andersen et al., 2001).

Although since 1930 several systems of visual evaluation with various ranking techniques, have been elaborated by several authors, they have shown to be capable of judging crystallograms by connecting them correctly, e.g., to different farming systems or different processing techniques (Engqvist, 1961; Mäder et al., 1993; Balzer-Graf,
1996; Schudel et al., 1980; Lieblein, 1993; Hagel et al., 2000; Weibel et al., 2000; Andersen, 2001). Herewith the literatures about these results still lacks a clearly define and communicable language. Anyhow, Huber et al. (2006) and Kretschmer (2003) succeeded to develop the visual evaluation method for discrimination the biocrystallogram image from organic and conventional farm and/food by establish a scientific method according to principles of sensory analysis.

1.3.2 Computerized image analysis

The digital imaging (scanning resolution) and computer calculation power increased during the recent decade. The computerized image analysis becomes a possible interesting tool. It is a method that analyzes the image by using the computer science. The goal of image analysis is the “construction of scene descriptions” on the basis of information extracted from image or image sequence (Rosenfeld, 1984). This method does not provide only the information which represents the visual characteristics, but also those which can not visually be differentiated (Basset et al., 2000).

Nowadays this method is also applied in many fields of research, for instance food technology, medical diagnostic tool, remote sensing, navigator and geosciences aspects etc. For example, Gupta and Undrill (1995) used the texture analysis to delineate suspicious masses in mammography as a diagnostic tool. Copper (2004) studied the texture analysis of gravity data by using co-occurrence matrix in the field of geosciences. Anderson and Busscher (2001) examined the effect of different levels of mineral N and light intensity on the pictomorphological properties of barley.

Concerning the application of this approach to the biocrystallization method, there are also several researchers who have evaluated and interpreted the food images via the computerized image analysis. Earlier studies have applied texture analysis, based on 256 gray levels, and a reduced area (region of interest; ROI). Both mono- and polycentric biocrystallogram were applied (Le Gia, 1995; Le Gia et al., 1996; Andersen et al., 1999; Andersen, 2001). However, up-to-now, there are still some queries about the parameters of the texture analysis like histogram matching and color to gray level transformation.

1.4 Objective of study

In the field of food quality, the biocrystallization method still has a major limitation. That is the lack of a standardized evaluation method in quantification and differentiation of the morphological features to discriminate samples from different treatments. In order to go a next step on the standardization for the biocrystallization evaluation as a routine analysis, therefore this thesis is based on samples from the project 020E170/F (German Research Project that is financed by the Federal Ministry of Food, Agriculture and Consumer Protection). Here using texture image analysis to differentiate wheat and carrot samples from different field trials and from the market.

1 The work was partially guided from Mrs. Machteld Huber at the Louis Bolk Institute in Netherlands
The sakes of the thesis are

1. To optimize the statistical model according to the experimental design, because the experimental design is dealing with repeated measurements. The former statistical model was established calculating the P value with a very conservative approach. Therefore, the new statistical model should be corrected and verified. Moreover, it should be accomplished with the existing ACIA (Applied Crystallogram Image Analysis) program which is used for the texture analysis.

2. To investigate the effect of the parameters of texture analysis on biocrystallization images based on the question whether these parameters show an effect on the differentiation of the samples, e.g., farming system of processing step. Moreover, the optimization is necessary to optimize the discrimination of these samples. The parameters for optimization are: region of interest (ROI), color transformation and histogram matching.

3. To get a better understanding of the texture analysis and check its limitation for the differentiation of wheat and carrot samples from different farming treatments. Furthermore, to consider the relation of texture parameters and visual evaluation criteria in order to clarify how the involvement of the textural and visual characteristics are for evaluation.

The goal will be a routine standard method of texture image analysis for the discrimination of wheat and carrot samples via the biocrystallization method.

1.5 Outline of the thesis

Chapter 1 gives an introduction to the motivation, biocrystallization method, the evaluation of this methods and objectives of study, followed by Chapter 2: the method description. Chapter 3 contains the theory part of a mixed effect statistical model with repeated measurement and computerized image analysis. Chapter 4 presents the study methodology and the results, it consists of (a) optimized the statistical model, (b) effect of Region of Interest (ROI), (c) effect of color transformation, (d) effect of histogram matching on biocrystallization image in order to distinguish wheat or carrot samples from different field trials and farm pairs, (e) the relation of texture parameters and visual evaluation. Chapter 5 presents the discussion with conclusions and future work, followed by the abstract in Chapter 6, references in Chapter 7. Chapter 8 contains the appendix.
Chapter 2: Method description

2.1 Biocrystallization method

2.1.1 The biocrystallization phenomena

The biocrystallization method, also called sensitive crystallization or copper chloride crystallization, was originally introduced by Pfeiffer (1930) but the biocrystallization term was later introduced by Engqvist (1970). The basic principal of the method is based on the crystallographic phenomenon that when an aqueous solution of CuCl₂.2H₂O (dihydrate copper chloride) is added with organic or inorganic substances, the crystallograms will be a reproducible dendritic structure (Kleber and Steinike-Hartung, 1959; Andersen and Laursen, 1998). By that, the biocrystallograms which are produced from agricultural products, such as vegetables, grains, fruits and milk samples are on the basis of three components: (a) an aqueous solution or extract of the sample in question; (b) an aqueous solution of dihydrate copper chloride; (c) purified water (Andersen, 2001).

(a) copper chloride solution                  (b) wheat sample                  (c) carrot sample

Figure 2.1: the biocrystallogram images. (a) Copper chloride solution. (b) Wheat sample. (c) Carrot sample.
The phenomena of biocrystallograms are taken place on the basis of ramification pattern that may be divided into three major stages, extending from the mostly one center in all directions to the periphery of the image (figure 2.2). In the initial or 1-zonal biocrystallogram, when increasing concentrations of agricultural extracts are applied relative to a given fixed concentration of CuCl$_2$, the transparent needles with enormous star-like formations are extending in all directions to the periphery. The second, 2-zonal structure, when divided by using a vertical and horizontal axis going through the crystallization center. The needles are pointed, predominantly transparent and relatively equal length in the middle zone. These morphological features may be described by means of plant morphological terms, such as stems, branches and needles. The last stage of the biocrystallogram is to an optimal degree divided into a 3-zonal structure with a center zone, middle zone and marginal zone. In the third stage the biocrystallogram exhibits various macro- and microscopic morphological features which reflect the quality of sample in question. (Andersen, 2001; Engqvist, 1970)

2.1.2 The biocrystallization method

The biocrystallization method comprises of 2 main parts. The first is the pattern formation which starts from the laboratory until the crystallogram picture is completely generated in the chamber. The second part is pattern recognition, herewith, the evaluation tools will be responsible to perceive and differentiate all of the images (figure 2.3).
There are two tools to evaluate the image, i.e., visual evaluation and computerized image analysis (figure 2.4). The visual evaluation determines the images in questions by trained human with the judgment of discrete reference scales arranged in connection with picture phenomena (Huber et al., 2006 and Kretschmer, 2003).
The other evaluation tool is computerized image analysis. It interprets the image by using the fundamental knowledge of texture analysis. Even though such a technique has been explored and applied with the biocrystallization method (Le Gia 1995; Le Gia et al., 1996; Andersen et al., 2001a and b). But, there are still some queries about the effect of texture analysis parameters on the evaluation of the crystallogram image which are not yet deeply investigated. They are histogram matching, color to gray level transformation and the ROI dependency. Hence, it leads to the main task of this research work.

Table 2.1: the definition of terms in this research.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region of Interest (ROI)</td>
<td>The region that is specified to examine which zone of an image contains relevant information for classification purpose (Andersen, 2001).</td>
</tr>
<tr>
<td>Histogram matching</td>
<td>A method to enhance the image contrast by matching an original histogram of gray scale to another histogram distribution, e.g., Gaussian, equalized histogram etc. (Hilger, 2004).</td>
</tr>
<tr>
<td>Color transformation</td>
<td>The color that transform RGB (red, green and blue color) three dimensional information to the desired color which can provide more enhancements (Lillesand et al., 2004).</td>
</tr>
</tbody>
</table>
2.2 The main factors of the biocrystallization method

- **The physical aspect in the crystallization apparatus**

The physical conditions in the crystallization apparatus (so called chamber) that influence the biocrystallogram are air temperature, air movement, air humidity, air pressure and mechanical vibrations. So, various types of crystallization apparatuses have been applied in order to regulate and control these major influencing factors (Andersen, 2001).

The concept is to control the temperature and humidity in the chamber so the conditions are reproducible (figure 2.5), i.e., coming from the outside, 4 areas are shielded from each other.

(a) The outer atmosphere including the local meteorological conditions which can not be controlled, only measured.

(b) The outer chamber where temperature and humidity are controlled.

(c) The inner chamber where the evaporation and crystallization takes place, where the temperature is controlled.

(d) The evaporation or crystallizing surface which is shielded by a ring (Ballivet et al., 1999; Andersen, 2001).

![Diagram](image)

Figure 2.5: the concept to control the temperature and humidity in the chamber.

From this concept, the original chamber was then developed by E. Pfeiffer and others, various techniques were applied including a vessel with a volume of 0.07 m$^3$ designed for 10 biocrystallograms (Nickel, 1968), a combination of 12 crystal growth units each designed for a single biocrystallogram (Shibata et al., 1998), and actual chambers that are entered when setting up an experiment with volume ranging from 3.6 m$^3$ (Teisseron and Neumann, 1998). Later, Andersen (1992) studied the experiment in connection with the control of the physical conditions in a cube-like chamber type. It was concluded that in
case of the absence of control of temperature and humidity in the outer chamber, the atmosphere variations were reflected strongly in the evaporation or crystallization time in the inner chamber. Therefore a chamber was developed in 1998 in order to control air temperature and humidity gradients between inner and outer chamber (Andersen, 1998) and the latest chamber was succeeded in between 2002 and 2003 by triangle researchers group (University of Kassel, Germany; Louis Bolk Institute (LBI), Netherlands and Biodynamic Research Association Denmark (BRAD), Denmark).

Ballivet et al. (1999) argued that relative humidity above the ring influences the evaporation time of the plates. When decreasing the relative humidity, the evaporation time is reduced. The longer evaporation time is always connected with shorter crystallization duration.

In addition Busscher et al. (2003) noticed that the crystallization time affects the crystal picture, therefore the effect of evaporation time on morphological characteristics was studied with wheat, carrot and PVP pictures. The result showed that short evaporation times are relatively characterized by more or less numerous and pronounced “heat ring”, ramification structure with a reduced diameter, rather linear stems as well as a general thinning out of ramification stems. The optimal evaporation time is 12-14 hours (Kahl et al., 2003).

- The concentration matrix

In the case of the so called solution concentration matrix, the optimal mixing ratio between sample extract and copper chloride is another effect influencing the crystallization image. In order to provide the optimum of desired morphological features which enables the differentiation of various samples, the optimum of sample and CuCl₂ concentrations were investigated. The concentration vector is determined by increasing the amount of sample (juice or extract) that is added to a fix amount of reagent (200mg CuCl₂ per plate), see Selawry and Selawry (1957); Selawry (1961) and Engqvist (1970). After that, this ratio concentration matrix study has been developed for wheat and carrot samples, the optimal mixing ratio for carrot is 115/90 meanwhile 90/90 for wheat sample at the chamber condition of 30 °C, 50 % relative start humidity. Where the first number represents the amount of extracted sample (mg per plate) and the second is the amount of reagent (mg per plate) (Andersen et al., 2003; Kahl et al., 2003).

2.3 Application of biocrystallization method

In recent years the biocrystallization method has been applied in many researches as mentioned before. They can be divided into 2 main categories, i.e., medical and agro-industrial approaches (Cocude, 1998).

- The medical approach

Concerning of the medical, it was applied as a medical diagnostic tool, especially, on the human health aspect via the human blood from healthy and unhealthy patients. It can be supported with Shibata et al. (1998). They noticed that the blood solution with copper chloride has an enormous effect on the growth of hydrated cupric chloride crystals and the difference in crystal shapes form were found when exposed to diabetic patients and healthy subjects blood.
As well as Piva et al. (1994) and Piva (1998) compared patients, from the Service of Internal Medicine at the Montpellier hospital, with healthy subjects by using the crystallization method. The result showed that for the healthy subjects, the crystallization was more regular than for non healthy subjects. In pathological groups there were crystallization alterations leading to different geometrical figures.

(a) the crystallogram of healthy subjects. (b) the crystallogram of non healthy subjects.

Figure 2.6: the crystallogram of healthy and non healthy subjects (Piva et al., 1994; Pp 30).

- **The agro-industrial approach**

   Regard to agro-industrial, the aim is to determine the product’s quality or its biological compatibility with the human organism. Experiments now being carried out in various areas (Cocude, 1998). The effect of different stages of freshness and degradation on carrot quality was studied on consecutive several days such as 1st, 15th and 30th day (Le Gia, 1995; Le Gia et al., 1996). As well as the carrot samples were exposed to autolysis degradation over consecutive days at 6°C (Andersen et al., 1999). Their finding was that the storage or degradation treatment induced systematic changes in morphological features over a number of successive days.

   Concerning to the plant fertilization research, studies have been reported a relation between a low level of N-availability and picture developing properties. Schudel et al. (1980) found that the conventional spinach grown with 100 or 300 kg N ha-1 exhibited less coordinated structures than organic samples grown with N- equivalent manure and plant compost. These low N organic samples contained three to six times less nitrate-N than the high-N conventional samples. Additionally, pictomorphological properties were also affected by level of mineral and light intensity. Barley samples originated from 4 levels of mineral - N fertilizers (0, 45, 90 and 135 kg N ha-1, respectively) and three levels of light intensity (0, 30 and 55% reduction relative to daylight) provided several of morphological features. These various pattern can be divided into three groups, via visual evaluation. The first was the pattern from light intensity 100%. The second was light intensity 70, 100 with N-level 135, and light intensity 70 with N-level 135. The last group
was light intensity only 50 and light intensity 50 with N-level 135 (Andersen and Busscher, 2001).

Regard to the food processing field, the connection between food processing and quality was also examined. One of them is the effect of freeze processing. The experiment was done not only the effect of deep freeze storage of carrot extract at -20 °C over one year, but also the liquid-N-freezing treatments on pictomorphological properties of carrot. The result noticed that increasing liquid N- treatment and long storage with deep frozen were monotonically decreasing the visual evaluation scores, that mean there is an effect to morphological phenomena of the pictures and contributed to the quality of carrot (Andersen et al., 2001a).

Besides, the effect of microwave heating was also determined on the picture developing properties of food via carrot juice, the result showed that the heating treatments (done by a microwave) have a marked derivative effect on the picture properties (Andersen, 2005).

2.4 Evaluation of biocrystallization method

In connection with assessing the quality of products from evaluated biocrystallograms, it should be ideally based on the following:
(a) precise description of the individual biocrystallogram;
(b) a comparison between different groups of biocrystallograms originating from different treatments, experiments and years;
(c) a correlation with other experimental data, including quantitative data from soil analysis and chemical analysis of nutritional constituents (Andersen, 2001, Andersen and Busscher, 2006).

The 2 main evaluations of biocrystallograms are visual evaluation and computerized image analysis.

2.4.1 The visual evaluation

The visual evaluation method demands a trained panel and the accumulation of a great number of data which are based on various morphological features (Barth, 1998). The different categories of biocrystallogram judgment based on a scale of 1 - 9 have been elaborated by the triangle researcher group (appendix E). They succeeded to use it for discrimination the sample quality via the biocrystallogram image. The main categories are:

(1) quantifiable separate morphological features, also called local features, like curly needles;

(2) interconnectedness of separate morphological features also called textural features, like dense radial formation, regularity of ramifications;

(3) quality of the image as a whole, also called integration, coordination and “beweglichkeit”;

(4) interpretation of the above features by connecting them to concepts from plant – physiology like ripeness.
2.4.2 The computerized image analysis

The biocrystallogram structures are ramified structures which are homogeneous within confined zones rather than throughout the crystallized object (Engelmann and Ersboell, 2004). The evaluation method may demand an image processing approach to analyze such a kind of crystalline growth whose complexity delivers a large amount of texture information (Teisseron and Neumann, 1998). So the image analysis procedure ideally should reflect all the characteristics of a biocrystallogram as a three dimensional, colored crystal structure, coordinated with zones relative to a center (Andersen et al., 1999).

The computerized image analysis has been explored by several scientists, Le Gia (1995) and Le Gia et al. (1996) followed two major groups of biocrystallogram images, i.e., (a) image originating from leek, carrot, milk and egg yolk samples. (b) Images originating from comparable carrot samples with different stages of freshness, termed 1st, 15th and 30th day. The application so called texture analysis was based on image acquisition by means of 256 gray-levels, 256×256 pixels with 70% radial extension around geometric center of the biocrystallogram. It was noticed that high classification rate of this evaluation method were found when combining first and second order parameters.

Kahl et al. (2003); Kahl (2006) and Kahl et al. (2006) have been worked on the computerized image analysis to differentiate biocrystallogram picture from controlled field trial sample (DOC trial) as well as from farm pairs of organic and conventional farming (apple, carrot and wheat samples). The results showed that the computerized image analysis approach can differentiate samples as statistically significant. Moreover, the correlation of the results with parameters measured from the other methods on the same material shows that the biocrystallization does not depend on other analysis method parameters such as dry matter or single compounds but gives a holistic pattern of the product.

The texture approach was explored further with particular focus on the performance of the five resolution scales and the 23 first and second parameters to discriminate 5 groups of polymerization levels of PVP(polyvinyl pyrrolidone), the best performing parameters showed a monotonic relationship to the polymerization levels (Andersen et al., 2001b).
Chapter 3: Theory

3.1 Repeated measurement statistical model; mixed effect model

3.1.1 Introduction

Statistics is a science to solve the problem of variability. It starts with a problem, continues with the collection of data, proceed with the data analysis and finishes with the conclusion (Faraway, 2002)

Figure 3.1 symbolizes the fact that statistics should play a role in every data collection and analysis, from initial problem formulation to the drawing of the final conclusion. From this figure, an experiment studies the variables of interest which often can be controlled and fixed at predetermined values for each test run in the experiment. The unit which treatments are applied or assigned is called experimental unit, e.g., in a clinical trial where different patients are given different drugs, each patient is an experimental unit. If, on the
other hand, each patient is given a different oilment on each arm, then each arm constitutes an experimental unit.

In observational studies, many of the variables of interest cannot be controlled but they can be recorded and analyzed (Mason et al., 2003; Hinkelmann and Kempthorne, 1994). Therefore, statistical experimental design will be shown to be effective in eliminating known source of bias, guarding against unknown source of bias, ensuring that the experiment provides precise information about responses of interest, and guarantee that excessive experimental resources are not needlessly wasted through the use of an uneconomical design (Mason et al., 2003).

In many experiments in agriculture influences are usually analyzed based on linear statistical models that incorporate both fixed and random effects. The fixed effect is the effect which parameters are associated with an entire population or with certain repeatable levels of experimental factors. While, random effect is the effect which is associated with individual experimental units drawn at random from a population. Random effects arise when there is more than one observation on one experimental unit since we expect the unit to vary independently. A statistical model with both fixed and random effects is called mixed effect model (Mahony, 1986; William and Ripley, 2002; Pinheiro and Bates, 2000). This model extends linear models by incorporating random effects, which can be regarded as an additional error term, to account for the correlation among observations within the same group (Pinheiro and Bates, 2000). From this consequence, the overall significance of treatment (P-value) will be more increase when the main observation unit is random from the population instead of specified it as fixed. A mixed model is an important advance in the statistical methodology and frequently used to analyze group data of agriculture because its flexibility and includes not only several effects representing different random source of variation, but also the development of widely available software, e.g. the Mixed procedure of SAS (Davis, 2002; Stroup, 1989). A few examples for such model are as follows: (1) a split-plot experiment requires two error terms for main plot and sub-plot. (2) Repeated measurements taken at different point in time at the same unit. (3) Experiments replicated at several sites and/or in several years (Piepho et al., 2003).

### 3.1.2 The principal concept of mixed effect model

Concerning a principle of linear statistical model, it is to express the observations, generally denoted by $Y$, in term of effect which contributes to $Y$. These effects or components fall basically into three categories: (a) treatment effects, (b) design effects and (c) error effects.

The treatment effects are a reflection of the intervention procedure or treatment design as either single treatments or factor combinations are applied. The design effects are determined by the error control design, in particular effects due to the various kinds of blocking. Finally, the error effects represent different kinds of random variation (Hinkelmann and Kempthorne, 1994)

$$Y = \text{treatment effects} + \text{design effects} + \text{error effects} \quad (3.1)$$

In the following, we will show the advantage of mixed model starting from the fixed way of handling a problem.
Pinheiro and Bates (2002) gave a simple example called rail way study. Six rails were chosen at random and tested three times each by measuring the time that took for a certain type of ultrasonic wave to travel the length of the rail. From this study, there is only a fixed effect and can be written as follows

\[ Y_{ij} = \beta_i + \varepsilon_{ij} \quad i = 1, \ldots, m \]
\[ j = 1, \ldots, n \]  

(3.2)

Where \( Y_{ij} \) represents observed value travel time for observation \( j \) on rail \( i \).
\( \beta_i \) represents mean travel time across the population of rail being sampled.
\( \varepsilon_{ij} \) represents error term which are independent distributed as
\[ N(0, \delta^2) \]

Even though the fixed effect model accounts for the rail effect. But it does not provide a useful representation of the rail data as it only model that specific sample of rail used in the experiment, while the main interest is on the population of rail from which the sample was drawn. Additionally, such fixed effect models does not provide an estimate of the between rail variability. Therefore, to circumvent these problems, the rail effect can be treated as random effect. The new model for this experiment, so called mixed effect model, can be written to

\[ Y_{ij} = \bar{\beta} + (\beta_i + \bar{\beta}) + (\varepsilon_{ij} + \beta_i) \]

(3.3)

that is,
\[ Y_{ij} = \bar{\beta} + b_i + \varepsilon_{ij} \quad i = 1, \ldots, m \]
\[ j = 1, \ldots, n_i \]

(3.4)

Where \( Y_{ij} \) represents observed value for observation \( j \) on \( i \).
\( \bar{\beta} \) represents mean travel time across the population of rail.
\( b_i \) represents the deviation of group mean from grand mean, that is the deviation from the population mean of the mean travel time for the \( i^{th} \) rail.
\( \varepsilon_{ij} \) represents the deviation of observation from group mean. It is a deviation in travel time for observation \( j \) on rail \( i \) from the mean travel time for rail \( i \).

As stated before, from this experiment, then the fundamental mixed effect model that contain both fixed and random effects, can be also referred and described by David (2002), i.e., assume that a continuous, normal distributed response variable is measured at each of \( t \) time points for each of \( n \) experimental units (subjects). So the model is
\[ Y_{ij} = \mu_{ij} + \pi_{ij} + \varepsilon_{ij} \quad i = 1, \ldots, n \]
\[ j = 1, \ldots, t \]  

Where \( Y_{ij} \) represents the response from subject \( i \) at time \( j \)

\( \mu_{ij} \) represents mean at time \( j \) for individual randomly selected from the same population as individual \( i \).

\( \pi_{ij} \) represents the consistent departure of \( Y_{ij} \) from \( \mu_{ij} \) for the \( j \)th subject.

\( \varepsilon_{ij} \) represents the departure of the \( \mu_{ij} + \pi_{ij} \) for individual \( i \) at time \( j \).

It is a mixed effect model since the parameter \( \mu_{ij} \) is a fixed effect as \( \mu_{ij} \) has a fixed value irrespective of the particular individual. While \( \pi_{ij} \) varies randomly over the population of individual, the \( \pi_{ij} \) parameter is called random effect. The \( \varepsilon_{ij} \) parameter is random error term.

Pinheiro and Bates (2002) gave another example in mixed effect model. This experiment is designed in randomized block design, i.e., the experiment records the effort required by each of nine different subjects to arise from each of four types of stools. The comparison is these four particular types of stool, therefore, the type factor is fixed effect while the nine different subjects are random effect as it represents a sample from the population about which the experimenter wish to make an inferences. The mixed model should be written as

\[ Y_{ij} = \beta_i + b_i + \varepsilon_{ij} \quad i = 1, \ldots, 9 \]
\[ j = 1, \ldots, 4 \]

Where \( Y_{ij} \) is the observed value for effort with observation type \( j \) on subject \( i \).

\( \beta_i \) is mean effort across the population of type.

\( b_i \) is deviation in effort mean for subject \( i \).

\( \varepsilon_{ij} \) is error term which show deviation in effort for observation type \( j \) on subject \( i \) from the mean effort from subject \( i \).

More examples and details can be found via above author, Davis, 2002; William and Ripley, 2002; Piepho et al., 2003; Piepho et al., 2004.

### 3.1.3 The nested and crossed effects in the experiment.

Designed experiments conducted by crop scientists often give rise to several random source of variation. Relevant examples are split plot design, series of experiments and repeated measurements taken on the same field plot etc. From these experiments, the nested and crossed factors are particular related to design in order to find out the best solution of problems.

The nested factors are the factors that have unique levels within each level of one or more other factors. That is they have levels that differ within one or more of the other
factors in the experiment or only one of treatments is given to each subject. The nested effects often are included in experiments when it is desired to study components of response variation that can be contributed to the nested factor (Mason et al., 2003). An example of an experiment where factors are nested can be seen in figure 3.2.

The experiment that the test method requires the polyethylene pellets be heated, melted and pressed into a plaque. A small disk is then cut from the plaque and the density of disk is measured. As the plaques made from one shift have no physical relationship with those made on another shift, so plaques are nested within shift. Similarly, because the disks cut from one plaque, herewith, disks are nested within plaques (figure 3.2).

Crossed factor is the factor that contains levels by having a physical or fundamental property that is the same for all levels of the other factors included in the experiment (Mason et al., 2003). It is also often used in an experiment when the levels of the factor are the specific value of the factors such as the complete factorial experiments as each level of each factor occurs in the experiment with each level of every other factor.

These kinds of both factors can be included in an experimental design (figure 3.3). From the figure, the experiment conducts to investigate automatic cutoff times of lawnmower. The crossing of the manufacturer (A, B) and speed (high, low), the nesting of lawnmowers within manufacturers.
More experiments for such both kind of factors are given by Sanders (1989), Cornelius and Archbol (1989); Stroup (1989); Pinheiro and Bates (2000); Piepho et al. (2003); Piepho et al. (2004); William and Ripley (2002) and Davis (2002). Such as a split plot experiment, it is a mixed effect design which the main plot factor A (fertilizer) is laid out in randomized complete block, while the sub plot factor B (genotype) is completely randomized within main plot. Factor A is the main plot factor, i.e., levels of A are randomly allocated to main plot within block. Levels of B are randomly allocated to sub plot within main plot. So there are two nested randomization units: main plot within block and sub plot within main plot but the treatment is crossed factor between fertilizer (A) and genotype (B). A full model is written as

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \gamma_k + b_{ik} + \varepsilon_{ijk}$$  \hspace{1cm} (3.7)

Where $Y_{ijk}$ is the yield of $i^{th}$ fertilizer and $j^{th}$ genotype in $k^{th}$ block.

- $\mu$ is general mean.
- $\alpha_i$ is main effect of $i^{th}$ fertilizer.
- $\beta_j$ is main effect of $j^{th}$ genotype.
- $\alpha\beta_{ij}$ is the fertilizer by genotype interaction.
- $\gamma_k$ is the effect of the $k^{th}$ block.
- $b_{ik}$ is the error of $i^{th}$ main plot within $k^{th}$ block.
- $\varepsilon_{ijk}$ is the sub plot error(residual).
3.1.4 Setting up the mixed effect model

The aspects to set up the mixed effect model which is appropriated in an experiment can be briefly explained as 2 main parts: block and treatment part. The block part can be expressed exclusively in term of block factor and represent the structure innate to the observational units. This block factors are comprised of randomly selected sampling unit (e.g. plants, soil sample etc) and blocking units, hence, block factors are needed to uniquely identify each observational unit. Another one part is treatment factor and its levels that are chosen by the experimentater to answer scientific research questions.

The main proceeding is to keep these 2 parts together and derive them into statistical model. The derivative method can be followed from Piepho et al., 2003 and 2004. Their arrangement base on 4 main operators, i.e.,

1. Dot operator (.) the dot operator is used to define the crossed effects part. That is

\[ A \cdot B = B \cdot A \]  \hspace{1cm} (3.8a)
\[ (A \cdot B) \cdot C = A \cdot (B \cdot C) \]  \hspace{1cm} (3.8b)
\[ (A \cdot C) \cdot (B \cdot C) = A \cdot B \cdot C \]  \hspace{1cm} (3.8c)

2. Product term operator \([pt(.)]\), If M is the model part, then \(pt(M)\) is the product term (by using dots) of all effects in M, i.e.,

\[ pt(A+B) = A \cdot B \]  \hspace{1cm} (3.9a)
\[ pt(A+B+C) = A \cdot B \cdot C \]  \hspace{1cm} (3.9b)

3. Nesting operator (/), if a factor B is nested within another factor A, then the model contain the term A and A.B, i.e.,

\[ A/B = A + A.B \]  \hspace{1cm} (3.10a)
\[ A/ (A/C) = (A/B) / C \]  \hspace{1cm} (3.10b)
\[ A/ (B+C) = A/B + A/C \]  \hspace{1cm} (3.10c)

4. Crossing operator (x), the model for two crossed factor A and B can be presented by \(A \times B = A+B+A.B\), i.e.,

\[ A \times B = A+B+A.B \]  \hspace{1cm} (3.11a)
\[ A \times (B+C) = A+B+C+A.B+A.C \]  \hspace{1cm} (3.11b)
\[ (A \times B)/C = A+B+A.B.C \]  \hspace{1cm} (3.11c)
The example to illustrate how these operators derived and to set up the mixed effect model can be followed by the split plot trial in Piepho et al. (2003). The split plot experiment contains the main-plot factor A (fertilizer) that laid out within block. And the sub-plot factor B (genotype) within main plot. So the block part of it can be noted as Block / (main plot/ sub-plot) and treatment part is A x B. The derived method of them will be followed as

1. Block part is Block/ (main-plot/sub-plot)

   = Block / (main-plot + main-plot. sub-plot) using equation 3.10a.
   = Block + Block . main-plot + Block . main-plot. sub-plot using equation 3.10a.

   Setting the main-plot with A and sub-plot with B, then
   = Block + Block . A + Block . A . B

2. Treatment part is Ax B = A + B + A.B where A.B is the interaction of A factor and B factor.

From 2 derivative parts, so the full mixed effect model of this experiment is

   A + B + A.B + Block + Block . A + Block . A . B

And then, it can be expressed to the completely statistical model as

\[ Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \gamma_k + b_{ik} + \varepsilon_{ijk} \] (3.12)

Where
- \( Y_{ijk} \) is the yield of \( i^{th} \) fertilizer and \( j^{th} \) genotype in \( k^{th} \) block.
- \( \mu \) is general mean.
- \( \alpha_i \) is main effect of \( i^{th} \) fertilizer.
- \( \beta_j \) is main effect of \( j^{th} \) genotype.
- \( \alpha\beta_{ij} \) is the fertilizer by genotype interaction.
- \( \gamma_k \) is the effect of the \( k^{th} \) block.
- \( b_{ik} \) is the error of \( i^{th} \) main plot within \( k^{th} \) block.
- \( \varepsilon_{ijk} \) is the sub plot error(residual).
3.2 Computerized image analysis

3.2.1 Introduction

As digital image and computer power increasingly develop during the recent years, image processing analysis and computer vision techniques have also become widely involved in many scientific field such as in the medical science in order to categorize tumor cells, remote sensing to interpreted the images from satellite and food technology to develop differential characteristic in food quality etc (Patton et al., 2006; Srivasteva et al., 2003). The image analysis or processing is one of the computerized image analysis techniques that have the main goal “the construction of scene descriptions on the basis of information” extracted from images (Rosenfeld, 1984; Pp 634). It computes statistics and measurements based on the gray level intensities of the image pixels. Thereby, it can determine whether the image quality is good enough for the inspection task (Rosenfeld, 1984; National instruments (http://zone.ni.com/devzone/cda/tut/p/id/3470)).

As mentioned before in chapter 2, the biocrystallograms are the images that are generated by different additives and based on the ramified structures which are homogeneous in structure. To sort out the quality and quantity of such image, it would benefit by induced the issue of texture to computerized image analysis (Engelsmann and Ersboell, 2004)

3.2.2 Texture analysis

Texture analysis is important in many applications of computer image analysis for the classification and analysis of many types of pictures. It can be found and applied in all pictures from multispectral scanner images obtained from aircraft or satellite platforms (in the field of remote sensing community analysis) to microscopic images of cell cultures or tissue samples (in the field biochemical community analyzes) (Haralick, 1979; Ojala et al., 2001; Wu et al., 2003; Coburn and Robert, 2004; Carstensen, 2002).

The texture of an image reflects changes of intensity values of pixels which might be contain information of geometric structure of objects (Zheng et al., 2006). So this technique does contain not only the information which is representative for visual characteristics, but also characteristics which can not be visually differentiated (Basset et al., 2000). Thus, the main objective of texture analysis is to extract useful textural information from an image and measures texture directly from the image based on the local spatial variation of intensity or color (Ojala et al., 2001; Wu et al., 2003; Coburn and Robert, 2004).

Several definitions of texture have been proposed over many literatures. Merriam - Webster dictionary (http://www.m-w.com/dictionary/texture ) provides five definitions of texture, of which four are applicable

1. a: “something composed of closely interwoven elements; specifically : a woven cloth”
b: “the structure formed by the threads of a fabric”

2. a: “essential part : substance”
b: “identifying quality: character”
3. a: “the disposition or manner of union of the particles of a body or substance”
b: “the visual or tactile surface characteristics and appearance of something, the texture of an oil painting”

4. a: “basic scheme or structure”
b: “overall structure”

Carstensen (2002) define the term texture is a region in 2D or 3D that can be perceived as being spatially homogeneous in some sense. Textural feature is the tonal or gray level variation of an image, it contains information about the spatial distribution of tonal variation or express of the local spatial structure in an image (Gupta and Undrill, 1995; Ferro, 1998; Tuceryan and Jain, 1998; Rikxoort, 2004). And this is the main definition of this research.

3.2.3 Texture statistics

The evaluation and development of new approaches for calculating texture have been a particular focus. Historically there have been two major approaches, i.e., structural and statistical approaches. The structural approach describes a texture by a sub pattern or primitive and spatial distribution of primitives, so called the placement rule. The primitives are also called texture elements. For instance, to consider the brick wall the primitive is a brick and the placement rule specifies the arrangement of bricks in the wall.

The statistical approach does not presume in term of primitive but it draws on the general set of statistical tool (Ferro, 1998; Carstensen, 2002). It is the most widely used and more generally applied method because of its high accuracy and less computation time (Zheng et al., 2006).

![Figure3.4: brick wall (Carstensen, 2002; Pp218)](image)

Texture statistics is frequently classified into first-order, second-order and high-order statistics. They are referring to the gray level distribution of pixel on an image. The gray scale is a black and white image at any given focus of pixel, typically there is a corresponding intensity on a range from 0 (black) to 255 (white). That means an image is composed of an array of pixels of varying intensity across the image, the intensity corresponding to the level of grayness from black (0) to white (255) at any particular point in the image (Carstensen, 2002; Patton et al., 2006).
3.2.3.1 First-order statistics

First-order statistics refers to the marginal gray level distribution. It can be computed from histogram of pixel intensity in the image, so it depends only on individual pixel values and it independents on their interaction of neighboring pixel values (Tuceryan and Jain, 1998; Basset et al., 2000; Aguila, 2004). First order statistics is used to characterize the statistical properties such as mean or average intensity, variance, skewness, kurtosis, energy and entropy.

The first-order gray level statistics can be derived from the gray level histogram \((h_i)\). \(h_i\) is the number of pixels in an image with gray levels \(i\), \(N\) is the total number of pixels and \(G\) is the number of gray levels, then

\[
\sum_{i=0}^{G-1} h_i = N. \text{ The normalized histogram } (H_i) \text{ with } H_i = h_i / N \tag{3.13}
\]

is the empirical probability density function for single pixel. Statistics computed from \(H_i\) are mentioned by several authors such as Carstensen (2002), Adan et al. (2003) and Harris (1987) as followed as

1. The mean gray level

\[
\mu = \sum_{i=0}^{G-1} iH_i \tag{3.14}
\]

\(\mu\) measures the averages intensity in the image

2. The gray level variance

\[
\sigma^2 = \sum_{i=0}^{G-1} (i - \mu)^2 H_i \tag{3.15}
\]

where \(\sigma\) is the standard deviation. The variance and the standard deviation measure the global contrast in the image.

3. The coefficient of variation

\[
cv = \frac{\sigma}{\mu} \tag{3.16}
\]
4. The gray level skewness

\[ \gamma = \frac{1}{\sigma} \sum_{i=0}^{G-1} (i - \mu)^3 H_i \]  \hspace{1cm} (3.17)

Skewness measures the extent to which outliers favor one side of the distribution or the departure from symmetry about the mean gray level.

5. The gray level kurtosis

\[ \gamma_2 = \frac{1}{\sigma^4} \sum_{i=0}^{G-1} (i - \mu)^4 H_i - 3 \]  \hspace{1cm} (3.18)

Kurtosis measures the peakness or tail prominence of the distribution or measure of the spread of gray tones about the mean.

6. The gray level energy

\[ e = \sum_{i=0}^{G-1} H_i^2 \]  \hspace{1cm} (3.19)

where \( G^{-1} < e \leq 1 \). Energy measures the non uniformity of the histogram or how much intensity variation is.

7. The gray level entropy

\[ s = -\sum_{i=0}^{G-1} H_i \log H_i \]  \hspace{1cm} (3.20)

where \( 0 < s \leq \log G \). Entropy measures the uniformity.
Nevertheless, sometime, first-order statistic can not distinguish some patterns, for instance, the figure 3.5 shows how the first orders measurement of texture can not differentiate the pattern.

![Figure 3.5](attachment:image.png)

Figure 3.5: (a) different patterns with the same first order statistical value (Ferro, 1998; Pp3). (b) The same histogram of different patterns with the same first order statistical value.

The figure 3.5a contains twenty-five numbers with the value 5, 10, and 20 in different patterns of texture, but the same histogram distribution (figure 3.5b). These textures show how some measures of texture can not distinguish pattern by using first order statistics (Ferro, 1998). So the alternative to circumvent and measure this kind of texture is gray-level co-occurrence matrix (GLCM) as it measures the relationships of pixel intensity to its neighboring pixels.

### Gray-level co-occurrence matrices

The gray-level co-occurrence matrix (GLCM), a frequency matrix, is a useful method for enhancing details and is frequently used as an aid for interpretation of an image. The GLCM indicated the frequency of a pair of pixels that at “exactly the same distance and direction of the displacement vector” (Park et al., 2004; Cooper, 2004). From this principal, it uses to computes the relationships of pixel intensity to the intensity of its neighboring pixels which are based on hypothesis that the same gray level configuration
is repeated in a texture and pixels that are close together tend to be more related than pixels that are far away from each other (Coburn and Robert, 2004; Mena and Malpica, 2003; Fernandez et al., 2005).

GLCM was introduced in Haralick (1979) and some authors, Carstensen (2002); Cooper, 2004; Basset, 2000; Barber et al., 1993 and Lefebvre et al., 2000. They have suggested that the GLCM can describe the probability of finding pixels of gray level value \(i\) and \(j\) at a given displacement \(h\). The GLCM, \(c\), is defined with respect to given (row, column) displacement \(h\). And element \((i, j)\), denoted \(c_{ij}\), is the number of times a point having gray level \(j\) occurs in position \(h\) relative to a point having gray level \(i\). Let \(N_h\) be the total number of pairs, the \(C_{ij} = c_{ij}/N_h\) is the elements of the normalized GLCM, \(C\).

The meaning of the above mentioned can be followed to refer the above case, i.e., figure 3.5.

If \(h = (0, 1)\), i.e., one step in the horizontal direction, then \(c\) (GLCM) will be

The position of each element in the matrix indicates which pixel values are being compared (figure 3.6). The value at row \(i\) and column \(j\) gives the number of times that a pixel with the value \(j\) was to the immediate right of a pixel with the value \(i\). Hence, the value 3 in the 2\(^{nd}\) column, 3\(^{rd}\) row of figure 3.6a indicates that a pixel value of 10 was to the right of a pixel with the value 20 counted 3 times. Meanwhile, the value 2 in the same position of GLCM from figure 3.6b means a pixel value of 10 was to the right of a pixel with the value 20 counted 2 times. Certainly, both the sum of these matrices \(N_h\) will be equal to 20. It can be seen that different pattern distributions of an image make dissimilar GLCM because it is derived from the intensity of pixels and its neighboring pixels. More different distributions of 256 gray values image can be seen via 15 Brodazt textures and their GLCM (figure 3.7).
Carstensen (2002) noticed that one of the main problems associated with the using of co-occurrence matrices is that they have to be computed for many gray values, thus providing with an immense amount of data so let

\[ C_i^x = \sum_{j=0}^{G-1} C_{ij} \]  \hspace{1cm} (3.21)  

\[ C_j^y = \sum_{i=0}^{G-1} C_{ij} \]  \hspace{1cm} (3.22)

where \( G \) is the number of gray levels and let \( \mu_x, \mu_y, \sigma_x \) and \( \sigma_y \) be the means and standard deviation of \( C_i^x \) and \( C_j^y \) over \( i \) and \( j \). Then, the second order statistical features—one of the advantage statistical approach to interpret images—can be computed from GLCM as follows.
3.2.3.2 Second-order statistic

1. Energy or Angular second moment

\[ \varepsilon = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij}^2 \]  \hspace{1cm} (3.23)

where \( G^{-2} < \varepsilon \leq 1 \). \( \varepsilon \) take the value \( G^{-2} \) for a uniform distribution over \( C \), and the value 1 if only one cell is nonzero. It will be large when \( C_{ij} \) are concentrated in some position in the GLCM. While a near random or much more local variation will have low energy.

2. Entropy

\[ s = -\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij} \log C_{ij} \]  \hspace{1cm} (3.24)

where \( 0 \leq s \leq \log G^{-2} \). \( S \) takes the value \( \log G^{-2} \) for a uniform distribution over \( C \), and the value 0 if one cell is nonzero. The entropy is larger for an image with an evenly distributed GLCM, therefore, a near random or orderless image will have a larger entropy (Newsam and Kamath, 2004 and Barber et al., 1993).

3. Maximum Probability

\[ \mu = \max C_{ij} \]  \hspace{1cm} (3.25)

where \( G^{-2} \leq \mu \leq 1 \). \( \mu \) takes the value \( G^{-2} \) for a uniform distribution over \( C \), and the value 1 if only one cell is nonzero.

4. Correlation or Autocorrelation

\[ \rho = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{(i - \mu_x)(j - \mu_y)C_{ij}}{\sigma_x \sigma_y} \]  \hspace{1cm} (3.26)

where \(-1 \leq \rho \leq 1 \). \( \rho \) takes the value 1 if only value on the main diagonal of \( C \) are nonzero and the value 0 if the gray value are uncorrelated. It measures how close the GLCM is to the main diagonal. The diagonal and near diagonal entries of a GLCM will be large for smooth texture or image that composed of patches with the same or similar pixel value, on the other hand, the off-diagonal entries will be large for rough texture or image in which the pixel values vary locally (Newsam and Kamath, 2005). So if it is high value the image will be complex than the lower value (Rikxoort, 2004).
5. Diagonal moment

\[ D = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)(i + j - \mu_x - \mu_y)C_{ij} \]  

(3.27)

the diagonal moment basically measure the difference in the correlation for high gray level and for low gray levels (Carstensen, 2002). If the low gray levels or bright areas are rougher and the high gray levels or dark areas are smooth then the diagonal moment will have positive value. If the dark areas are rougher and the bright are smooth, then the diagonal moment is negative. At the same time, if there is no difference between correlation in the dark and bright areas then the diagonal moment is close to zero.

Beside the second order statistical approach, the alternations of the usual co-occurrence matrices are also the sum and difference matrices.

3.2.3.3 Gray level difference histogram (GLDH)

GLDH is a histogram of the absolute differences of gray levels from pairs of pixels. It is computed from GLCM by summing the two-dimensional density \( C_{ij} \) over constant value of \(|i-j|\). The GLDH can be concerned as a histogram of the distance to the main diagonal in the GLCM (Carstensen, 2002)

\[ D_k = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij} \text{ where } |i - j| = k \quad k = 0, \ldots, G - 1 \]  

(3.28)

![Figure 3.8: the GLDH calculation methodology according to GLCM.](image)

1. Difference energy

\[ D_\varepsilon = \sum_{k=0}^{G-1} D_k^2 \]  

(3.29)

where \( G^{-1} < D_\varepsilon \leq 1 \)
2. Difference entropy

\[ D_s = -\sum_{k=0}^{G-1} D_k \log D_k \]  

(3.30)

where 0 \leq D_s \leq \log G

3. Inertia, contrast or variogram

\[ I = \sum_{k=0}^{G-1} k^2 D_k = 2\sigma^2(1 - \rho) \]  

(3.31)

where \( \sigma \) is the gray level variance and \( \rho \) is the correlation. The contrast is large for a GLCM with larger off-diagonal values. Contrast value is larger when the pairs of pixels being compared have different values with quickly varying intensities on an image (Cooper, 2004; Newsam and Kamath, 2005)

4. Inverse difference moment (Local homogeneity)

\[ IDM = \sum_{k=0}^{G-1} \frac{D_k}{1 + k^2} \]  

(3.32)

It will be larger for a GLCM with large diagonal value, therefore, it will higher for an image with constant or near constant patches (Newsam and Kamath, 2005). That is when the pixels of crystal structure on image are the same. The IDM value will be 1 and relied along diagonal of GLCM.

3.2.3.4 Gray level sum histogram (GLSH)

GLSH is a histogram of the sum of pairs of pixels. It is computed from the GLCM by summing the two-dimensional density \( C_{ij} \) over constant value of \((i+j)\), i.e.,

\[ S_k = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij} \text{ where } |i + j| = k \quad k = 0, \ldots, 2G - 2 \]  

(3.33)

\[ S_A = \sum_{k=0}^{2G-2} kS_k = \mu_x + \mu_y \]  

(3.34)
Figure 3.9: the GLSH calculation methodology according to GLCM.

1. *Sum energy*

   \[ S_e = \sum_{k=0}^{2G-2} S_k^2 \]  \hspace{1cm} (3.35)

   where \((2G - 1)^{-1} < S_e \leq 1\)

2. *Sum entropy*

   \[ S_s = -\sum_{k=0}^{2G-2} S_k \log S_k \]  \hspace{1cm} (3.36)

   where \(0 \leq S_s \leq \log (2G - 1)\)

3. *Sum variance*

   \[ S_v = \sum_{k=0}^{2G-2} (k - S_A)^2 S_k^2 = 2\sigma^2 (1 - \rho) \]  \hspace{1cm} (3.37)

   where \(\sigma\) is the gray level variance and \(\rho\) is the correlation

4. *Cluster shade*

   \[ A = \sum_{k=0}^{2G-2} (k - S_A)^3 S_k \]  \hspace{1cm} (3.38)
5. Cluster prominence

\[ A = \sum_{k=0}^{2G-2} (k - S_A)^4 S_k \]  

when cluster prominence is low, it shows a peak around the mean value so an image will have a little variation in gray scales (Rikxoort, 2004).

The GLSH and GLDH are the subset of GLCM features. Their computation has lower storage requirements and lower computational complexity. So only their benefit way of comparing the two sets of textures is to determine the loss information when going through GLCM to these GLSH and GLDH (Carstensen, 2002). Conners et al. (1984) found that cluster shade and cluster prominence was a useful supplement to the GLCM and GLDH features. However, Ojala et al. (2001) noticed that the GLDH appeared to be much more powerful than GLSH.
3.2.4 Histogram matching

3.2.4.1 Introduction

Suppose that the brightness range within an image is very small, there may not be enough contrast to assure visibility, herewith the nature tone of the image has to be manipulated by increasing the contrast of its (Richards, 1993; Russ, 1995). The contrast of images can be enhanced by matching or transforming the histogram of the gray scale values to make it match a given distribution such as histogram equalization, Gaussian distribution and so on (Hilger, 2004). The other enhancement is that histogram matching makes it more insensible to lighting effect (Carstensen, 1992).

3.2.4.2 Image histogram

Image histogram is a preliminary step for successful manipulation which provides basic information about the appearance of an image. It describes the statistical distribution of gray levels in an image in terms of the number of pixels having the same gray level, or it consists of a graph indicating the number of times each gray level occurs in the image. Across the horizontal axis of this graph is the range of possible pixel intensity values, e.g., 0-255. The vertical axis represents a measure of the frequency of occurrence of each intensity value. In the case of an excessively dark or bright image, the gray level would be clustered to the extremes of the histogram. But in a well contrasted image, these levels would be well spread out the range and without significantly large bars at black or white (Patton et al., 2006; Scozzafava et al., 2004; Richards, 1993). As mentioned, one can see that the histogram is containing only tone and no spatial distribution information of gray levels throughout the image (Schowengerdt, 1983; Richards, 1993).

![Image histogram example](image.png)

Figure 3.10: 6x6 image and its histogram (Carstensen, 2002; Pp 25).

3.2.4.3 Principle of histogram matching

Histogram matching is a standard method which was found to be useful in visualizing, increasing the sensitivity of an image. The simplest and most used to enhance the contrast of the image is histogram equalization (Hilger, 2004; Gupta and Undrill, 1995; Carstensen, 2002). Gaussian distribution is another possibility and has been shown in Carstensen (1992).
**Histogram equalization**

The goal of histogram equalization is to make the histogram of an image as uniform as possible, that means it construct the new histogram of image intensity and compute new image values - directly from the cumulative frequency diagram- which can generate to be equalized from local probability distribution see figure 3.9 (Patton et al., 2006; Yoo, 2000; Belward and Valenzuela, 1991)

As before, let $h(x)$ is the histogram function of the original image and $ho(y)$ represents the modified histogram, which is to be uniform. If the image contains a total of $N$ pixels and there are $L$ histogram bins or brightness value, then each of the brightness value in the modified histogram should have a bar of $N/L$ pixels associated with it. Recall also that the bars in a discrete histogram have the values $ho(y)dy$. In the case of $L$ available brightness value, $dy = (L-1)/L$ so that a uniform histogram.

$$ho(y)(L-1)/L = N/L \quad (3.40)$$

giving $ho(y) = N/(L-1)$, therefore $\frac{dy}{dx} = \frac{d}{dx} (f(x)) = \frac{L-1}{N} hi(x) \quad (3.41)$

in which $y = f(x)$ is the transformation of brightness values that takes the original histogram of an image into a uniform histogram. Hence, call “$y$” is a scaling factor

$$y \sim \text{ scaling factor} = (L-1)/N \quad (3.42)$$

The histogram equalization transform, thereby, is the integral of the original histogram function times a scaling factor. The integral is just the continuous cumulative histogram. Look up table 3.1 and figure 10 that can be used to move histogram bars to new brightness value location which uniform as possible (Richards, 1993).
Table 3.1: Table generation for histogram equalization of figure 3.12 (Richards, 1993; Pp 99).

<table>
<thead>
<tr>
<th>Original brightness</th>
<th>Unscaled new value or original cumulative</th>
<th>Scaled new value</th>
<th>Nearest available brightness value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.63</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1.25</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>3.13</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>5.63</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>8.75</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>11.25</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>19</td>
<td>11.88</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>12.50</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>23</td>
<td>14.40</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>24</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>
Figure 3.12: example of histogram equalization (Richards, 1993; Pp 98).

The new brightness value location of a histogram bar is provided by finding its original location on the abscissa of the cumulative histogram(x) and then reading its unscaled new location(y) from the ordinate. Multiplication by the scaling factor, then, it is generated the required new value as show in the figure 3.10 and table 3.1. It can be done to re-map the brightness value to give a histogram that is as uniform as possible. Sometimes some bars from an original histogram being moved to either new nearest available brightness location or same new location (Richards, 1993; Russ, 1995; Schowengerdt, 1983). Therefore, some pixels that originally had different brightness values are now assigned the same value, which represent loss information, while other values that were once very close together have been spread out, leaving gaps in the histogram. When performing
histogram equalization on the image, it spreads the peak out while compressing another part of the histogram by assigning the same or very close brightness values. Herewith it is leading to possible loss of detail but this is compensated by increased contrast and average brightness (Russ, 1995; Mather, 2004). For example, in an image that contains a main dark region, the histograms equalization will not reproduce only an overall brightness, but also spread out the brightness values so that the small variations in an original are moved. Meanwhile, in the white or bright area, an overall brightness also is reduced, so more variation is showed (Russ, 1995; Mather, 2004).

**Gaussian distribution matching**

Richards (1993) refer to another application of histogram matching that sometime an image is required to match with the desired shape in order to give a modified image with particular distribution of brightness values. An example of useful shape is to match an image with Gaussian distribution. It contains a modified version with a few black and white areas. Most of the detail is located in the middle of gray range. The consider requirement of Gaussian matching are the mean of distribution, which is placed usually at the middle point of brightness scale. Note that the procedure use the cumulative histogram of the original image to obtain new brightness values is reading ordinate value (y axis of original cumulative histogram) corresponding to original brightness value on the abscissa. The new values, then, are entered into the ordinate value of the cumulative reference histogram, i.e., Gaussian cumulative histogram. Finally, the brightness values (for the bars of the original histogram) are read from the abscissa. That means the cumulative reference histogram is used in reverse function. Look up figure 3.13 for another possibility matching.
Figure 3.13: an illustration of the steps of histogram matching (Richards, 1993; Pp 104).

Notice that the new brightness value of histogram matching result, value 2, came from cumulative reference histogram. Meanwhile, number of pixels of its obtained from the original histogram, i.e., brightness value of cumulative original histogram equal to 6 has a number of pixels from original histogram equal to 2.
3.2.5 The color to gray level transformation

3.2.5.1 Introduction

Color refers to the range of human visual sensations that can be produced by mixtures of various wavelengths of visible light as it interacts with eye, brain and human experience. Furthermore, it is also a basic feature that can often be used for various information in an image (Gose et al., 1996; Park et al., 2004; Rikxoort, 2004).

The human visual system is being on human eye perception. The range of seeing for humans starts around 630 nm, which is red color. After that, come green and blue tones and finally violet which has approximately 400 nm at wavelength. There are many parts being responsibility of human eye which reflect light and generate the image, which is interpreted by the brain. Light reaching the eye passes through the pupil and is focused onto retina by the lens (Aguila, 2004; see figure 3.14).

![Anatomy of the Eye](image)

Figure3.14: anatomy of the human eye (Aguila, 2004; Pp 11).

The retina contains cells of light sensitive photoreceptors, namely cones and rods, figure 3.15. There are around 100 million rod shaped cells on the retina and 5 million cone shaped cells. Cones detect color such as red, green, blue, or hue and perception of the degree of saturation of each hue as well as the intensity level, however, it vision requires a higher illumination level than rods cells. Meanwhile, rods cells respond to light or brightness and do not provide any color information (Mather, 2004; Aguila, 2004).

![Rods and Cones Cell](image)

Figure3.15: rods and cones cell (Aguila, 2004; Pp 12).
A model of color space can be derived from the idea that color are set by adding together differing amounts of red, green and blue light. Additionally, in the field of image analysis, in many situations, even though color images are acquired but the absolute color information may be not useful for image analysis so it is not necessary to store the entire color image. Thereby, various kinds of color model are available (Russ, 1995).

### 3.2.5.2 RGB model

RGB model is a common basically to scan and display normal color of television and computer. The name is given by the abbreviation of red, green and blue. There are a peak sensibility in the wavelengths around 630 nm (red), 530 nm (green) and 450 nm (blue) and the vector represent black color is (0,0,0), while (1,1,1) corresponds to white as well as (0.5,0.5,0.5) is gray color, red color is (1,0,0), green is (0,1,0) and blue is (0,0,1) (Aguila, 2004).

![RGB cube](Mather, 2004; Pp 122).

If the proportions of red, green and blue are equal at each point, a gray color would be seen and a color picture is obtained when the amounts of red, green and blue at each point are unequal (figure 3.16). So the color at any pixel is represented by a point that is located away from the black-white diagonal line (Mather, 2004; Lillesand et al., 2004).

### 3.2.5.3 HIS model

An alternative to come closer to the nature of human color perception limits, especially to detects color differences, is HIS (Hue-Saturation-Intensity) or HSV (Hue-Saturation-Value), HLS (Hue-Lightness-Saturation) or La*b* model as they are closely related to each other on the concept of tint, shade and tone (Russ, 1995).

When observing the RGB cube along the main diagonal that goes from white to black, the cube contour seem like a hexagon (figure 15a,b and 16). The border of the HIS hexcone give the different hues, where hue is the wavelength of the color such as red, green, blue. The hexcone hues have a separation of 60 degree from each other, red is at origin, yellow is 60°, green is 120°, cyan in 180°, blue in 240° and magenta in 300° (Aguila, 2004).

Saturation or chroma is measured along a horizontal axis that means the distance from white to that point at the same angle (Gose et al., 1996) or the degree of purity of a color is its (Mather, 2004). Saturation varies from 0 to 1 and represents the purity of a selected hue, where 0 means color is gray and 1 is maximal is its or highly saturated.
Intensity or luminance measures brightness of a color or the color lightness that varies from 0 (black) to 1 (white). Therefore, when \( S = 1 \) and \( I = 1 \), pure hue are acquired, meanwhile white is at the point \( S = 0 \) and \( I = 1 \) (Aguila, 2004).

Figure 3.17: RGB color cube including HIS co-ordinate system (Carstensen, 2002; Pp 39).

Figure 3.18: RGB color cube including HIS co-ordinate system. (Carstensen, 2002; Pp 39).

Figure 3.19: HIS hexcone (Mather, 2004; Pp 122).

Hue is represented by the top edge of a six-sided cone (hexacone) with red at 0°, green 120°, blue 240°. Pure unsaturated and maximum intensity colors are around the top edge
of the hexacone. The white color produces less saturation. Intensity is shown as a distance above the top of the hexcone (figure 3.19). The relationship between RGB and HIS is also shown in table 2 (Russ, 1995).

Table 3.2: conversion from RGB color co-ordinate to HIS co-ordinate (Russ, 1995; Pp 39).

| H | \( \frac{\pi}{2} - \arctan\left(\frac{2(R - G - B)}{G - B}\right) + \frac{\pi}{2} \) \( \forall G < B \) |
| I | \( \frac{R + G + B}{3} \) |
| S | \( 1 - \frac{\min(R, G, B)}{I} \) |

Because color has a three dimensional structure and the texture analysis can only handle one dimensional “gray level” data, the color scan has to be transformed to a gray level.
3.2.6 Region of interest (ROI) and texture analysis

Region of interest (ROI) is the region of the image specified for texture analysis in order to examine which zones contain the relevant information for classification purposes. A main region that is a primary of the general study of biocrystallogram image is circle-ROI (Andersen et al., 1999). It represents the radius of a circle around the geometric center. Thereby 0% equals to the radius of 0 from the geometric center while 100% equals the whole image area. Circle-80, therefore, is a circle of 80% radial extension around the geometric center.

The relation between the region of interest and their GLCM (Figure 3.20 and 3.21) can be shown that when the ROIs are changed, then the numbers of pixels that are perceived and calculated as GLCM are quite difference. As the numbers of pixels in that specified area of an image contribute to GLCM. The ROI changes the GLCM as a whole.

![Circle-40](image1)
![Circle-80](image2)
![Circle-100](image3)

**Figure 3.20:** the GLCM (with Gaussian distribution matching) of different ROIs of sample H.
Figure 3.21: the GLCM (with Gaussian distribution matching) of different ROIs of sample G.
Figure 3.22: the texture analysis results of the comparison sample G and H by plotting F and P value over the different ROIs.

From the steadiness of the curve in the graph, we decided which variable we can trust to showing something of the image (figure 3.22).

3.2.7 Texture analysis applications

The texture or image analysis technique has been interested and utilized in many application domains. In some of the mature domains (such as remote sensing), the texture analysis approach has already played a major role, while in the other disciplines are being found and studied (Tuceryan and Jain, 1998). Thereby, a wide range of applications of this method to characterize images shall give a briefly overview.

The plant tissue culture technique was followed in the area of agriculture by Ibaraki and Kenji (2001) the desirable properties of culture cells or cell lines was successfully selected by using image analysis to enhance microscopy texture and select culture cell by relate to the color, growth rate shape, aggregate size distribution.

Regarding to the medical application, nowadays, image analysis are increasing important in all fields of medical science. For instance, they are using both of first- order and second order statistics to analyses their image, Gupta and Undrill (1995); Lefebvre et al.(2000) using them to followed breast cancer. Vince et al.(2000) have also been used for diagnostic coronary plaques in intravascular ultrasound images. The different bones mineral density image by using texture analysis was also studied by Materka and Tuliszkiewicz (1999).

In the field of remote sensing, as mentioned, it has been played an important role in the field of image analysis. Coburn and Robert (2004) followed texture analysis both of first and second order for forest stand classification. Cooper (2004) use GLCM measured the strength of the earth’s gravity field for mineral exploration purpose. Jobanpatra and Clausi (2004) analyzed SAR images which used for classification of land use categories such as
water, agriculture areas by using also texture analysis. Further application, e.g., texture aerial photographs, carpet and wood field, can be found more in summarized from Tuceryan and Jain, 1998 and Carstensen, 2002.

In the field of food technology and quality, Fernandez et al. (2005) analyzed the effect of drying dehydrated apple by using second-order statistical. Mezreb et al. (2003) investigated the structure properties of both corn and wheat extrudes, the result showed that the image analysis based on statistical approach associated with physical as well as chemical measurements. It would appear to be an effective means for controlling extrude texture, processing parameter and formulation. Faucitano et al. (2005) and Basset et al. (2000) have been used GLCM matrix to characterize quality of meat sample images such as measuring boving meat or pork marbling characteristics. More detail of texture analysis application in the food industrial can be found in Zheng et al. (2006).

Concerning the biocrystallization method, as the image analysis should be also taken into account all the characteristics of biocrystallogram pictures which were generated from differing food quality. There are many scientists who have been studied and correlated this approach to illustrate the image from biocrystallization method. Andersen et al. (2001); Andersen and Busscher (2001) and Andersen et al. (1999) studied and worked by using computerize image analysis to discriminate biocrystallograms from several treatments, e.g., liquid-N-freezing of carrot extracts, mineral N and light intensity on the pictomorphological properties of barley as well as PVP biocrystallograms. Their results show that the image analysis can differentiate all of sample in questions.
Chapter 4: Study methodology and result

This chapter is able to be divided into 3 main parts, namely,

(1) Optimize the statistical model is to set up a suitable statistical model which can describe all the effects that contribute to the experiment. Because the old statistical model which was prior established, calculated the P value with a very conservative approach. Aftermath, this suitable statistical model is embedded with the main programs of this research, i.e., the ACIA and R statistical programs so that they can co-work properly for the differentiation of all samples in questions.

(2) Investigate the effect of image parameters on the texture analysis of the biocrystallogram images, i.e., region of interest (ROI), color transformation and histogram matching. These parameters may play roles on the analysis of interested images. The best set of these combined 3 parameters are selected and applied then to differentiate samples, i.e., wheat and carrot by using the texture analysis approach.

(3) Consider the strongest effect of texture parameter with the visual evaluation criteria that have been developed by triangle researcher group (University of Kassel, Germany; Louis Bolk Institute (LBI), Netherlands and Biodynamic Research Association Denmark (BRAD), Denmark) in order to depict how the relation of the texture parameter and visual characteristic on an image is.
4.1 Optimization of the statistical model

4.1.1 Introduction

Essentially, this main part will involve the following consecutive steps;

(a) An existing statistical model should be optimized by using linear mixed effect model (lme), as lme can illustrate all of the treatment effects that are able to contribute to the experiment.

(b) Verify the statistical model which is obtained from (a) through comparing the results of ANOVA from the different two statistical programs (R and SAS) in order to ensure that the statistical model which is generated can precisely works.

(c) Embed this improved model into the ACIA program (one of the main programs that is applied to calculate the texture parameters in this research) until this model is reasonably to use for the next step.

So this chapter will describe not only how to implement and verify the statistical model which fits best with the experiment, but also coupling with ACIA and R programs until they are able to work properly.

With regard to this research, there are several experiments which are studied. Hence, a prototype data should be selected in order to implement the correct statistical model for all experiments. This prototype data is a subset of carrot from University of Kassel in 2004. The investigation is done based on the old subset of parameter ROI, color transformation and histogram matching which are characterized as circle-60, equal color and Gaussian matching, respectively. The experimental design synopsis can be interpreted that 4 different samples, e.g., carrots with the combination of 2 different levels nitrogen fertilizer and 2 varieties, call them A, B, C and D. Each of these carrots is sampled and then juiced, for example, sample A generates then the solutions A1, A2. The step of sampling called the step of sample sampling and juicing, i.e., sample preparation or sample sampling. After that, from each of these repetition solutions was originated 6 pictures in chamber 1 and 6 pictures in chamber 2. Aftermath, 15 texture parameters were analyzed per picture. The experiment was repeated on 2 days, an experimental diagram was shown in figure 4.1. Then, sum entropy- one of the 15 texture parameters - was determined through the Analysis of Variance (ANOVA) results.
Figure 4.1: the 1st day experimental diagram of carrot from University of Kassel in 2004 as a prototype experimental design, the 2nd is repeated in the same method.

Table 4.1: the overall basic details of a prototype data of carrot from University of Kassel in 2004

<table>
<thead>
<tr>
<th>2 days</th>
<th>Fixed effect (or Random effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 samples</td>
<td>Fixed effect</td>
</tr>
<tr>
<td>2 sample sampling</td>
<td>Random effect</td>
</tr>
<tr>
<td>2 chambers</td>
<td>Fixed effect</td>
</tr>
<tr>
<td>Pictures in chamber</td>
<td>Random effect (repeated measurement)</td>
</tr>
</tbody>
</table>

From above details, the experimental design shows that it is comprised of 2 main effects, i.e., fixed and random effects or so called “mixed effect”. Moreover, it is also repeated in measurement (table and figure 4.1). Fixed effects are the parameters which are associated with an entire population or with certain repeatable level of experimental factor, meanwhile, random effects are associated with the individual experiment units draw at random from a population (Pinheiro and Bates, 2000).

This experiment is necessary to implement a suitable statistics so that it can describe all of treatment effects which are contributed. By that, it is utilized through the statistical program to differentiate all of the samples in questions, i.e., wheat and carrot. The methodology for setting up a mixed effect model can be found in Piepho et al. (2003); Piepho et al. (2004) and in chapter 3.
4.1.2 Statistical model implementation

As stated before, all of the experiments are designed with mixed effects cooperating repeated measurements. Therefore, the improved statistical model should be implemented and set up by using mixed model methodology.

Piepho et al. (2003) presented a hitchhiker’s guide to set up mixed models for randomized experiments which included the repeated measurements. The measurements were taken at different points of time and/or space are correlated. Herewith, the model for this research was generated by following their procedure which is processed as follows

1. When is a factor random or fix?

Random factor is defined when the observed levels can be regarded as randomly sampled from a population (e.g., environments and sampling units), alternatively, a factor is random if it represents a randomization unit, e.g., plot. Otherwise the factor will be fixed, e.g., treatment.

Herewith, recall to a prototype experimental data, it can be argued that

a) Fixed effects are days, samples and chambers.

b) Random effects are sample preparations and replications in chamber.

2. Separate 2 types of factor

There are 2 main types of factor which are important for setting up mixed model, they are block and treatment factors. Block factors comprise of randomly selected sampling units and need to uniquely identify each observational unit, i.e., block factor are innate to the observational units. Contrarily, treatment factors are treatments which are chosen by the investigator to answer a research question. Therefore, this research experiment can be classified as

a) A treatment is sample

b) A block factor is

day/ sample/ sample preparation/ chamber/ replication.

This block factor means an observational unit that originates from a replicated biocrystallogram image in a chamber which related to a sample preparation within a sample and a day, respectively. Furthermore, the replications of sample were imported into the same 2 chambers, and then the experiment is repeated on several days. From this consequence, the day and chamber are crossed effect because each level of factors has a fundamental property that is the same for every level of the other factors in the experiment (Mason et al., 2003), that means each levels of sample and sample preparation within a day is the same for 2 levels of chambers. It is, thereby, also an additional term which must be postulated in the full statistical model. Let it conclude all of factors which are important to set up statistical model as follows
1) A treatment is sample.

2) A block factor is

day/sample/sample preparation/chamber/replication

It can be derived into

day + T1 + T3 + T4 + T5

where the first order interaction, i.e., T1 is day. sample.
the higher order interactions comprise of

T3 is day. sample. sample preparation.
T4 is day. sample. sample preparation. chamber.
T5 is day. sample. sample preparation. chamber. replication.

3) An additional factor is chamber x day (crossed effect). It can be also derived into chamber + day + T2 where T2 is day. chamber.

More details of derivation and the meaning of operators, such as (/) or (.) can be found in Piepho et al. (2003).

As mentioned earlier, the mixed statistical model is able to be written by including those above 3 factors and comprised of

day + sample + chamber + T1 + T2 + T3 + T4 + T5

First order interactions are T1 and T2,

where T1 is day. sample.
T2 is day. chamber.

Higher order interactions are T3, T4 and T5,

where T3 is day. sample. sample preparation.
T4 is day. sample. sample preparation. chamber.
T5 is day. sample. sample preparation. chamber. replication.

After that, the full mixed effect statistical model with crossed factor can be summarized into

\[ Y_{ijklm} = \mu + \alpha_i + \beta_j + \eta_k + (\alpha\eta)_{ik} + (\alpha\beta)_{ij} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijklm} \quad (4.1) \]

Where \( Y_{ijklm} \) is the response of \( i^{th} \) day, \( j^{th} \) sample, \( k^{th} \) chamber, \( l^{th} \) sample preparation and \( n^{th} \) replication

\( \mu \) is the general or overall mean
\( \alpha_i \) is the mean effect of \( i^{th} \) day
\( \beta_j \) is the mean effect of \( j^{th} \) sample
\( \eta_k \) is the mean effect of \( k^{th} \) chamber
4.1.3 Statistical model verification

The statistical model must be verified and proved in order to ensure this model can be precisely generated. Therefore, this part aims (a) to set up the statistical model within the statistical program, i.e., R program version 2.1.0, as it is the main program which has been run for this research and (b) to compare its ANOVA result with another statistical program, namely, SAS program version 9.1. The statistical model is reasonably verified for the next step when the ANOVA results from both statistical programs turn out the same results. Otherwise it must be finer-tuned until the same ANOVA results are achieved.

Firstly, the statistical model must be embedded into R program. From previous literature reviews, Pinheiro and Bates (2000); William and Ripley (2002) fitted the random effects model with lme (linear mixed effects models). This lme model has a list of arguments separated by ",". The main ones are the first three arguments that specify the fixed effect, data and random effect. The fixed effect refers to the fixed effects part of the statistical model. Data is the object containing the data which the model should be fit. The last one is random effect that describes the random effect part as well as grouping structure of the model. Detail of their forms, all the possible arguments and examples can be found by the above authors. Thereby, the statistical model which was earlier generated from 4.1.2 should be put into the R program by following their methodology. The used lme model with R program neglected the higher order interactions as their variance components are either zero or very small (the complete random grouping design was examined via lmer formula). Additional, the lme and lmer which compared the highest F value for all samples present very resembling F and P values (appendix L). From this consequence, the complete lme model with the first order interactions should be established as follows;

\[
\text{lme( value~ day+sample+chamber+day:sample+day:chamber, data, random~1|groupVektor) .}
\]

The first argument indicates the response which is named “value” means one value of the 15 texture parameters. Then there are 5 fixed effects, i.e., day, sample, chamber, interaction between day and sample as well as interaction between day and chamber. The second argument is the name of the data which will be used for calculation. The last one refers to a single random effect for each group and that grouping is named groupVektor. Due to this experimental design is mixed effect with crossed factor. Thereby, it is necessary to create a grouping structure that has unique levels for each factor within each level of treatment. From this reason, groupVektor is produced by using getgroup function (Pinheiro and Bates, http://biostat.hitchcock.org/FacultyandStaff/OnlineManuals/PDF%20Files/Imesas.pdf) in order to set a grouping structure of day/sample preparation/sample/chamber. That means each value will be calculated and used by unique levels of
From the above complete lme model, it is the full lme model which will be applied for this research—that means it consists of days, samples, sample preparations and chamber. But, in fact, some of the interest experiments might not be set up like above mentioned, but they have a reduced data set, i.e., 1 day or 1 chamber etc. Its goal is that lme model is more flexible for using object oriented programming in R program, hereby, 16 formula cases of it are fitted and shown in appendix B.

### 4.1.3.1 The comparison of ANOVA results between SAS and R program.

As earlier mentioned on the verification, the comparison of ANOVA results from the statistical R and SAS program which input the same statistical model is necessary.

Through comparing their ANOVA, the data which are used for this comparison is a subset of carrot from University of Kassel in 2004. A value of the image analysis variable named *sum entropy*. Both of the ANOVA results are shown in the figure 4.2 and 4.3

```r
> lmeRet4<-lme(Value~Sample+Chamber+Day+Sample:Day+Day:Chamber,random=~1|Day/repIndexVektor/SamplePreps/Chamber)
> gg4<-getGroups(lmeRet4,level=4)
> lmeq4<-lme(Value~Sample+Chamber+Day+Sample:Day+Day:Chamber,random=~1|gg4)
> print(anova(lmeq4))

<table>
<thead>
<tr>
<th>numDF</th>
<th>denDF</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>1</td>
<td>112.7205</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Sample</td>
<td>3</td>
<td>22</td>
<td>17.5260</td>
</tr>
<tr>
<td>Chamber</td>
<td>1</td>
<td>22</td>
<td>3.4212</td>
</tr>
<tr>
<td>Day</td>
<td>1</td>
<td>22</td>
<td>8.6200</td>
</tr>
<tr>
<td>Sample:Day</td>
<td>3</td>
<td>22</td>
<td>0.1067</td>
</tr>
<tr>
<td>Chamber:Day</td>
<td>1</td>
<td>22</td>
<td>1.5344</td>
</tr>
</tbody>
</table>

> print(levels(gg4))
```

Figure 4.2: the ANOVA result from R program version 2.1.0.
Figure 4.3: the ANOVA result from SAS program version 9.1, its result was calculated by STATCON Company, Witzenhausen, Germany.
The ANOVA results from R and SAS programs show that their particular degree of freedom (df), F values are nearly same, but P values are different. This is caused by the difference in denominator degree of freedom (denDf) which may be originated from the different calculation methods from each program. By that, SAS program uses Kenward-Roger method, meanwhile R program method is referred in Pinheiro and Bates (2000). So it is difficult to decide exactly how this denDf should be done for the models (Bates, 2005). However, the denDf results illustrate that they produce identical number within each program that means their difference are not from the statistical model. Therefore, it can be summarized that the mixed effects statistical model which was implemented, up-to-now, is reasonably verified and properly to use or so called “refined statistical model”. The R program calculation gives a more conservative P value. With this it is on the safe side.

4.1.4 The ACIA program investigation: old and new versions with refined statistical model

The ACIA program version 1.2.3 is the major program for calculating all texture parameters in this research. At this step, ACIA program is calibrated and modified with the refined statistical model until it can properly work. The carrot from University of Kassel in 2004 is also used as input data for the old (without crossed factor) and refined statistical model (with crossed factor). The parameters which were checked are the basic data and the texture analysis by plotting F and P value over the different ROIs (region of interest) that vary from circle-30 to circle-100 (figure 4.4a and b). The entire data should unveil the same. At the same time, F and P values should provide also the same pattern behavior. Furthermore, the refined statistical model should turn out better differentiation than the old one.

Figure4.4: (a) the region of interest (ROI) of carrot from University of Kassel data at circle-50 and circle-80. (b) The texture analysis by plotting F and P value versus ROI, vary from circle-30 to circle-100.
The texture analysis is calculated on the specified area of image where is initially expanded from the geometrical center, such as, circle-50 and circle-80 (figure 4.4a; they are not located in the blue color area). The F and P value of texture parameters were plotted over the different ROIs and they can be shown in which area the pictures are most different (figure 4.4b).

The results of this section divided into 4 data sets of carrot from University of Kassel in 2004. The first data shows the complete experimental design (table 4.2a, b and c). The second data shows the reduced data that is designed merely 2 days and 1 chamber (table 4.3a, b, and c). The third experiment is designed with 1 day and 2 chambers (table 4.4a, b, and c) and the last is designed with 1 day and 1 chamber (table 4.5a, b and c).

As all of the data results disclose the same behavior, so this can be deduced as follows (a) the investigation (table 4.2, 4.3, 4.4 and 4.5a) shows that the basic data are same, as they are from the same source, i.e., carrot from University of Kassel in 2004. Additionally, the lme statistical models that were analyzed for each data set are correct and (b) the comparisons between an old and refined statistical models (table 4.2, 4.3, 4.4, and 4.5b, c and d) are found that the refined statistical model does not provide only similar pattern for F and P value, but also improves the statistical result, as its F value and P values show higher significance (except the result of 1 day and 1 chamber that shows an identical result as the lme model are the same). As described before, it can be summarized that the refined statistical model performs better performance than the old model and suitable for future work by using ACIA program.
Table 4.2a: the overall basis data of carrot from University of Kassel in 2004 with complete experimental design.

<table>
<thead>
<tr>
<th>Series name</th>
<th>Assignment</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Refined statistical model (with crossed effect)</strong></th>
<th><strong>OLD statistical model (without crossed effect)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>analysis option: DF.CR Day fix crossed effect</td>
<td>analysis option: DF.OLD</td>
</tr>
<tr>
<td>name for analysis: Default.DF.CR.r.G1G2H1H2all.ROI30-100</td>
<td>name for analysis: default.G1G2H1H2all.30-100</td>
</tr>
<tr>
<td>Statistical model: lme.formula Value~Quality+Chamber+Day+Chamber:Day + Quality:Day, random ~1</td>
<td>groupvektor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1-1. Übersicht Basisdaten:</th>
<th>Name</th>
<th>Mean</th>
<th>Value</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nprocess</td>
<td>72</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nkammern</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NdatumDerserie</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NserienNamen</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Npictures</td>
<td>216</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAssignments</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NStatisticCase</td>
<td>2.011</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NrSamplePrepsPerDay</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NrRegPerDay</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>NO</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1-1. Übersicht Basisdaten:</th>
<th>Name</th>
<th>Mean</th>
<th>Value</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nprocess</td>
<td>72</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nkammern</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NdatumDerserie</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NserienNamen</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Npictures</td>
<td>216</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAssignments</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NStatisticCase</td>
<td>2.011</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NrSamplePrepsPerDay</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NrRegPerDay</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>NO</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with complete experimental design.

<table>
<thead>
<tr>
<th>Refined statistical model (with crossed effect)</th>
<th>OLD statistical model (without crossed effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Table 4.2c: F and P values of different nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with complete experimental design (1, 2 are Rodelika variety).
Table 4.2d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with complete experimental design (3,4 are Rothild variety).
Table 4.3a: the overall basis data of carrot from University of Kassel in 2004 with 2 days and 1 chamber experimental design.

<table>
<thead>
<tr>
<th>Series name</th>
<th>Assignment</th>
</tr>
</thead>
</table>

**Refined statistical model (with crossed effect)**

<table>
<thead>
<tr>
<th>analysis option: DF.CR Day fix crossed effect</th>
<th>OLD statistical model (without crossed effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>name for analysis:</td>
<td>name for analysis:</td>
</tr>
<tr>
<td>default.DF.CR.r.2Day1Chamber.ROI30-100</td>
<td>default.DF.OLD.r.2Day1Chamber.ROI30-100</td>
</tr>
</tbody>
</table>

**Statistical model:**

<table>
<thead>
<tr>
<th>lme.formula Value~Quality+Day+Quality:Day</th>
<th>lme.formula Value~Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>random~1</td>
<td>grouvektor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1-1. Uebersicht Basistaten</th>
<th>Table 1-1. Uebersicht Basistaten</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Name</td>
</tr>
<tr>
<td>MeanValue StdDev</td>
<td>MeanValue StdDev</td>
</tr>
<tr>
<td>Nprocess</td>
<td>Nprocess</td>
</tr>
<tr>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Nkammer</td>
<td>Nkammer</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NdatumDerSerie</td>
<td>NdatumDerSerie</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>NserienNamen</td>
<td>NserienNamen</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Npictures</td>
<td>Npictures</td>
</tr>
<tr>
<td>108</td>
<td>108</td>
</tr>
<tr>
<td>Nassignments</td>
<td>Nassignments</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>NStatisticCase</td>
<td>NStatisticCase</td>
</tr>
<tr>
<td>2.001</td>
<td>2.001</td>
</tr>
<tr>
<td>NrSamplePrepPerDay</td>
<td>NrSamplePrepPerDay</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>NrRepPerDay</td>
<td>NrRepPerDay</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>GlobnoDays</td>
</tr>
<tr>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>
Table 4.3b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with 2 day and 1 chamber experimental design. The same ANOVA result is because of 1 chamber.

<table>
<thead>
<tr>
<th>Refined statistical model (with crossed effect)</th>
<th>OLD statistical model (without crossed effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph 1" /></td>
<td><img src="image2.png" alt="Graph 2" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph 3" /></td>
<td><img src="image4.png" alt="Graph 4" /></td>
</tr>
</tbody>
</table>
Table 4.3c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 2 days and 1 chamber experimental design (1, 2 are Rodelika variety).

<table>
<thead>
<tr>
<th>Refined statistical model (with crossed effect)</th>
<th>OLD statistical model (without crossed effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Table 4.3d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with 2 days and 1 chamber experimental design (3,4 are Rothild variety).
Table 4.4a: The overall basis data of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design.

<table>
<thead>
<tr>
<th>Series name</th>
<th>Assignment</th>
</tr>
</thead>
</table>

**Refined statistical model with crossed factor**

- **Analysis option**: DF.CR Day fix crossed effect
- **Name for analysis**: default.DF.CR.r.1Day2Chamber.ROI30-100
- **Statistical model**: lme.formula Value~Quality+Chamber random~1|groupvektor

**OLD statistical model without crossed factor**

- **Analysis option**: DF.OLD
- **Name for analysis**: default.DF.OLD.r.1Day2Chamber.ROI30-100
- **Statistical model**: lme.formula Value~Quality+Chamber random~1|ffSamplePrep

---

**Table 1-1. Uebersicht Basisdaten**

<table>
<thead>
<tr>
<th>Name</th>
<th>Mean</th>
<th>Value</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nprocess</td>
<td>24</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nhammen</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NdatumDerSerie</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NaerienNamen</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Npictures</td>
<td>72</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nassignments</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NStatisticoCase</td>
<td>2.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NrSamplePrepPerDay3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NrRepPerDay</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>NO</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**Table 1-1. Uebersicht Basisdaten**

<table>
<thead>
<tr>
<th>Name</th>
<th>Mean</th>
<th>Value</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nprocess</td>
<td>24</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nhammen</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NdatumDerSerie</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NaerienNamen</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Npictures</td>
<td>72</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nassignments</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NStatisticoCase</td>
<td>2.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NrSamplePrepPerDay3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NrRepPerDay</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>NO</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 4.4b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design.

Refined statistical model with crossed factor

OLD statistical model without crossed factor
Table 4.4c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design (1, 2 are Rodelika variety)

<table>
<thead>
<tr>
<th>Refined statistical model with crossed factor</th>
<th>OLD statistical model without crossed factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Graph 1" /></td>
<td><img src="image2" alt="Graph 2" /></td>
</tr>
<tr>
<td><img src="image3" alt="Graph 3" /></td>
<td><img src="image4" alt="Graph 4" /></td>
</tr>
</tbody>
</table>

Refined statistical model with crossed factor:

OLD statistical model without crossed factor:
Table 4.4d: F and P values of different nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) of carrot from University of Kassel in 2004 with 1 day and 2 chambers experimental design (3,4 are Rothild variety).

<table>
<thead>
<tr>
<th>Refined statistical model with crossed factor</th>
<th>OLD statistical model without crossed factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics: 2.01 Assignment= 1Day/Chamber</td>
<td>Statistics: 2.01 Assignment= 1Day/Chamber</td>
</tr>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

Refined statistical model with crossed factor

OLD statistical model without crossed factor
Table 4.5a: the overall basis data of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design.

<table>
<thead>
<tr>
<th>Series name</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACIA.test.datensatz.Moehre.Fleck.2004.11590.GJ.GK.GL</td>
<td>1Day1Chamber</td>
</tr>
</tbody>
</table>

**Refined statistical model (with cross factor)**

<table>
<thead>
<tr>
<th>analysis option: DF.CR Day fix crossed effect</th>
<th>OLD statistical model without cross factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>name for analysis:</td>
<td></td>
</tr>
<tr>
<td>default.DF.CR.r.1Day1Chamber.ROI30-100</td>
<td>default.DF.OLD.r.1Day1Chamber.ROI30-100</td>
</tr>
</tbody>
</table>

**Statistical model:**

<table>
<thead>
<tr>
<th>lme.formula</th>
<th>value~Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>random~1</td>
<td>groupVektor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lme.formula</th>
<th>value~Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>random~1</td>
<td>SamplePrep</td>
</tr>
</tbody>
</table>

---

Table 1-1. Uebersicht Basisdaten:

<table>
<thead>
<tr>
<th>Name</th>
<th>Mean</th>
<th>Value</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nprocess</td>
<td>12</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nkammern</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NdatumDerSerie</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NserienNamen</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Npictures</td>
<td>98</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nassignments</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NSamplePrepsPerDay3</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NrRepPerDay</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>NO</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

---

Table 1-1. Uebersicht Basisdaten:

<table>
<thead>
<tr>
<th>Name</th>
<th>Mean</th>
<th>Value</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nprocess</td>
<td>12</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nkammern</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NdatumDerSerie</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NserienNamen</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Npictures</td>
<td>36</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nassignments</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NSamplePrepsPerDay3</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NrRepPrepsPerDay3</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GlobnoDays</td>
<td>NO</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.5b: F and P values of combination of nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) and varieties (Rodelika and Rothild) of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design.
Table 4.5c: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design (1,2 are Rodelika variety).
Table 4.5d: F and P values of different in nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) of carrot from University of Kassel in 2004 with 1 day and 1 chamber experimental design (3,4 are Rothild variety).

<table>
<thead>
<tr>
<th>Refined statistical model with crossed factor</th>
<th>OLD statistical model without crossed factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Graph 1" /></td>
<td><img src="image2" alt="Graph 2" /></td>
</tr>
<tr>
<td><img src="image3" alt="Graph 3" /></td>
<td><img src="image4" alt="Graph 4" /></td>
</tr>
</tbody>
</table>
4.1.4 Summary of the optimization of the statistical model

The research experiments are mixed effects with repeated measurements that are incorporated with both fixed and random effects. Therefore, the mixed model should be accomplished in order to describe all effects which are involved in the experimental design. Furthermore, it is more possible to handle via many statistical programs such as R, S and SAS programs, etc.

The important results can be listed as follows:

1. Implement the mixed effect statistical model,
2. Verify ANOVA results of the statistical model from 2 statistical programs, i.e., SAS, R program in order to ensure that this model is correct, and
3. Run this model with ACIA program and elaborate until it is reasonably for use in the future work.

Their details are concluded as follows:

1. The mixed effect statistical model is completely carried out and implemented with crossed factor by following the methodology from Piepho et al. (2003). The full refined statistical model is

\[ \text{day} + \text{sample} + \text{chamber} + T1 + T2 + T3 + T4 + T5 \]

First order interactions are T1 and T2,
where T1 is day. sample.
T2 is day. chamber.

Higher order interactions are T3, T4 and T5,
where T3 is day. sample. sample preparation.
T4 is day. sample. sample preparation. chamber.
T5 is day. sample. sample preparation. chamber. replication.

2. The comparison of ANOVA results of R and SAS gives identical results. R program calculates a more conservative approach, so it is in the safe side. This result shows that the mixed statistical model is correct and so-called “refined statistical model”.

3. The refined statistical model was also executed via ACIA program. The comparison result shows that the refined statistical model produce better performance than the old one.
4.2 Investigate the effect of image parameters on texture analysis of biocrystallization image: ROI, color transformation and histogram matching.

4.2.1 Introduction

The sake of this section is to investigate the parameters that play an influence on the biocrystallization image analysis. These parameters are ROI (Region of Interest), color transformation and histogram matching.

ROI is the region to examine which zone of an image contains relevant information for differentiation purpose. It is initially enlarged from the geometric centre towards the periphery of the image at the percent value of total radius. Color transformation is to transform 3 dimensional information RGB (red, green and blue) to the desired 1 dimensional “color” which can be handled by the texture analysis. The last one is histogram matching which is defined as the method to enhance and normalize the image contrast by matching an original histogram of gray scale to another histogram distribution, e.g., Gaussian histogram, equalized histogram etc.

The biocrystallogram image is based on the crystallographic phenomenon that adding a specific substance to an aqueous solution of dehydrate CuCl$_2$, crystallograms with reproducible dendritic texture are formed during crystallization (Andersen et al., 1998). Therefore, one of the methods that can interpret and analyze this kind of image is the texture analysis. It extracts important textural information from an image and measure texture directly based on local spatial variation of intensity or color (Ojala et al., 2001; Wu et al., 2003; Coburn and Robert, 2004). The advantage of this technique does not provide only the information which represents the visual characteristics, but also is able to analyze the non-accessible information through optical view (Basset et al., 2000). This research will use the texture analysis to discriminate between biocrystallograms from conventional and organic samples.

The samples which are obtained are from BLE project 020E170 and 020E170/F (financed by Federal Ministry of Food, Agriculture and Consumer protection 020E170 and 020E170/F). These samples are from controlled field trials and farm pairs. The controlled field trials comprise of the wheat DOC in 2003 and 2005 as well as carrot from University of Kassel in 2004, 2005 and 2006. The wheat DOC samples are measured in all 4 field replicates. The F and P values are giving the information not only about the differentiation of samples, but also the farming systems. The carrot from University of Kassel are from the replicated field trial at the university of Kassel. For the measurement through the biocrystallization method, all 4 field replicates were pooled together to a bulk. Therefore, the F and P value are determined according to the differentiation of the samples merely. The samples are named in accordance to the cultivars tested: Rodelika and Rothild.

The farm pairs are the pairs of samples that compare the difference between organic and conventional cultivated on neighbor farms. They are called wheat market and carrot market in 2004 and 2005. The aim is to differentiate samples based on the information about the differentiation of single farm paired only, not as an overall discrimination of organic from conventional systems. The details of all samples are shown in appendix A. All images from each sample are scanned in 600 dpi 8 bits of RGB color in transmission of the Umax PowerLook III. A total of 15 texture parameters per image are then calculated and evaluated by the statistical model obtained from the 1st study by using.
ACIA program version 1.2.3. They are derived from the gray level co-occurrence matrix (GLCM at (0, 1) where 0 means direction of pixel pairs on a picture, this means direction \( \theta = 0^\circ \) and 1 refer to distance between the pixel pairs, i.e., one step in the horizontal direction) with the resolution scale 1. These 15 texture parameters are consisted of energy, entropy, maximum probability, correlation, diagonal moment, kappa, difference energy, difference entropy, inertia, inverse difference moment, sum energy, sum variance, sum entropy, cluster shade and cluster prominence.

The differentiation of each image parameter is analyzed based on ANOVA results, these best image parameters are chosen by the best overall of F and P value of 15 texture parameters (the color coding can be found in appendix D). By that, P value was plotted via log of P value, when the P value of texture parameters presents more significance than 0.0001 then the value will be dipped and shown underneath the line of -4.

![Biocrystallogram](a) GLCM (b) ANOVA result (c)

Figure 4.5: the methodology to image analysis. (a) The scanned biocrystallogram image. (b) The gray level occurrence matrix (GLCM) at 0, 1 for the texture analyzed calculation. (c) The ANOVA result by plotting F and P value of 15 texture parameters over the image parameters, e.g., ROI.
4.2.2 Effect of Region of Interest (ROI) on texture analysis

4.2.2.1 Introduction

Region of interest (ROI) is the region of the image specified for analysis. The circle-ROI, the main type of ROI, represents the radius of a circle around the geometric center. Herewith, 0% equals to the geometric center while 100% equals the whole image. Circle-80, therefore, is a circle of 80% radial extension around the geometric center (figure 4.6)

![Figure 4.6](image)

(a) (b) (c)

Figure 4.6: (a) the region of interest (ROI) of circle-50 (non blue area). (b) The region of interest of circle-80 and (c) circle-100.

As the number of pixels in that specified area of an image contribute to the texture analysis it will show which total area has the strongest effect. Hence, this part aims to investigate the main region of interest (ROI) that provides the main area on images and gives the strongest effect to differentiate sample in questions.

4.2.2.2 ROI investigation

7 circle-ROIs are used by varying from 40 to 100% of the total image area, i.e., circle-40, 50, 60, 70, 80, 90 and 100. In order to circumvent all effects that may be caused from the other parameters, thereby, color transformation and histogram matching are kept on the same value. The color transformation is equal color and the histogram matching is Gaussian distribution. The histogram of original image is fixed by matching to Gaussian distribution, as this kind of distribution can well perform for the analysis of standard textures, i.e., 15 Brodatz textures (Carstensen, 2002).
4.2.2.3 ROI result of wheat DOC

Wheat DOC has 5 samples that are different in the grouping system of controlled trial. Each sample is replicated in 4 field replication blocks. They were sampled in 2003 and 2005. The details of primary data are shown in appendix A.1. The ROI results can be mainly considered in 2 parts, i.e., firstly, the overall assignments is checked to unveil for the overview significance of wheat DOC (figure 4.7a and b) and secondly, each pair of assignments is also checked (appendix H.1.1 and H.1.2).

Their checking results show that there is the monotonic pattern of F and P value over different ROIs, but this pattern gives the opposite movement between 2003 and 2005. The different pattern can be described that when the ROI is increased, F value in 2003 goes down and reaches eventually the lowest at circle-100, meanwhile F value in 2005 tends to move up with enlarging ROIs. This pattern also conforms to P value that appears to give low significance in 2003, but it does not in 2005. The unconformity of F and P value of samples in 2003 and 2005 may be reasoned from the different climates in each year. Zörb et al. (2006) explained that the variation in the growing season affects to single compound in wheat DOC, especially, valine and alanine amino acid. From this consequence, it might influence the difference in image pattern that was generated from various samples in different years.

Even the results are unconformity in different years, but the most effective ROI from both years can be viewed at the middle area of images, i.e., between circle-40 to -80 and circle-80 till -100 in 2003 and 2005, respectively. In order to select the best representative ROI that is optimized for the differentiation of wheat DOC in these 2 years and mitigate the effects which may be consequent from specified different ROIs in each year. For an example, if ROI is different, the number of pixel in that specified area of the picture would be unequal and it then contributes to be an erroneous diagnosis result. Furthermore, a region is perceived to have texture when the number of primitive objects in the region is large. If only a few primitive objects are present, then a group of countable objects is perceived instead of textured image (Tuceryan and Jain, 1998). Hence, the best ROI of wheat DOC finally appears at circle-80 (figure 4.7 and 4.8), as it provides the highest numbers of pairs of assignments that play significant role, i.e., 9 pairs in 2003 and 5 pairs in 2005 (table 4.6, appendix H.1.1 and H.1.2). Anyhow, this research result is related to the concentration at 70/90 because of to study the effect of image parameters in the same level. But with the concentration at 90/90 in 2005 the pictures shows a better differentiation.
Figure 4.7: (a) F and P value of wheat DOC for overall assignments in 2003. (b) F and P value of wheat DOC for overall assignments in 2005. They are plotted on the Y axis versus the ROI on the X axis. Starting on the left with circle-40 and ending on circle-100.
Figure 4.8: some of texture parameter over ROI. (a) Average diagonal moment value in 2003 and 2005. (b) Average energy value in 2003 and 2005: control-(A), mineral-(B), bio-dynamic-(C), organic-(D) and conventional sample(E) in 2003, control-(K), mineral-(H), bio-dynamic-(G), organic-(J) and conventional sample(F) in 2005.
Table 4.6: The significance for paired of wheat DOC assignments in 2003 and 2005.

<table>
<thead>
<tr>
<th>Paired assignment</th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control~mineral</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Control~bio-dyn</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Control~organic</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Control~conventional</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Mineral~bio-dyn</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Mineral~organic</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Mineral~conventional</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Bio-dyn~organic</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Bio-dyn~conventional</td>
<td>_</td>
<td>/</td>
</tr>
<tr>
<td>Organic~conventional</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9 significant pairs</strong></td>
<td><strong>5 significant pairs</strong></td>
</tr>
</tbody>
</table>

“/” presents the significant pairs of assignment (P ≤ 0.05)
“–” presents the non significant pairs of assignment (P>0.05)

Figure 4.9: The best ROI at circle-80 of wheat DOC (non blue area).
4.2.2.4 ROI result of wheat market

The 6 pairs of wheat market in 2004 and 2005 were organized from the different organic and conventional farm pairs. Each assigned pair is obtained from the same variety, but different in adjacent cultivated lands. The details of data are shown in appendix A.2.

The sake of the wheat market experiment is to compare organic and conventional samples that are obtained from the same variety but different farms. However, this experimental design still has the other objectives that are (a) whether it can discriminate different varieties of wheat or not and (b) whether it can discriminate general organic and conventional wheat or not. Therefore, this research part is studied and summarized in 3 parts, i.e., the result of ROI for (a) overall assignments is observed for an overview differentiation as well as the comparison of each pair of assignments, (b) the different varieties and (c) the comparison between general organic and conventional wheat market.

The first part of ROI result is considered via the overall and each pair wise assignments in both 2004 and 2005. The results of these years provide the monotonic phenomenon of texture parameters. That is, when the ROIs are even enlarged, the F value is even rose up, attains the peak value at circle-70, after that recedes down from circle-80 till -100. It is conformed to P value that also tends to give the highest significance at circle-70 (figure 4.10a and b). The best ROI result can be viewed in the middle area of images at circle-70, because it does not provide only the strongest effect of F and P value, but reveals also the majority significant number of pairs of assignments. By that, 4 significant pairs are from total 6 pairs in 2004, but non significance in 2005. However, the F and P value results in 2005 still give the strongest trend at circle-70 (appendix H.2.1 and H.2.2).
Figure 4.10: (a) F and P value of wheat market for overall assignments in 2004. (b) F and P value of wheat market for overall assignments in 2005. It is plotted on the Y axis versus the ROI on the X axis. Starting on the left with circle-40 and ending on circle-100.
The second part is the differentiation of general organic and conventional wheat market (figure 4.11 a and b), it can be illustrated that the F and P value have a uniform stronger effect when ROIs are increased from circle-40 to -80 and their recessions are in the extents of circle-90 and -100 in both years. Herewith, the primary optimum ROIs to discriminate organic and conventional wheat market should be viewed in between circle-40 till -80.

Figure 4.11: (a) F and P value of wheat market for the differentiation of general organic and conventional sample in 2004. (b) F and P value of wheat market for the differentiation of general organic and conventional sample in 2005.
Concerning to the 3rd section of the differentiation of wheat varieties (figure 4.12a and b), the varieties consisted of Capo, Bussard, Asketis in 2003, Ludwig, Bussard and Capo in 2005. The overall of their results provides a similar pattern and conforms to the results of organic versus conventional comparison as mentioned above. That is the F and P values are stronger when the ROI are changed from circle-40 to -80, and then appears to give lower effects at circle-90 and -100. The optimum ROI of different varieties can be also viewed in between circle- 70 to - 80.

Figure 4.12: (a) F and P value of wheat market for differentiation of varieties in 2004. (b) F and P value of wheat market for differentiation of varieties in 2005.
As above all results, the optimum ROI of wheat market manifestly appears in the middle area of image as a representative region at circle-70, because it presents the strongest effect of F and P value, as well as provides the highest number pairs of assignments which are significant as mentioned earlier.

Figure 4.13: the best ROI at circle-70 of wheat market (non blue area).

4.2.2.5 ROI result of carrot from University of Kassel

Carrot from University of Kassel are samples obtained from controlled field trials. They were sampled in 2004, 2005 and 2006 in order to study the effect of different nitrate fertilizer levels and varieties. The detail of samples is in appendix A.3. The main ROI should be determined for the differentiation in prime 3 parts, i.e., varieties, nitrate fertilizer levels and combination of these 2 treatments.

The first results are to observe the ROI effect of different varieties and the combination of nitrate levels and varieties. In 2004 and 2006, the ROI effect shows that their F and P value movement gives the same pattern of texture parameters over various ROIs. That is when the ROIs are bigger, consequently, F value rises up and, eventually, achieves the peak in between circle-60 and -80. Aftermath, its recession limb locates from circle-90 to 100. The F pattern conforms to P value that provides higher significance when the ROI are larger, and then P value shows non-significance at circle-90 and -100 (figure 4.14a and c; appendix H.3.1 and H.3.5).

The ROI results in 2005 provide a similar behavior, but they presents the contradict pattern with ROIs from 2004 and 2006 that is the F and P value give stronger effect when enlarge ROI from circle-40 to -60. The next, they are suddenly drop and shows non-significance at circle-70. Aftermath these values turn out stronger effect at circle-80 to -90 again and suddenly move down at circle100 (figure4.14b; appendix H.3.3 and H.3.4). This unconformity phenomenon of F and P value over 3 years can be described via figure 4.15. The figure shows that all texture parameters, such as, energy and diagonal moment values are equal at circle-70, while the other ROIs do not. Its main effect may be arose from the different climates in each year, as ROIs in 2004 and 2006 still depict the same ROI effect, but they do not show the similarity with ROIs in 2005. From this consequence, the optimal ROI which can discriminate varieties, the combination of nitrate fertilizer levels and varieties should be used at circle-80, as it provide the optimal F and P value for the differentiation.
Figure 4.14: (a) F and P value of different varieties of carrot from University of Kassel in 2004. (b) F and P value of different varieties of carrot from University of Kassel in 2005. (c) F and P value of different varieties of carrot from University of Kassel in 2006.
Figure 4.15: (a) The texture parameters value (diagonal moment and energy value) versus ROI changes of different varieties in 2004. (b) The texture parameters value (diagonal moment and energy value) versus ROI changes of different varieties in 2005. (c) The texture parameters value (energy and inertia value) versus ROI changes of different varieties in 2006.
The last ROI results concern on the comparison of different nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)). They reveal the opposite results in different years too. In 2004, the F and P value of texture parameters from Rodelika variety give strongest effect at circle-70, and they then are steadily decreased until the circle-100 (figure 4.16a). Meanwhile Rothild variety does not show any differentiation. This result conforms to the result from 2006 that the differentiation over ROIs of sample A also does not express any differentiation (figure 4.16c). The F and P value of texture parameters from Rodelika and Rothild varieties provide different patterns of graphs in 2004, 2005 and 2006 (figure 4.16b). This phenomenon may be affected from the different climates in each year.

Considering on all ROI effect, it is found that there is a similar pattern of variety and the combination of nitrate fertilizer levels plus varieties (figure 4.14 and appendix H.3). Meanwhile it does not provide strong effect in different nitrate levels (figure 4.16). Herewith, it can be deduced that the ROI effect can differentiate with different varieties more effective than nitrate levels.

From all earlier mentioned results, there are uncertainties of region that can differentiate carrot from University of Kassel samples over 3 years. However, it can be seen that the best ROI should locate in the middle area of images at circle-80, because the optimal strong effect of F and P value present there. Moreover, it can efficiently determine the difference all samples over 3 years.
Figure 4.16: (a) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2004. (b) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2005. (c) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2006.
4.2.2.6 ROI result of carrot market

Five pairs of carrot market are studied to discriminate organic from conventional farm pairs. The experiment was designed like wheat market. That is each assigned pair is obtained from the same variety, but different cultivated lands in both 2004 and 2005. The detail of data shows in appendix A.4. From this experiment, this part is oriented to (a) compare organic and conventional samples that come from the same variety, but different farm cultivation, so called “pair of assignments”, (b) investigate whether ROI effect can discriminate different varieties of carrot and (c) examine whether ROI effect can differentiate general organic from conventional carrot.

The first part of ROI result should be considered on the overall and pairs of assignments comparison in order to view the ROI effect that takes place with carrot samples both in 2004 and 2005. The ROI results of these years provide a similar pattern, i.e., the F value of texture parameters tends to be even higher when the ROI are steadily increased from circle-40 to -70. After that, F value trends to slope down from circle-80 till -100. The P value tends also to be the highest significance at circle-70 (figure 4.18). The best ROI result can be viewed at circle-70, as it provides the strongest effect of F and P value and gives the major number of pairs of assignments that significance, i.e., 4 pairs and 3 pairs from total 5 pairs in 2004 and 2005, respectively (appendix H.4).
Figure 4.18: (a) F and P value of carrot market for overall assignments in 2004. (b) F and P value of carrot market for overall assignments in 2005. It is plotted on the Y axis versus the ROI on the X axis. Starting on the left with circle-40 and ending on circle-100.
For the differentiation of varieties and general organic versus conventional carrot comparison, the ROI results disclose a similar behavior, i.e., F and P value of texture parameters are stronger effect when ROIs are increased from circle-40 to -70 and then they are starting to decline from circle-80 till -100 (figure 4.19 and 4.20). So the optimal ROI for both comparison should be viewed from circle-60 till -70.

Figure 4.19: (a) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2004. (b) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2005.
Figure 4.20: (a) F and P value of carrot market sample for differentiation of varieties in 2004. (b) F and P value of carrot market sample for differentiation of varieties in 2005.
Summary, all the ROI patterns from carrot market are giving the same behavior, i.e., once the ROI is even higher, the F and P value tend to be also higher effect, and they then likely go down from circle-80 till -100. Hence, the optimal ROI which is able to discriminate carrot market should be finally selected in the middle area of images at circle-70.

Figure 4.21: the best ROI at circle-70 of carrot market (non blue area).
### 4.2.2.7 Summary of the effect of ROI on texture analysis

Region of Interest (ROI) is the region of an image that is specified in order to examine which zone of an image contains in total the relevant information for differentiation purpose. The main type of ROI, circle-ROI, represents the radius of a circle around the geometric center. Therefore, a goal of this part is to determine the main differentiation region that takes place on the image. All results can be deduced as follow:

1. The ROI results show an important issue, i.e., that the changing of texture parameters by plotting F and P value is smooth and has the same movement over various ROIs. The differentiation of samples will be even better when the regions of image are enlarged, because a region is perceived to have texture when the number of primitive objects in the region is large enough. If only a few primitive objects present, then a group of countable objects is perceived instead of textured image (Tuceryan and Jain, 1998).

2. The ROI patterns of market samples provide nearly identical phenomena. That is once ROI is even higher, then the F and P value of texture parameters illustrate monotonic, tend to be even stronger and reach the strongest effect at circle-70. By that, the ROI can differentiate both in different varieties and differentiating general organic from conventional samples.

3. The best ROI of wheat DOC is circle-80.

4. The best ROI of carrot from University of Kassel is circle-80. The main effect of ROI is able to differentiate varieties more efficient than nitrate fertilizer levels.
4.2.3 Effect of color transformation on texture analysis

4.2.3.1 Introduction

Color is the range of human visual sensations that can be generated by mixtures of many wavelengths of visible light. It does not only play an important role in computer, television monitors display, but also is a basic feature that can be used for giving various information in an image. Generally, image colors are typically displayed as additive color composites by using the three primary colors, i.e., red, green and blue. But, in truth, the absolute color information may be not useful for the image analysis at all. So the alternative to describe the image colors by their RGB components should be derived. There are several alternative color models that are closely related to each other on the concept of tint, shade and tone, that are LCH (luminance-chroma-hue), HIS (hue-intensity-saturate) and HLS (hue-luminance-saturation) etc. The luminance or intensity refers to the total brightness of a color. Hue refers to the dominant or wavelength of light contributing to a color. And saturation or chroma presents the degree of purity of a color is its relative to a gray scale.

![RGB cube](image1.png) ![HIS hexcone model](image2.png)

Figure 4.22: (a) RGB cube and (b) HIS hexcone model (Mather, 2004; Pp 122).

The goal of this part is to study the effect of color transformation that may perform a central role on the differentiation samples in questions. The image is scanned based on three primary colors (RGB color), but the texture analysis is calculated based on gray levels. By that, the hypotheses of this section are (a) whether the effect of elements colors, for instance intensity and color wavelength, contribute to the image analysis and (b) whether the transformation of color can effectively differentiate samples and the extraordinary sample treatments such as different nitrate fertilizer levels in carrot from University of Kassel.
4.2.3.2 Color transformation investigation

The 8 colors transformation which are focused on contains of equal, red, green, blue, default, luminance, chroma and hue colors transformation. The original image histogram matching is specified as Gaussian distribution, while the Region of Interest (ROI) of each sample is fixed by following the study result from 4.2.2. By that, ROI of wheat DOC and carrot from University of Kassel are set at circle-80, meanwhile ROI of wheat and carrot market are specified at circle-70. The best color transformation was selected through plotting the F and P value versus color changes. The most efficient color is selected through concerning on the best overall F and P value of texture parameters.

The graphs for F and P value are plotted versus the “colors” change. By that, the color transformation moves from the left toward right as follows equal, red, green, blue, default, luminance, hue and chroma, respectively. The names of color with number which are along the X-axis mean those color names with the best ROI, for example, “equal 70” is equal color at circle-70 calculation.

1. Equal color is a color which derived from the sum red, green and blue with equal coefficient. The formula is \( E = 0.33r + 0.33g + 0.33b \).

2. Red, green, blue (RGB) are color only.

3. Default is a gray color which is converted from red, green and blue. The quantization formula is \( 0.299r + 0.587g + 0.114b \). It is according to the available receptors in the human eye.

4. Luminance, chroma and hue (LCH) are converted from red, green and blue color space. The luminance presents the brightness that relies on the gray color scale. Chroma presents the degree of purity of color on an image. Hue refers to the dominant of light that contributing to the color on an image.

4.2.3.3 Color transformation result of wheat DOC

Wheat DOC is sampled in 2003 and 2005 in order to investigate the difference between organic and conventional samples from controlled field trials. The color transformation is studied by specifying the ROI at circle-80 and histogram matching to Gaussian distribution. The results can be considered in 2 key parts. The first is the differentiation for the overall of wheat DOC assignments in order to consider the overview of color transformation effect on wheat DOC (figure 4.23). And the second is the differentiation of each pair of wheat DOC assignments (appendix I.1) in order to consider in the details of each pair comparison.

The two main results in each year produce monotonic behavior of F and P value over different colors. But these patterns give discordant movements when they are compared between 2003 and 2005. These contradicting patterns show that the major movements of equal, red, green, blue, default and luminance colors are at the same level of strong effect
and show highly significant to differentiate all wheat DOC in 2003. The next, the F and P value movement tend to decrease at hue and chroma color and do not provide any effect to differentiate images. Meanwhile the color effect in 2005 presents an opposite pattern, i.e., the F and P value give an equal effect through the colors that vary from equal to luminance and chroma colors, but they do not illustrate the strongest effect like 2003. On the other hand, the strongest color effect suddenly shows at merely hue color in 2005. This different behavior of color transformation effect in both years may be caused from the different climates as mentioned in the ROI effect.

Figure 4.23: (a) F and P value of wheat DOC for the overall assignment in 2003. (b) F and P value of wheat DOC for the overall assignment in 2005. It is plotted on the Y axis versus the color transformation changes on the X axis.
Table 4.7: The significance for paired of wheat DOC assignments in 2003 and 2005.

<table>
<thead>
<tr>
<th>Paired assignment</th>
<th>2003 with equal color</th>
<th>2005 with equal color</th>
<th>2005 with hue color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control~mineral</td>
<td>/</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Control~bio-dyn</td>
<td>/</td>
<td>-</td>
<td>/</td>
</tr>
<tr>
<td>Control~organic</td>
<td>/</td>
<td>-</td>
<td>/</td>
</tr>
<tr>
<td>Control~conventional</td>
<td>/</td>
<td>/</td>
<td>-</td>
</tr>
<tr>
<td>Mineral~bio-dyn</td>
<td>/</td>
<td>-</td>
<td>/</td>
</tr>
<tr>
<td>Mineral~organic</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Mineral~conventional</td>
<td>/</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bio-dyn~organic</td>
<td>/</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bio-dyn~conventional</td>
<td>_</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Organic~conventional</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Total</td>
<td>9 significant pairs</td>
<td>5 significant pairs</td>
<td>6 significant pairs</td>
</tr>
</tbody>
</table>

“/” presents the significant pairs of assignment (P ≤ 0.05)
“–” presents the non significant pairs of assignment (P > 0.05)

Although the color transformation effect in both years of wheat DOC gives the different behaviors. But their main results are the purity and the original color (chroma and hue color) do not contribute to wheat DOC image analysis, as they do not show strong effect. In fact, the main effective color transformation is only occupied on the gray color scale. Even though the hue color presents distinguish behavior in 2005, but the numbers of significance for paired assignments do not show differentiation more efficient than equal color (table 4.7).
From above results, it can be interpreted that the color effect of wheat DOC is not from the dominant effect of 3 primary colors but primarily depends on the gray color. Even the hue color unveils the strong effect in 2005, but the numbers of paired assignment which are significance are not higher than equal color. At the same time, an equal color shows more stable movement(table4.7).

Summarily, the best representative of color transformation which is effective to differentiate wheat DOC in these 2 years should be carried on equal color because it shows more stable effect than the other colors and remain also the strongest effect for the differentiation purpose.

4.2.3.4 Color transformation result of wheat market

Six paired assignments of wheat market are sampled in 2004 and 2005. The details of these samples can be found in appendix A. The color transformation is investigated by setting the ROI at circle-70 and original image histogram matching to Gaussian distribution. Owing to these 6 paired assignments are organized from organic and conventional farm pairs, so the differentiation is not only the comparison of organic and conventional from the same variety, but also focuses on the different wheat varieties and general organic versus conventional wheat market. This part is mainly divided into 3 parts as follows (a) determine the color overview of wheat market so that their overall color transformation effect and each farm pair can be visual and differentiated between the conventional and organic samples from the same variety, (b) consider the color transformation effect on the different varieties and (c) concern on the general difference of organic and conventional wheat market.

The first summary of color transformation is the comparison of organic and conventional farm pairs from the same variety. The results are investigated by considering on the overall (figure 4.24) and each pair of wheat market assignments in both 2004 and 2005 (appendix I.2).
Figure 4.24: (a) F and P value of wheat market for the overall assignment in 2004. (b) F and P value of wheat market for the overall assignment in 2005. It is plotted on the Y axis versus the color transformation changes on the X axis.
Their results disclose that the color transformation effect in these 2 years expresses an identical pattern of texture parameters. This phenomenon can be illustrated that when the color changes from the equal to luminance colors, the F and P value of these 2 years reveal the strongest effect and produce the same effect. At the same time, F and P value do not show any effect when the original color is transformed to hue and chroma colors. These results conform to the wheat DOC sample results that the main effect of color transformation does not reason from the purity of color and the dominant wavelength of light. But, in fact, it is from the gray color as mentioned earlier. From this consequence, it can be manifested the optimal color transformation which can discriminate overall and each farm pair of wheat markets is equal color, since it provides more stable movement than the other colors and shows the strongest performance to differentiate all wheat farm pairs in both years.

The second summary part is focused on the general organic and conventional comparison (figure 4.25a and b). The plots can be illustrated that the F and P value in these 2 years result the same phenomenon, i.e., they present an uniform stronger effect when the color changes from the equal to luminance color and they tend to be weaker at hue and chroma color. This phenomenon conforms to the overall and each farm paired assignments as above mentioned. Hence, the main effect of color transformation which is applied to differentiate organic and conventional wheat does not arise from the original color and their purity of color. But, in truth, the main effect relies only on the gray color.

From these above results, equal color should be selected as a representative optimal color to differentiate general organic and conventional wheat market.
Figure 4.25: (a) F and P value of wheat market for differentiation of organic and conventional comparison in 2004. (b) F and P value of wheat market for differentiation of organic and conventional comparison in 2005.
Figure 4.26a: F and P value of wheat market for differentiation of varieties samples in 2004; Capo (ass 1), Bussard (ass 2) and Asketis (ass 3) in 2004.
Figure 4.26b: F and P value of wheat market for differentiation of varieties samples in 2005; Ludwig (ass 1), Bussard (ass 5) and Capo (ass 7).
The final section of color transformation effect on wheat market is to differentiate varieties (figure 4.26a and b), their results seem to be no pattern and varied on the different color transformation behaviors. That means some pair of different varieties show stronger effect at hue and chroma colors (Capo and Asketis in 2004 as well as Ludwig and Capo in 2005), meanwhile the others do not. By that, it can point that the color transformation effect is unstable to differentiate the varieties. However, the main results of color effect in these 2 years still conform to the overall assignments, each farm pair of assignments and organic versus conventional comparison.

As above all wheat market results, they can be finally manifested that (a) the optimum color transformation that can differentiate wheat market should be run through an equal color and (b) the color transformation is able to differentiate general organic from conventional wheat market more effective than the differentiation in varieties.

4.2.3.5 Color transformation result of carrot from University of Kassel

Carrot from University of Kassel is determined in 2004, 2005 and 2006 to focus on the effect of different nitrate fertilizer levels and varieties. The color transformation study is investigated by setting the ROI at circle-80 and original histogram matching to Gaussian distribution. The main 3 results are the difference in varieties, nitrate fertilizer levels and the combination of these 2 treatments.

The first results of color transformation describe both of the different varieties (figure 4.27) and the combination of nitrate fertilizer levels versus varieties (appendix I.3).
Figure 4.27: (a) F and P value of different varieties of carrot from University of Kassel in 2004. (b) F and P value of different varieties of carrot from University of Kassel in 2005. (c) F and P value of different varieties of carrot from University of Kassel in 2006.
All 3 years results found that the F and P value provide an identical effect when the color transforms to equal, red, green, blue, default, luminance and chroma color, whilst F and P value are then sharply decreased at hue color. This phenomenon exposes that the main effective color transformation which can discriminate carrot varieties (Rothild and Rodelika) and the combination of nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) against varieties are on 2 dimensions of color space, i.e., gray color and the purity of its color on image(chroma color). It may be depicted that, actually, hue and chroma colors are always relevance to each other, but hue color has a high degree of variation (Lambert and Carron, 1999). In case of carrot sample, if hue color is less involved with chroma and then it does not show stronger effect like chroma color. Furthermore, one advantage of chroma color is contained of some information on image that does not present merely on the gray color (Shyu and Leou, 1998).

The last result is regarded to the comparison of the different nitrate fertilizer levels (0 and 150 kg N ha\(^{-1}\)) over 3 years. It can be seen in both the same and different color transformation patterns(figure 4.28). The identical color behavior shows that F and P value of texture parameters are same effect with different colors, i.e., equal, red, green, blue, default and luminance. Whereas the different behavior exists at hue and chroma color movement, i.e., the F and P value give strong effect at only hue color of 2004 (Rothild variety) and some in 2006 (variety of sample A), but there is no any effect and non significance at both of hue and chroma color in 2005 (Rodelika variety) and 2006 (variety of sample B).This kind of pattern may be reasoned from either the different climates in these 2 years or the truly effect of hue color. Therefore, it should be deeply focused and studied more than 3 years in order to grasp the real effect of nitrate fertilizer levels. As above discussions, the equal color is most appropriate to discriminate nitrate fertilizer levels.
Figure 4.28a: F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2004.

Rodelika in 2004

Rothild in 2004
Figure 4.28b: F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2005.
Figure 4.28c: F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2006.

The color transformation results of carrot from University of Kassel can be finally summarized that (a) the main color transformation effect is 2 dimensions of color space, i.e., gray and chroma color, (b) The color transformation effect can discriminate different varieties more effective than nitrate fertilizer levels, as the color pattern presents more consistent and (c) the optimal color transformation which can differentiate all carrot from University of Kassel over 3 years should be finally selected at equal color.
4.2.3.6 Color transformation result of carrot market

Five pairs of carrot market which comprised of organic and conventional farm pairs were investigated in 2004 and 2005, their details are expressed in appendix A. The color transformation was observed by specifying the ROI at circle-70 and original histogram matching to Gaussian distribution. The results of carrot market consisted of 3 main parts, i.e., firstly, consider of overall and each farm pair of carrot market assignments obtained from the same variety but different cultivated land. Secondly, consider the comparison of general organic and conventional carrot market and the last consideration is the differentiation of carrot varieties.

(a)

(b)

Figure 4.29: (a) F and P value of carrot market for overall assignments in 2004. (b) F and P value of carrot market for overall assignments in 2005. It is plotted on the Y axis versus the color transformation changes on the X axis.
The overview of carrot market plots unveils a monotonic pattern of texture parameters of the color transformation effect in these 2 years (figure 4.29). By that, the F and P values provide an equivalent strong effect with the varying from equal to luminance and chroma color. And then they are sharply declined at hue color. This result provides the conformity to the comparison of each farm pair of carrot market that a hue color shows non significance and does not show any effect on image analysis (appendix I.4). From this consequence, the main color transformation effect which can differentiate paired assignment of carrot market is relied on 2 dimensions of color space, i.e., gray and chroma.

The same pattern of the overall carrot and each paired assignment appear on the comparison of general organic and conventional carrot in these 2 years (figure 4.30). This can imply that there is no effect from a hue color, but the real effect is based on the gray and the purity of its color on image (chroma color).

On the other hand, when the different carrot varieties result is considered (figure 4.31), it shows that the color transformation effect is variant. By that, there are no effects of hue and chroma color in 2004, meanwhile the effect exists only on hue color in 2005.

From all carrot market results, the main color transformation effect which can feasibly differentiate all carrot farm pairs depends on the 2 dimensions of color space, i.e., gray and chroma colors. Except the differentiation of different carrot varieties that seems to be depended merely on a gray color. As mentioned above, the optimum color transformation effect of carrot market should be finally used an equal color.
Figure 4.30: (a) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2004. (b) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2005.
Figure 4.31: (a) F and P value of carrot market sample for differentiation of varieties in 2004. (b) F and P value of carrot market sample for differentiation of varieties in 2005.
4.2.3.7 Summary of effect of color transformation on texture analysis

Color is the range of human visual sensations that can be generated by mixtures of many wavelengths of visible light. It does not play only an important role in computer and television monitors display, but also is a basic feature that can be used for giving information in an image.

From this consequence, the sakes of this part, then, are (a) investigate the effect of color transformation that may be influenced on biocrystallogram image and afford the strongest effect to differentiate all sample in questions. And (b) investigate the transformation of color can enhance more effective to differentiate samples as well as the extraordinary sample treatments.

The investigated “colors” comprise of equal, red, green, blue, default, luminance, hue and chroma “colors”. All the results of wheat and carrot can be manifested as follows

1. The main effect of color transformation on image analysis is relied on the gray color, especially with wheat. At the same time, the color transformation effect reasonably relied on 2 dimensions of both gray and chroma colors with carrot.

2. The best color transformation which can efficiently differentiate all samples in questions is equal color as it provides the stronger effect and more stable movement than the others.

3. The original color or the wavelength of light (hue color) does not provide the strong effect to the biocrystallogram image analysis. Except with the extraordinary sample treatments such as different nitrate fertilizer levels in carrot from University of Kassel. The hue color present uncertainty effect so it can not be summarized the truly effect with merely 3 years data.

4 The result of wheat and carrot market that are focused to differentiate organic and conventional farm pairs, exposes that the pattern of color transformation gives better performance to the differentiation general organic from conventional samples than the difference in varieties. This result shows unconformity to the ROI effect which can provides the strongest effect both in different varieties and organic versus conventional comparisons.

5 The result of carrot from University of Kassel is studied to compare the influence of the cultivation on carrot quality. The results unveil that the color transformation provides stronger effect with the difference in varieties than the different levels of nitrate fertilizer.
4.2.4 Effect of histogram matching on texture analysis

4.2.4.1 Introduction

Image histogram is a fundamental step to provide basic information about the appearance of an image. It analyzes the distribution of pixel values in an image and expresses as a graph that consists of the number of times (vertical axis) of each gray level in the image (horizontal axis) (Patton et al., 2006). Although the image histogram can be assessed to express the tonal or radiometric quality of an image. But, actually, some image shows poor contrast since its histogram utilizes a restricted range of brightness value, or saturation in the black or white regions. Thereby, the image enhancement should be manipulated to improve the contrast of an image. It can be done through matching or transforming the original histogram to the desired matching.

There are alternative histograms matching which are increasing the sensitivity on an image used to improve the contrast of image. The most applications are histogram equalization and Gaussian matching (Carstensen, 1992). The equalized histogram principle is to make the histogram as uniform as possible, meanwhile Gaussian is to match the histogram to Gaussian shape, i.e., it contrasts a histogram with a few black and white area.

The objective of this section is to determine the impact of histogram matching on the differentiation of all samples. Because the original histogram of biocrystallograms from wheat and carrot in this research work are rather skewed in the transparent (white) region and they cannot provide the differentiation property well. So the sake of this part then is based on the assumption that when the original histogram is transformed to the other matching, it may clarify the image more acceptable for viewing, processing and analysis.

4.2.4.2 Histogram matching investigation

Five histograms matching types are used, i.e., histogram equalization, Gaussian, triangle, parabolic and hyperbolic matching. In order to focus only the histogram effect, so the best color transformation is set to equal color for all samples (the study result from 4.2.3). The best Region of Interest (ROI) is fixed by following the study result from 4.2.2. Hereby, wheat DOC and carrot from University of Kassel are set at circle-80 whilst wheat and carrot market at circle-70. The most efficiency effect of histogram matching must provide the powerful effect to differentiate all samples and acceptable for viewing, analysis as well. It is culled by the best overall F and P values plotting versus histogram changes. By that, the histogram matching are plotted by starting on the left toward to the right with equalized histogram, triangle, Gaussian, parabolic and hyperbolic matching, respectively. The number which was followed the histogram matching name is the best ROI that was selected from 4.2.1, e.g., equal80 means matching the original histogram to an equalized histogram at circle-80.
1. Histogram equalization is a matching approach to make the histogram as uniform as possible (figure 4.32a).

2. Gaussian matching is a matching histogram to Gaussian shape, i.e., most of the image detail is located in the middle zone of brightness and gives a few black and white regions (figure 4.32b).

3. Triangle, parabolic and hyperbolic matching are the mapping by using an interval of triangle, parabola and hyperbola shape (figure 4.32 c, d and e).

Figure 4.32: (a) histogram equalization. (b) Gaussian matching. (c) Triangle matching. (d) Parabolic matching and (e) hyperbolic matching.
4.2.4.3 Histogram matching result of wheat DOC

Wheat DOC is analyzed in 2003 and 2005 for the purpose of differentiation organic from conventional plant by controlled field trials. The histogram matching is followed by confining the ROI at circle-80 and color transformation at equal color. Their result can be mainly considered in 2 parts, i.e., (a) the overall assignment part(figure 4.33) and (b) each pair assignments (appendix J.1).

Figure 4.33: (a) F and P value of wheat DOC for the overall assignment in 2003. (b) F and P value of wheat DOC for the overall assignment in 2005. It is plotted on the Y axis versus the histogram matching changes on the X axis.
Their results of 2 parts in both years can be primarily seen that when the histogram matching changes, the F and P value pattern provide 2 groups of behaviors, i.e.,

(1) The first group is the texture parameters can be divided into 2 main sub-groups by concerning on their dissimilar movement pattern (figure 4.33). When the histogram matching is dissimilar, each histogram matching may lead to be different details of gray scale on images, and then the texture parameters which are calculated from each histogram matching also provide different values. The first group is shown as follows

(1.1) The 1\textsuperscript{st} sub-group is diagonal moment, cluster shade and cluster prominence which are measured for the difference of the variation of gray levels in an image.

(1.2) The 2\textsuperscript{nd} sub-group is energy, different energy, inverse different moment and the others. This sub-group measures mainly the uniformity distribution, the image contrast, the correlation of high and gray level on an image.

(2) The second group is divided the main histogram matching into 2 main sub-groups according to the conformity movement of F and P value (figure 4.34).

(2.1) The 1\textsuperscript{st} sub-group is equalized histogram and Gaussian matching as their performance is equivalent.

(2.2) The 2\textsuperscript{nd} sub-group consists of triangle, parabolic and hyperbolic matching. The fundamental idea why the group is consisted of 2 sub-groups are due to the first sub-group are uniform and Gaussian shape while the 2\textsuperscript{nd} sub-group skewed rightward.

Figure 4.34: box plot of diagonal moment versus different histogram matching of wheat DOC. The result implies that when the histograms matching are changed, then their original histogram in an image also is adapted. From this reason, they contribute to the calculation of texture parameters and their behavior, i.e., differing histogram then texture parameter values are also different.
According to the histogram matching result of wheat DOC, the first consideration is the overall assignment result in both years (figure 4.33). Their F and P value of 2 texture parameter groups are dominant movement and show distinguish effect at Gaussian matching. The second consideration is to view the histogram pattern on each pair of assignments (appendix J.1). The results can be summarized in table 4.8, i.e., the Gaussian and histogram equalization is able to effectively discriminate each pair of assignments, as their F and P value are stronger and provide a major number of significant pairs in both years. By that, the Gaussian matching and histogram equalization expresses, in turn, 9 and 8 significant pairs in 2003, meanwhile, 5 and 2 significant pairs from total 10 pairs in 2005. On the other hand, the triangle, parabolic and hyperbolic matching provide only 8 significant pairs in 2003 and no one is significant in 2005. From these above results, it may be concluded that the best histogram should be determined from Gaussian matching, as it performs the best efficiency to differentiate all wheat DOC and generates the highest significant pairs in these 2 years.

This may be explained that the original image histograms of wheat biocrystallogram, in fact, dwells in the white region of brightness (figure 4.35). This histogram shape is different from equalization and Gaussian matching, but gives a similar shape with the triangle, parabola and hyperbola matching. Consequently, when triangle, parabola and hyperbola are applied to discriminate the wheat image, they could not well clarify the difference of image contrast like Gaussian. Furthermore, the major area under Gaussian curve is positioned at the middle of gray scale, whilst its dark and white region are occupied at the small 2 areas of leftmost and rightmost part off major central area of Gaussian. Thereby a new histogram shape which is established by Gaussian can obviously enhance the contrast of the image and provide consequently more efficient effect to differentiate all wheat images. This result conforms to Carstenson (2002) who argued that the Gaussian matching provides better performance with the differentiation of the 15 standard Brodatz textures. Richard (1993) noted that the well contrast of an image will exist through the well spread out of histogram range and without significantly large bars at black or white area.
Table 4.8: The summary result of histogram matching for wheat DOC in 2003 and 2005.

<table>
<thead>
<tr>
<th>Pair of assignments</th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistical result</td>
<td>Energy value</td>
</tr>
<tr>
<td>control ~ mineral</td>
<td>/ all</td>
<td>all</td>
</tr>
<tr>
<td>Control ~ biodynamic</td>
<td>/ all</td>
<td>G, E</td>
</tr>
<tr>
<td>control ~ organic</td>
<td>/ all</td>
<td>G, E</td>
</tr>
<tr>
<td>control ~ conventional</td>
<td>/ all</td>
<td>G, E</td>
</tr>
<tr>
<td>mineral ~ biodynamic</td>
<td>/ T,P,H</td>
<td>all</td>
</tr>
<tr>
<td>mineral ~ organic</td>
<td>/ G,T,P,H</td>
<td>all</td>
</tr>
<tr>
<td>mineral ~ conventional</td>
<td>/ all</td>
<td>G, E</td>
</tr>
<tr>
<td>biodynamic ~ organic</td>
<td>/ G</td>
<td>-</td>
</tr>
<tr>
<td>biodynamic ~ conventional</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>organic ~ conventional</td>
<td>/ G,E</td>
<td>T,P, H</td>
</tr>
</tbody>
</table>

The “/” means statistical result express significance (P ≤ 0.05).
The “-” means statistical result express non significance (P ≥ 0.05).
The “all” means all of histogram matching are significant.
The alphabet “E”, “G”, “T”, “P”, “H” refer to the stronger behavior of equalized histogram, Gaussian, triangle, parabolic and hyperbolic, in turn.

Figure 4.35: (a) Biocrystallization image of wheat DOC. (b) The original histogram of wheat DOC. (c) The histogram of wheat DOC after match to Gaussian matching.
From results, it can be finally manifested that the best effective histogram matching should be the Gaussian matching as its F and P value provide stronger effect and perform better than the others matching by originate high amount of significant pairs, i.e., 9 from 10 pairs in 2003 and 5 pairs in 2005.

4.2.4.4 Histogram matching result of wheat market

Six pairs assignments of wheat market from 2004 and 2005 were used to investigate (a) the difference in organic and conventional farm pairs. (b) The difference in varieties and (c) the difference in general organic and conventional wheat market. The histogram matching was determined by setting the best ROI at circle-70 and color transformation to equal color (refer to the result in 4.2.2 and 4.2.3).

The first histogram matching result is to compare the difference between organic and conventional farm pairs through the overview of all assignments and each farm pair comparison. The F and P value from the overview assignment provide stronger effect both in histogram equalization and Gaussian than triangle, parabolic and hyperbolic matching (figure 4.36).

Figure 4.36: (a) F and P value of wheat market for the overall assignment in 2004. (b) F and P value of wheat market for the overall assignment in 2005. It is plotted on the Y axis versus the histogram matching on the X axis.
When it is assessed through the comparison of each pair assignments. Their all comparison results show non significance in 2005 (appendix J.1.2), but expresses the resemble effect of the result in 2004 to the overall assignment (appendix J.1.1), that is F and P value disclose strong effect via Gaussian matching, as it provides the major numbers of highest values of significant pair among tested matching, i.e., Gaussian results 4 highest significant pairs of total 6 examined pairs, meanwhile the others matching give only 3 significant pairs. These results can be clarified that the initial image histograms of wheat biocrystallogram, in truth, resides in the brighter area (its mean skewed toward right), while the major area under Gaussian curve is shaped as a symmetric bell (its mean positions at the center). By that, a new histogram which is implemented by Gaussian can obviously enhance the contrast of image and more effective differentiate all wheat images than the original histogram. This result conform to Richard (1993) and Carstensen (2002) as previously mentioned in wheat DOC result.

By that, it can be primarily deduced that the optimal histogram matching that is able to discriminate an overall and each pairs of wheat market should be Gaussian matching.

The second summary is emphasized on the comparison of general organic and conventional wheat (figure 4.37). The plots can be described that the F and P value express a similar manner and revealed identical significant effect to an image. However, this result unconformably turn out to the overall and each farm pair comparison results – the comparison between organic and conventional from different cultivated land but same varieties- that Gaussian matching plays an important role. It might be argued and deduced that there is less sensitive of histogram matching to differentiate the general organic and conventional wheat than the comparison organic and conventional samples from the same variety. Anyway, the plots are still show that Gaussian matching provides more efficiency to differentiate general organic and conventional wheat than the others.
Figure 4.37: (a) F and P value of wheat market for differentiation of organic and conventional comparison in 2004. (b) F and P value of wheat market for differentiation of organic and conventional comparison in 2005.

The last results are the different wheat varieties (figure 4.38a and b). Their results in both years found that the F and P value from Gaussian matching unveil the strongest pattern to discriminate wheat varieties in 2004 meanwhile the histogram matching pattern are uncertain in 2005. However, it can be seen that, most of the time, Gaussian matching exists the strongest and plays significant role to different varieties comparison.
Figure 4.38a: F and P value of wheat market for differentiation of varieties samples in 2004; Capo (ass 1), Bussard (ass 2) and Asketis (ass 3).
Figure 4.38b: F and P value of wheat market for differentiation of varieties samples in 2005; Ludwig (ass 1), Bussard (ass 5) and Capo (ass 7).
All wheat market results can be finally manifested that (a) the best effective histogram matching should be carried on by using Gaussian matching. This result also coincides to wheat DOC result, therefore, it might be definitely concluded that the optimum histogram matching to differentiate all wheat sample is Gaussian. (b) The best Gaussian matching produces stronger effect to differentiate organic and conventional wheat from same variety than general comparison. (c) The histogram matching plays more powerful to differentiate varieties than general organic and conventional comparison.

4.2.4.5 Histogram matching result of carrot from University of Kassel

Carrot from University of Kassel was investigated in 2004, 2005 and 2006 in order to focus on the effect of cultivated combination, i.e., nitrate fertilizer levels and varieties. This section, therefore, can be summarized in 3 main parts, i.e., the different varieties, nitrate fertilizer levels and the combination of these 2 treatments. Moreover, their results are also related to the visual evaluation criteria.

The first results are describing the differentiation in carrot varieties, i.e., Rothild and Rodelika (figure 4.39) and the combination of nitrate levels versus varieties (appendix J.3).
Figure 4.39: (a) F and P value of different varieties of carrot from University of Kassel in 2004. (b) F and P value of different varieties of carrot from University of Kassel in 2005. (c) F and P value of different varieties of carrot from University of Kassel in 2006.
Their result over 3 years show that F and P value for the histogram types give significant effect to discriminate both in different variety and the combination of treatments. Anyhow, if it is scrutinized through the F and P value behavior, they reveal the stronger effect and highly significance at Gaussian and histogram equalization, particular with diagonal moment value.

They may be reasoned that when the original histogram— an original histogram of Rodelika and its combination are nearly Gaussian matching whilst Rothild is skewness curve are changed to the new histogram shape, i.e., matching them to be uniform as much as possible or converting them to be Gaussian shape. Then the contrast of image is more increasing than the mapping to triangle, parabolic and hyperbolic matching. Furthermore, the reason why diagonal moment expresses the strongest effect of texture parameters as it measures the correlation in different high and low gray level on image. Therefore, when a pair of image is quite different via the visual evaluation criteria such as more substance spirals and side needle in Rodelika variety image, but less in Rothild (figure 4.40), this value consequently provides dominant effect than the others.

Figure 4.40: (a) the original histogram and biocrystallogram of Rodelika and (b) Rothild variety in 2004.
From these results, it can be fundamentally deduced that the optimum histogram matching to differentiate the difference in varieties and the combination of nitrate versus varieties in 3 years are both equalized histogram and Gaussian matching. But, in truth, the best histogram shape should be decided via Gaussian as it gives more powerful overall effect, i.e., F and P value are stronger, than another one.

The last result is to determine the different nitrate fertilizer levels (0 and 150 kg N ha$^{-1}$) in 3 years (figure 4.41). The results unveil that the histogram matching in 2003 and 2006 show non significant effect meanwhile a few significance levels in 2005. It might be discussed that histogram effect is uncertainty and has variable pattern to discriminate different nitrate levels. And it might be not stronger enough to show the different in nitrate levels. However, it can be seen that F and P value of Gaussian matching still provides strong behavior to differentiate this treatment via energy value with Rodelika variety, while Rothild through diagonal moment value. These texture parameters provide the valuable meaning which can be described that the energy value plays an important role with the pair of images which has the same visual evaluation criteria of more substance spirals or side needle, as appeared in Rodelika (figure 4.42a), because it measures the uniformity distribution on an image. If there are some pixel concentrate at the same value in an image then this value will be higher. Meanwhile, the diagonal moment value will be effective with a pair of biocrystallogram image that present the same visual evaluation criteria of less substance spirals or side needle, but dominant in clear stem and obviously different in black and white area as shown in Rodelika (figure 4.42b).
Figure 4.41: (a) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2004. (b) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2005. (c) F and P value of different nitrate fertilizer levels of carrot from University of Kassel in 2006.
Figure 4.42a: the biocrystallogram image and original histogram of different nitrate levels in Rodelika in 2004.
From all histogram matching results of carrot from University of Kassel, they can be finally summarized that (a) a strong histogram effect to discriminate all of carrot from University of Kassel is Gaussian matching, because it provides the strongest efficiency to separate all these samples. It discloses a more powerful effect through diagonal moment value when a pair of samples has different original histogram of each sample, i.e., one is nearly Gaussian and visual evaluation criteria presents more substance spirals or side needle on image, while another one is quite skewness, the visual evaluation criteria presents more clear stem or distinguish dark and white area on image. Inversely, diagonal moment also presents a strong effect with a pair of images which shows the same skewness original histogram, visual evaluation unveils less side needle, more clear stem and dominant black-white area in both paired images.
Furthermore, a Gaussian matching effect will be more powerful via an energy value when a compared original histogram image is nearly Gaussian and exposes more substance spirals and side needle on crystallogram. (b) The histogram matching expressed better performance through different varieties than the different in nitrate levels as the histogram behavior presents more certainty and strong effect.

### 4.2.4.6 Histogram matching result of carrot market

Six pairs of carrot market samples which consisted of different organic and conventional farm pairs were examined in 2004 and 2005. The histogram matching was investigated by fixing the best ROI and color transformation according to the result from 4.2.2 and 4.2.3, i.e., circle-70 and equal color, respectively. Their results can be considered in 3 main parts, i.e., the differentiation of each farm pair assignments, different varieties and general organic versus conventional carrot comparison.

The first result is considered the differentiation of each pair assignment which obtained from the same variety but differences in organic and conventional cultivated land. Its result can be considered via the overall assignment to view the histogram matching effect on carrot and then go through the differentiation within each farm pair of assignments.

The overall result reveals all of histogram matching plays significant effect, especially Gaussian matching shows stronger effect with both energy and diagonal moment value in 2004. Meanwhile it exposes strong effect with only diagonal moment in 2005 (figure 4.43). When it is finely examined to each farm pair of assignment (appendix J 4), they found that the high potential histogram matching should be determined from Gaussian and equalized histogram in both 2 years. Because their F and P value provide stronger effect and express higher numbers of significant pairs, i.e., 4 pairs in 2004, 3 pairs in 2005 for Gaussian matching and only 2 pairs for equalized histogram in 2005. From this consequence, it might be initially concluded that the optimum histogram matching should be Gaussian matching.

However, Gaussian matching exposes the mixture of significant texture parameters in which there are 3 significant pairs with energy value and a significant pair with diagonal moment in 2004, while 2 significant pairs with energy and one significant pair with diagonal moment in 2005. These mixture phenomena depend on the original histogram and crystallogram pattern of each assigned pair. It can be illustrated that an energy value plays a significant effect with a pair of assignment that visual evaluation criteria have different quernadeln, less side needle, small main stem character and their original histogram are somewhat located in white region (figure 4.44). Likely, if a paired image's both original histograms are nearly Gaussian matching with visual evaluation criteria that dominant substance spirals and/or dense radial (figure 4.45).

On the other hand, the diagonal moment will more or partially influence if a paired assignment of original histogram is rather skewed toward the white area and expresses visual evaluation criteria via more clear stem, quite different in black and white area (figure 4.46).
Figure 4.43: (a) F and P value of carrot market for overall assignments in 2004. (b) F and P value of carrot market for overall assignments in 2005. It is plotted on the Y axis versus the color transformation changes on the X axis. The color transformation is starting on the left to right with equal, red, green, blue, default, luminance, hue and chroma color, respectively.
Figure 4.44: (a) the crystallogram and original histogram of paired assignment, e.g., assignment 9 (P1) and assignment 10 (P2) of carrot market in 2004. The detail can be found in appendix A.4. (b) Their F and P value result.
Figure 4.45: (a) the crystallogram and original histogram of paired assignment, e.g., assignment 1(O1) and assignment 2(O2) of carrot market in 2004. The detail can be found in appendix A.4. (b) Their F and P value result.
Figure 4.46: (a) the crystallogram and original histogram of paired assignment, e.g., assignment 5 (Q1) and assignment 6 (Q2) of carrot market in 2004. The detail can be found in appendix A.4. (b) Their F and P value result.
From all mentions, it might be deduce that the optimum histogram matching to differentiate the difference in carrot farm pair is Gaussian matching as it express the highest amount of significant pair of assignments in both year, i.e., 4 from 6 pairs in 2004 and 3 pairs in 2005.

The second part is determined the comparison of different varieties both 2004 and 2005 (figure 4.47). They reveal that Gaussian matching provides stronger effect than the others by providing higher significant F and P value, especially via energy value. It might be illustrated that the distribution of Gaussian matching can enhance contrast of an image, by that a pair of original image is somewhat nearly Gaussian matching with larger dense radial, more side needle and smaller main stem size. Meanwhile, another original image is quite rightward skewed and expresses quernadein, less side needle and small main stem size (figure 4.48). Thereby, if it is matched histogram to Gaussian, and then the crystallogram can be shown more contrast in uniformity distribution.

Figure 4.47: (a) F and P value of carrot market sample for differentiation of varieties in 2004. (b) F and P value of carrot market sample for differentiation of varieties in 2005.
Figure 4.48: the original histogram and some of crystallogram from different varieties; upper is Nerac variety and Narbonne in the lower.

As stated result, the different varieties in these 2 years provide with Gaussian matching a certainty and strong effect of histogram matching. Their results express dissimilarly to general organic and conventional carrot comparison (figure 4.49). This different behavior can be shown that there are uncertainty and variable histogram pattern to discriminate general organic and conventional in both years, e.g., Gaussian matching express strong effect with diagonal moment in 2004, but with energy in 2005.
Figure 4.49: (a) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2004. (b) F and P value of carrot market sample for differentiation of organic and conventional comparison in 2005.

From all of results, it can be concluded that (a) Gaussian matching is an optimum histogram matching to differentiate all of carrot markets as it provides the strongest behavior and highest number of significant paired assignments, i.e., 4 pairs in 2004 and 3 from 6 pairs in 2005. (b) This Gaussian matching provides more efficiency to differentiate organic and conventional from the same variety than general organic and conventional comparison. (c) The histogram matching effect expresses stronger effect to discriminate different variety than organic and conventional comparison.
4.2.4.7 Summary of effect of histogram matching on texture analysis

Image histogram is a meaningful step to provide basic information of the appearance of an image. However, actually some image shows poor contrast since its histogram utilizes a restricted range of brightness value, or saturation in the black or white regions. Thereby, the image enhancement should be manipulated to develop and improve the contrast of an image. It can be done by matching or transform the original histogram to the desired distribution and make it more insensible to lighting conditions. Hence, the objective of this section is to determine the effect of histogram matching that unveil the strongest effect and more effectively improve the differentiation of all samples. The results can be deduced as follow.

1. The histogram matching can be divided into 2 groups according to their F and P values movement. The 1\textsuperscript{st} comprises of the histogram equalization and Gaussian matching. The 2\textsuperscript{nd} group comprises of triangle, parabolic and hyperbolic matching as their movement is equivalent.

2. The texture parameters can be classified in 2 main groups according to their inverse movement. The 1\textsuperscript{st} group is diagonal moment, cluster shade and cluster prominence. The 2\textsuperscript{nd} group is energy, sum energy, different energy and the others.

3. The histogram matching effect can be also divided into 2 groups according to the competence to enhance the contrast of paired original histogram. The 1\textsuperscript{st} is Gaussian and histogram equalization. They show more effectivity when the comparison of original histogram is either (a) these 2 samples are skewness in white region or (b) one is skewness shape and another one is nearly Gaussian shape. The 2\textsuperscript{nd} group consists of triangle, parabolic and hyperbolic matching. They perform well with the original histogram that is compared Gaussian and Gaussian matching. This behavior are found in corn DOC trial samples in 2003 and 2005 (appendix. G).

4. The best histogram which can discriminate all samples in questions is Gaussian matching as it provides a strongest effect and majority number of significant pairs of assignments.

5. When matching the histogram to Gaussian matching, the energy group can be efficiently augment the contrast of biocrystallogram images which the visual evaluation criteria are dominated by substance spirals, side needle, quernadeln and dense radial. Meanwhile the diagonal moment group shows a distinguish effect with more clear stem, less side needle, quite different in black and white region in crystallogram image.

6. The result of market samples unveil (a) the Gaussian matching histogram, perform with the comparison farm pairs better than the general organic and conventional comparison and (b) the histogram matching having a better efficiency to discriminate different varieties than general organic versus conventional comparison.
7. The results of carrot from University of Kassel which are investigated on the combination of nitrate and varieties, disclose that Gaussian matching performs with the discrimination varieties better than nitrate fertilizer levels.
4.3 The relation of texture parameters and visual evaluation

The main task of this section is to relate the texture parameters with the visual evaluation that have been developed by triangle researcher group (University of Kassel, Germany; Louis Bolk Institute (LBI), Netherlands and Biodynamic Research Association Denmark (BRAD), Denmark) in order to illustrate how the relation of the texture parameters and the visual characteristics are on an image. The images in this part are selected as a representative of the entire images which are generated in the laboratory.

This section relates the strongest texture parameters and visual evaluation criteria with carrot samples only. Because the carrot images present more distinguish differentiation than wheat pictures in this research (70/90 concentration).

4.3.1 Carrot

Carrot comprise of carrot from University of Kassel and carrot market. The best 3 criteria of texture analysis are selected. By that, histogram matching is Gaussian, color transform to equal color. Whereas ROI are circle-80 for carrot from University of Kassel and circle-70 for carrot market.

4.3.1.1 Carrot from University of Kassel

Carrot from University of Kassel in 2004, 2005 and 2006 are designed to study the effect of different nitrate fertilizer levels and varieties. Thus, the sample discrimination are different varieties, fertilizer levels and the combination of these 2 treatments comparisons.

The result of texture analysis shows that it can statistically differentiate carrot from University of Kassel, especially the combination of varieties and nitrate fertilizer levels. However, it unveils more significant effect to discriminate varieties than fertilizer levels (figure 4.55). The strongest texture parameters are cluster prominence, diagonal moment, difference energy and inertia values (Table 4.9).
Table 4.9: The strongest effect of texture parameters to differentiate carrot from University of Kassel in 2004, 2005 and 2006

<table>
<thead>
<tr>
<th>Treatment</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different varieties (Rothild and Rodelika)</td>
<td>Cluster prominence</td>
<td>Difference energy</td>
<td>inertia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different nitrate fertilizer levels</td>
<td>Rodelika</td>
<td>Cluster prominence</td>
<td></td>
</tr>
<tr>
<td>(0 and 150 kg N/ha)</td>
<td></td>
<td>and Diagonal moment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rothild</td>
<td>Diagonal moment</td>
<td></td>
</tr>
<tr>
<td>Combination of varieties and nitrate fertilizer</td>
<td>Cluster prominence</td>
<td>Cluster prominence</td>
<td>inertia</td>
</tr>
<tr>
<td>levels</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the part of the different varieties, i.e., Rodelika and Rothild, the most effective texture parameters are cluster prominence in 2004 (figure 4.50a), energy group in 2005 and 2006, i.e., different energy in 2005 and inertia in 2006 (figure 4.51a and 4.52a). The cluster prominence measures the peaks around the mean value, so an image will have a few variation in gray scales when the value is low (Rikxoort, 2004). Meanwhile, difference energy and inertia measures the uniform distribution and varying intensities off an image (Cooper, 2004; Newsam and Kamath, 2005). From this measurement definition, it can be found that the difference energy value of Rodelika images is higher than Rothild in 2005 (figure 4.51c) meanwhile inertia value is lower in 2006 (figure 4.52c). Because the number of pixels on Rodelika variety image might be concentrated in some position of the GLCM. Furthermore, the variation of intensity on an image is lower than Rothild variety. By that, it can be connected to the image which show more substance spirals, side needle and more dark areas. Contrarily, the inertia value gives higher and difference energy shows lower values with sample A, as its image shows an inverse visual characteristic, i.e., less substance spirals and side needle with more dense radial, quernadeln and dominant in black and white color (figure 4.51b and c). From this consequence, it might present more local variation or near randomness in GLCM, then, the inertia is higher and difference energy gives lower. The result may suppose that sample A in 2006 is Rothild variety while sample B is Rodelika.

In 2004, the cluster prominence presents strongest effect instead of different energy and inertia in 2005 and 2006, respectively. It may be reasoned from the different picture pattern, i.e., Rodelika variety in 2004 (figure 4.50b) shows less density and frequency of substance spirals than 2005 and 2006 (figure 4.51b and 4.52b). The different phenomenon might be taken place from the different climates in each year which contributes to the development of carrot growing.

The cluster prominence of Rodelika variety in 2004 reveals lower value than Rothild (figure 4.50a) as the image has more substance spirals and side needle then, the variation in gray scales of its is lower than Rothild image which has more dense radial, quernadeln, dominant in black and white color (figure 4.50b and c).
Figure 4.50: (a) F and P value of texture parameters versus ROI change of different varieties in 2004. (b) The biocrystallization image of different varieties. (c) Box plot of cluster prominence to differentiate the different varieties in 2004.
Figure 4.51: (a) F and P value of texture parameters versus ROI change of different varieties in 2005. (b) The biocrystallization image of different varieties. (c) Box plot of difference energy to differentiate the different varieties in 2005.
Figure 4.52: (a) F and P value of texture parameters versus ROI change of different varieties in 2006. (b) The biocrystallization image of different varieties. (c) Box plot of inertia to differentiate the different varieties in 2006.
The results of different fertilizer levels and the combination of varieties versus nitrate fertilizer levels reveal that the *cluster prominence* and *diagonal moment* values are truly showing strong behavior to differentiate all samples in 2004 and 2005. The *inertia* value is an additional value which can also differentiate the combination of varieties versus nitrate fertilizer levels in 2006.

These results may be concerned with the images which are rather different (a) more *clear stem, quernadeln*, less *side needle*, quite difference in black and white region in images (figure 4.54), (b) sometime, paired images have quite more *substance spirals* (figure 4.53), and (c) the comparison of paired images which one has more *substance spirals, side needle*, more dark areas and another image has more *clear stem, quernadeln*, different in black - white region (figure 4.55a). From these image characters, the *inertia, cluster prominence* and *diagonal moment* show strong effect as they measure the variation and correlation of gray scale on an image. By that, *inertia* measures the variation of intensities on images. The *cluster prominence* measures the peaks around the mean value as mentioned earlier. The *diagonal moment* measures the different in correlation for high and low gray levels, in this case the *diagonal moment* is positive so that means the dark areas on image are smooth meanwhile bright areas are rough (Carstensen, 2002).
Figure 4.53: (a) the biocrystallogram images, (b) F and P values of texture parameters versus ROI change, (c) box plot of cluster prominence and diagonal moment values of the different nitrate fertilizer levels of Rodelika variety.
Figure 4.54: (a) the biocrystallogram images, (b) F and P values of texture parameters versus ROI change, (c) box plot of cluster prominence and diagonal moment values of the different nitrate fertilizer levels of Rothild variety.
Rodelika             Rothild

Nitrate level at 0 kg N/ha           Nitrate level at 150 kg N/ha

Figure 4.55a: the biocrystallogram images of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha).

Figure 4.55b: F and P values of texture parameters versus ROI change of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer level at 0 kg N/ha. The other combination each year can be found in appendix H.3.
Figure 4.55c: box plot of *cluster prominence* of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha) in 2004.

Figure 4.55d: box plot of *cluster prominence* of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha) in 2005.

Figure 4.55e: box plot of *inertia* of the combination of varieties (Rodelika and Rothild) and nitrate fertilizer levels (0 and 150 kg N/ha) in 2006.
4.3.1.2 Carrot market

Carrot market both in 2004 and 2005 are investigated on the effect of different carrot farm pairs that focus on the same variety. The main discrimination is the different (a) farm pairs, (b) general organic and conventional samples and (3) varieties. The results and biocrystallogram of carrot market in 2004 show in table 4.10 and 4.11. Carrot market result in 2005 presents in table 4.12 and 4.13. The texture analysis can differentiate organic from conventional farm pairs as statistical significance. Hereby, 4 from 5 farm pairs are correctly differentiated (80%). Furthermore, it can be differentiated also the different varieties and general organic versus conventional samples.

Table 4.10: the strongest texture parameter to differentiate carrot market in 2004. The sample coding information can be found more in appendix A.4.

<table>
<thead>
<tr>
<th>Sample coding</th>
<th>Variety</th>
<th>Significant texture parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1 ~ O2</td>
<td>Nabonne</td>
<td><em>Sum entropy</em></td>
</tr>
<tr>
<td>P1 ~ P2</td>
<td>Nerac</td>
<td><em>Difference entropy</em></td>
</tr>
<tr>
<td>Q1 ~ Q2</td>
<td>Nerac</td>
<td><em>Diagonal moment</em></td>
</tr>
<tr>
<td>R1 ~ R2</td>
<td>Nerac</td>
<td><em>Sum entropy</em></td>
</tr>
<tr>
<td>S1 ~ S2</td>
<td>Nabonne</td>
<td>Non significance</td>
</tr>
<tr>
<td>General organic and conventional comparison</td>
<td>-</td>
<td><em>Cluster prominence</em></td>
</tr>
<tr>
<td>Different varieties</td>
<td>Nabonne and Nerac</td>
<td><em>Energy</em></td>
</tr>
</tbody>
</table>
Table 4.11: the biocrystallogram image, difference in visual characteristic and box plot of texture parameters for carrot market in 2004.

O1 (organic)  
More side needle, more thinning out, more dense radial

O2 (conventional)  
more quernadeln and irregularity ramifications

Sum entropy, Carrot market 2004
Less *quernadeln* and more curve of main stem  

More *quernadeln* and *irregurality* ramification
More substance spirals and curve

More thinning out and dense radial quite black and white region
R1 (organic)                                      R2 (conventional)

More substance spirals and center coordinate      More thinning out and quernadeln

Sum entropy, Carrot market 2004
S2 (organic)  
S1 (conventional)  
Non significance  

Sum entropy, Carrot market 2004
General organic and conventional comparison

Organic

conventional

More *substance spirals*
and more curve of main stem

More *thinning out* and
less *side needle*

Organic and conventional comparison, Carrot market 2004
Different varieties

Nabonne

More substance spirals and side needle

Nerac

less side needle, more quernadeln and irregularity ramification
Table 4.12: The strongest texture parameter to differentiate carrot market in 2005. The sample coding information can be found more in appendix A.4.

<table>
<thead>
<tr>
<th>Sample coding</th>
<th>Variety</th>
<th>Significant texture parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1 ~ E2</td>
<td>Nabonne</td>
<td><em>Diagonal moment</em></td>
</tr>
<tr>
<td>F1 ~ F2</td>
<td>Nerac</td>
<td><em>Cluster shade</em></td>
</tr>
<tr>
<td>G1 ~ G2</td>
<td>Nerac</td>
<td>Non significance</td>
</tr>
<tr>
<td>H1 ~ H2</td>
<td>Nerac</td>
<td><em>Inverse different moment</em></td>
</tr>
<tr>
<td>Y1 ~ Y2</td>
<td>Nerac</td>
<td><em>Cluster prominence</em></td>
</tr>
<tr>
<td>General organic and conventional comparison</td>
<td>-</td>
<td><em>Inverse difference moment</em></td>
</tr>
<tr>
<td>Different varieties</td>
<td>Nabonne and Nerac</td>
<td><em>Diagonal moment</em></td>
</tr>
</tbody>
</table>
Table 4.13: The biocrystallogram image, difference in visual characteristic and box plot of texture parameters for carrot market in 2005.

<table>
<thead>
<tr>
<th>E1 (organic)</th>
<th>E2 (conventional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>More quernadeln</td>
<td>more clear stem and side needle</td>
</tr>
</tbody>
</table>

Diagonal moment, Carrot market 2005
F2 (organic)  
More dense radial

F1 (conventional)  
More Quernadeln and clear stem

Cluster shade, Carrot market 2005
G2 (organic)  G1 (conventional)

Non significance

Inverse difference moment, Carrot market 2005
H2 (organic)

More clear stem, side needle and quernadeln

H1 (conventional)

More dense radial and thinning out

Inverse difference moment, Carrot market 2005
More dense radial and quernadeln

More clear stem

Cluster prominence, Carrot market 2005

Y1 (organic)                                                Y2 (conventional)
General organic and conventional comparison

Organic

Conventional

More clear stem and quernadeln

More dense radial

Organic and conventional comparison, Carrot market 2005
Different varieties

Nabonne

More clear stem and side needle

Nerac

More quernadeln

Different varieties, Carrot market 2005
The strongest texture parameters from carrot market results can be divided into 2 groups. The 1<sup>st</sup> group is energy, it reveals the strongest effect with dense radial, quernadeln and substance spirals. The 2<sup>nd</sup> group comprises of diagonal moment, cluster prominence, cluster shade and inverse difference moment. It discloses strong behavior for: different clear stem versus quernadeln, clear stem and quernadeln versus dense radial, clear stem versus more dense radial plus quernadeln.

4.3.2 Summary of the relation of texture parameters and visual evaluation

The main sake of this task is to discriminate the image textures for all samples and relate texture parameters with the visual evaluation. The strongest texture parameters of carrot can be divided into 2 groups.

1. The 1<sup>st</sup> group is energy which can unveil the strongest effect to discriminate a paired biocrystallogram images that are different visual evaluation criteria dense radial, quernadeln and substance spirals. The energy value seems to be higher with the substance spirals, quernadeln and dense radial, respectively.

2. The 2<sup>nd</sup> group comprises of diagonal moment, cluster prominence, cluster shade and inverse difference moment which can show the strongest effect to discriminate with different clear stem versus quernadeln, clear stem and quernadeln versus dense radial, clear stem versus more dense radial plus quernadeln.
Chapter 5: Discussion with conclusions and future work

The texture analysis has been applied to the biocrystallization method within the BMELV project 020E170/F (German Research Project that is financed by the German Federal Ministry of Food, Agriculture and Consumer Protection). This research is an innovative pilot method to examine the effect of image parameters and discriminate the organic from conventional food as well. From this consequence, there is not much specific scientist literature, except a project report, that has been related to this part to the work (Organic Eprint, http://orgprints.org/4815/). Therefore, the discussion with conclusions and future work part are disputed with the results from this study, available image analysis and statistics literatures. The chapter can be summarized into prime 3 parts which can attain for the dissertation objectives as follows

(a) Optimize the statistical model which can describe all the effects that contribute to the experiment. Because the old statistical model that was previously established calculates the P value with a very conservative approach. The optimal statistical model is then embedded into a cooperation between the ACIA program and the R statistical program. Next it is verified whether it can properly work for the differentiation of all samples in question.

(b) Investigate the effect of image parameters on the texture analysis of biocrystallogram image, i.e., region of interest (ROI), color transformation and histogram matching.

(c) Consider the strongest effect of texture parameter that is selected by using the best 3 criteria of image parameters -ROI, color transformation and histogram matching- with the visual evaluation so that it can clarify how the relation of the texture parameter and visual characteristic on an image is.
5.1 Optimization of the statistical model

The optimization of the statistical model aims:

(a) Optimize the statistical model by using linear mixed effect model (lme).

(b) Verify the same statistical model which is calculated from different two statistical programs (R and SAS) in order to ensure that the statistical model is correctly generated.

(c) Execute this corrected statistical model, so called “refined statistical model”, with the ACIA program – the main program to analyze image texture - until it reasonably works.

The discussion and conclusion can be followed as

1. All of this research experiments are designed as mixed effects with repeated measurements that incorporate both fixed effects-days, samples and chambers- and random effects-sample preparation and replication in the chamber. Furthermore, the experiments are also designed by crossed day and chamber together. Hereby, the full mixed effect statistical model (notation from the chapter 4.1) comprises of

   day+ sample + chamber + T1 + T2 + T3 + T4 + T5

First order interactions are T1 and T2,
   where T1 is day. sample.
   T2 is day. chamber.

Higher order interactions are T3, T4 and T5,
   where T3 is day. sample. sample preparation.
   T4 is day. sample. sample preparation. chamber.
   T5 is day. sample. sample preparation. chamber. replication.

2. The achieved statistical model is completely programmed and set up through R program version 2.1.0 as it is the main statistical program which is primarily operated for this research. The main function that is used to fit the statistical model is lme (linear mixed effect model). Herewith, the statistical model which can be elaborated with the R program (neglect the higher order interaction) is

   lme (value~ day+sample+chamber+day:sample+day:chamber, data, random ~1|groupVektor)

where groupVektor is generated with by putting day, sample preparation, sample and chamber in one group (day/ sample preparation/ sample/ chamber) in order to get a single random effect for each group which has unique levels for each value. It will be accounted by unique level of 1 image within 1 chamber to 1 sample preparation within 1 sample and then through within 1 day. Furthermore, the complete lme model can be also derived into 16 cases so that it is able to be induced more flexible for using such as reduced data set, i.e., 1 day or 1 chamber etc (appendix B).

The ANOVA results which are modeled by the R and SAS program with the same statistical model are compared in order to check that these two programs produce the
same ANOVA. If these two programs do, the statistical model is, thus, confident to use for the next step. In contrary, if these two programs do not, the statistical model is necessarily finer-tuned until they turn out the same ANOVA.

The ANOVA from both programs turn out identical result. Moreover, the P value from the lme model reveals more conservative than the result from SAS program. The statistical result, particular P value, from R program gives more realistic approach to the experimental problem than SAS program result, as the Dendf with R program is determined by the grouping level at which the term is estimated (Pinheiro and Bates, 2000). Meanwhile SAS program determines the entire group level. However, it is difficult to decide exactly how this denDf should be done for the models as it is still under research (Bates, 2005). From the comparison result, it can be concluded that the mixed statistical model that is implemented is correct and so called “refined statistical model”.

3. The refined statistical model with crossed factor is programmed and modified into ACIA program version 1.2.3. The calibration is done by comparing ANOVA results of the refined statistical model (with crossed factor) with the old statistical model (without crossed factor). Expectedly, all 4 checked data cases (chapter 4, topic 4.1.4) present that the refined statistical model provides similar pattern of F and P values and additionally, improves these values better than the old model. From this consequence, it can be concluded that the refined statistical model is completely accomplished and usable for future work with ACIA program.

**Summary of the optimization of the statistical model**

1. The statistical model is achieved by using lme model, because the experiments are designed as mixed effect with repeated measurements.

2. The achieved statistical model is set up and precisely generated by R program, so called “refined statistical model”.

3. The refined statistical model is successfully applied and suitable for future work by calibrate and modify via ACIA program.
5.2 Investigate the effect of image parameters on texture analysis of biocrystallization image: ROI, color transformation, histogram matching

The effect of Region of interest (ROI) on texture analysis can be discussed as follows:

The region of interest (ROI) is the region which is specified to examine which area of an image contains relevant information for the differentiation objectives. As the specified area contains the number of pixels from an image and contributes to the texture analysis, so the main ROI should be investigated for the strongest effect to discriminate all samples. The result can be summarized as follows:

1. The 7 circle-ROIs are determined by varying from 40 to 100% of the total area. The results present that when ROI changes, the number of pixels in that specified area of an image is different (figure 5.1a). And it then contributes to the different GLCM from which the different texture parameters are in that area are calculated (figure 5.1b). From the results, it can be seen that the strongest effect of the differentiation is steadily increased from circle-40 to the summit in between circle-60 and -80. After that, it declines and eventually reaches the lowest at circle-100 (figure 5.1c). This phenomenon shows an important point that is the changing of texture parameters is smooth and shows monotonic pattern over ROI changes. The differentiation of samples will be even better when the ROIs of image are enlarged and located in the middle area of images. Because a region will be perceived to have texture and shown the obvious difference when the number of primitive objects in the region is large enough. If only a few primitive objects are present, then a group of countable objects is perceived instead of texture image (Tuceryan and Jain, 1998).
Figure 5.1a: region of interest (ROIs) of carrot from University of Kassel in 2004 are varying from circle-40(a1), circle-80(a2) and circle-100(a3). When ROI changes, the number of pixels in that specified area of an image is different.
Figure 5.1b: different GLCM which was calculated according to the ROIs changes in figure 5.1a, circle-40(b1), circle-80(b2) and circle-100(b3). It contributes to the different texture parameters in that area are calculated.
Figure 5.1c: F and P value plotting of texture parameters versus ROIs changes in figure 5.1a. The ROI results show that the changing of texture parameters is smooth and shows the same movement. The differentiation is located in the middle area of image.

2. The best ROI that can discriminate biocrystallization images for all samples is located within the middle area of images. Herewith, the best ROI of wheat DOC is circle-80 and wheat market is circle-70, while the best ROI of carrot from University of Kassel and carrot market are also at circle-80 and circle-70, respectively. When the best ROI is related to the visual characteristics, the curly needles and gesture characters which are situated in the periphery of CuCl$_2$ crystallized image may be not included into the analyzed texture.

3. The best ROI can discriminate the market samples both in different varieties, different organic and conventional samples. Furthermore, it provides stronger effect to differentiate varieties than fertilizer levels in carrot from University of Kassel, as the texture parameters of different varieties present monotonic pattern and express significantly over 3 years, meanwhile the different fertilizer levels show uncertainty in texture parameters pattern.
The effect of color transformation on texture analysis can be described as follows.

Color which is the range of human visual sensation does not perform only a key role in computer, television and scanner, but also is a basis feature that can be used to interpret for the information on an image. The purposes of this section are

(a) investigate whether the different representative 3 dimensionalities of colors -for example, intensity, color wavelength- contribute and show some effects to the image analysis or not.

(b) investigate whether the transformation of color can enhance more effective to differentiate samples and images from the extraordinary sample treatments, such as, various nitrate levels in carrot from University of Kassel or not.

From above sakes, the color on biocrystallization image is monitored by transforming the original color images that are based on three primary colors (RGB color) to alternative colors, namely, equal, red, green, blue, default, luminance, chroma and hue colors. The color transformation can be discussed and deduced as follows

1. The best color transformation is an equal color, as it provides stronger effect and more stable movement than the other colors. The colors (hue and chroma) do not contribute to wheat biocrystallization image, while the intensity color contributes to carrot image. The effect of color wavelength (hue) does not produce any strong effect to wheat image but seems to enhance more effective with carrot images, especially, nitrate fertilizer levels.

2. The chief effect of color transformation of wheat sample is based on the gray color, as the main effective color only occupies on the gray color scale. Moreover, the chroma and hue colors do not provide any effect. Figure 5.2 can be explained that the effect of red, green and blue colors are nearly equal and imply as gray color. Additionally, the identical effect of red, green and blue results also conform to equal, default and luminance colors that their principles are based on gray color scale as well. From this consequence, it can be interpreted that the color effect of wheat biocrystallogram image is not from the dominant effect of the original color, but depends on the gray color.

3. The crucial color transformation effect of carrot exposed that the effect of color provides nearly an identical effect when the color transformation to equal, red, green, blue, default, luminance and chroma, meanwhile, suddenly diminish at hue color (figure 5.3). That means, the color transformation of carrot biocrystallogram image is relied on 2 dimensions of both gray and chroma colors (purities of its color on image). This phenomenon might be criticized that, although hue and chroma color are always relevant to each other the hue color has more very variable reliability (Lambert and Carron, 1999). But in case of carrot, hue color may be not shown enough relevance with chroma color until shown stronger effect like chroma and gray colors. Anyhow, the hue color shows some effect to different nitrate fertilizer levels in carrot from University of Kassel. But its behavior is uncertainty pattern. It might be reasoned from either different climate each year or truly effect of light wavelength. Therefore, in this case, the number of years should be focus and study more than 3 years in order to exactly clarify for this incident of hue color.
Figure 5.2: F and P value versus color transformation of wheat. The graph shows that the chroma and hue colors present no information. Meanwhile, equal, red, green, blue, default and luminance colors are nearly equal and relied on gray color.

Figure 5.3: F and P value versus color transformation of carrot. The graph shows that there is an identical effect when the color transformation to equal, red, green, blue, default, luminance and chroma, meanwhile, suddenly decrease at hue color.
4. The color transformation of wheat and carrot market displays better performance to differentiate organic and conventional samples than the difference in varieties. As the color transformation to differentiate organic and conventional samples shows the same and certainty pattern, whilst, different varieties found uncertainty pattern. This color result presents unconformity to the result of ROI effect that provides the strongest effect both in different varieties and organic versus conventional comparisons.

5. The result of carrot from University of Kassel unveils that the color transformation provides stronger effect by certainty pattern with different varieties than the different levels of nitrate fertilizer.

**The effect of histogram matching on texture analysis** can be discussed as follows

The image histogram is an important stride to provide basic information of the appearance of an image. Some images show poor contrast, because their histograms utilize a restricted range of brightness value or saturation in black or white regions. Therefore, this part aims to investigate whether the original histogram is transformed to the other matching types. They may enhance the differentiation images. By that, the histogram matching like Gaussian, equalization, triangle, parabolic and hyperbolic matching are focused in this study. It can be manifested as follows

1. The original histograms of biocrystallogram for all samples, most of the time, are rather skewed rightward within white region. When the histogram is matched to other distributions, the differentiation can be enhanced to differentiate all samples in questions. The best histogram matching is Gaussian matching, because it overall provides strongest effect and gives also the greatest significant pairs of assignments.

2. The histogram matching can be divided into 2 groups by using their F and P value movement and the competence to enhancing the contrast of original histogram. The 1st group is histogram equalization and Gaussian matching (figure5.4). Their competences enhance more effective when the comparison of paired original histogram is either (a) these 2 samples are skewness in white region or (b) one is skewness shape and another one is nearly Gaussian shape. The 2nd group consists of triangle, parabolic and hyperbolic matching (figure5.4). They perform well with the original histogram that is compared Gaussian and Gaussian matching. This behavior can be found in corn DOC trial samples in 2003 and 2005 (appendix. G). The reason why it is able to be divided into 2 groups may be criticized from the difference in fundamental distribution pattern, namely, the histogram distribution of the 1st group is not skewed, i.e., uniform and Gaussian distribution, while the 2nd group is skewed rightward.
Figure 5.4: (a) F value of texture parameters versus histogram matching changes. (b) P value of texture parameters versus histogram matching changes. (c) Box plot of \textit{diagonal moment} value versus histogram matching. The histogram matching divides into 2 groups, i.e., the 1\textsuperscript{st} group is histogram equalization and Gaussian matching and the 2\textsuperscript{nd} group is triangle, parabolic and hyperbolic matching.
3. The texture parameters can be divided into 2 prime groups by concerning on their inverse movements through the histogram matching effect. The 1st group is diagonal moment, cluster shade and cluster prominence. The 2nd group is energy, sum energy, different energy and the others. It might be caused from the dissimilarity measurement methodology, as the 1st group measures the different of the variation of gray scale in an image, meanwhile, the 2nd group mainly determines the uniformity distribution.

4. When the histogram is matched with Gaussian matching, the energy group can be efficiently augmented the contrast of biocrystallogram images which the visual evaluation criteria are remarkable through substance spirals, side needle, quernadene and dense radial. Meanwhile the diagonal moments group shows distinguish effect as can be observed at clear stem, less side needle, quite different in black and white region in crystallogram image.

5. The histogram matching results from market samples unveils that Gaussian matching performs with the difference in farm pairs and varieties better than the general organic and conventional comparison, as it shows more efficiency to differentiate sample from different varieties and farm pairs, whilst the general organic and conventional samples reveal that all histograms are identical (significant effect). This result does not coincide to the achievement of ROI and color effects, i.e., the ROI shows powerful to discriminate both difference in varieties and general organic versus conventional sample, whilst, the color transformation reveals stronger effect with general organic and conventional comparison than different varieties.

6. The histogram matching of carrot from University of Kassel discloses uniformity effect with both color and ROI results. That can be concluded that Gaussian matching carries out better with the different varieties than nitrate fertilizer levels(tendency).

Summary of the investigation the effect of image parameters on texture analysis of biocrystallization image: ROI, color transformation, histogram matching

1. The ROI result shows a valuable meaning, i.e., the alteration of texture parameters is smooth and monotonic pattern over ROI changes. The differentiation of sample will be even better when the ROI is enlarged and located in the middle area of images, i.e., between circle-60 till -80.

2. The best ROI of wheat DOC and carrot from University of Kassel is circle-80, meanwhile, wheat and carrot market is circle-70. When related to the visual characteristics, the curly needles and gesture characters may not be included to the analyzed texture.

3. The best color transformation is an equal color as it provides the stronger effect and more stable movement than the other colors.

4. The hue and chroma colors do not contribute to differentiate wheat biocrystallogram image. Furthermore, the intensity and chroma colors contribute to carrot image. The effect of color wavelength (hue) does not provide any strong
effect to wheat image analysis. However, it seems to be enhanced more effective with carrot image, especially for nitrate fertilizer levels.

5. The color transformation of farm pairs exposes better performance to differentiate organic and conventional samples than the difference in varieties. The result of carrot from University of Kassel unveils that the color transformation provides stronger effect with different varieties than the different levels of nitrate fertilizer.

6. The histogram matching can be divided into 2 groups by using their competence to enhance the contrast of the original histogram. The 1st group is histogram equalization and Gaussian matching. The 2nd group consists of triangle, parabolic and hyperbolic matching.

7. The best histogram matching to differentiate all samples is Gaussian matching.

8. The texture parameters can be devised into 2 main groups (a) the energy group that shows more effective through visual evaluation criteria substance spirals, side needle, quernadeln and dense radial. (b) The diagonal moment group shows distinguish effect with clear stem, less side needle, quite different in black and white region in crystallogram image.

9. Gaussian matching performs with the difference in varieties and farm pairs better than the general organic and conventional comparison. Regard to carrot from University of Kassel, the Gaussian matching acts better with the different varieties than nitrate fertilizer levels(tendency).
5.3 The relation of texture parameters and visual evaluation

The crucial task of this part is to discriminate all samples and relate texture parameters with the visual evaluation (appendix E) that have been developed by triangle researcher group (University of Kassel, Germany; Louis Bolk Institute (LBI), Netherlands and Biodynamic Research Association Denmark (BRAD), Denmark).

The texture analysis can statistically differentiate wheat DOC samples in both years, especially, discriminating organic from conventional samples. The samples from organic and conventional farm pairs (wheat and carrot market) can be also differentiated as obviously seen in the statistical significance (70-80% correct differentiation). Additionally, the different varieties can be also significantly discriminated. The carrot from University of Kassel, over 3 years, can be differentiated more significant in varieties than fertilizer.

The strongest texture parameters from carrot samples are able to be divided into 2 groups. The 1st group which is energy group unveils the strongest effect with the difference in dense radial, quernadeln and substance spirals in images. Herewith, the energy value turns out higher value with the substance spirals, Quernadeln and dense radial, respectively. The 2nd group which comprises of diagonal moment, cluster prominence, cluster shade and inverse difference moment shows strong behavior with the difference in clear stem versus quernadeln, clear stem and quernadeln versus dense radial, and clear stem versus more dense radial plus quernadeln.

The texture parameter groups which are above mentioned can be verified by sorting the strong visual characteristics on images that coincide with these 2 groups of texture parameters. After that, The ANOVA results of the strongest texture parameters are checked by plotting F value versus ROI changes (figure 5.5 for diagonal moment group checking and figure 5.6 for energy group checking).

Figure 5.5 does confirm that the diagonal moment group definitely shows distinguish effect with the paired images which are different visual evaluation criteria clear stem, dense radial and quernadeln. At the mean time, figure 5.6 also proves that the energy group shows a key effect to differentiate the different visual evaluation criteria, namely, substance spirals, quernadeln and dense radial in paired image.
Figure 5.5: (a) The different biocrystallogram image and their GLCM according to the diagonal moment group. (b) F value of texture parameters versus ROI change.
Figure 5.6: (a) the different biocrystallogram image and their GLCM according to the energy group. (b) F value of texture parameters versus ROI change.
Summary of the relation of texture parameters and visual evaluation

1. The majority effective texture parameters can be divide into 2 groups as follows

1.1 *Energy* group unveils strongest effect to discriminate paired images that are different in visual evaluation criteria *dense radial, quernadeln* and *substance spirals* in carrot, by that, the *energy* value provides higher value via the *substance spirals, quernadeln* and *dense radial*, respectively.

1.2 *Diagonal moment* group shows highly effective performance with paired images that are different in visual evaluation criteria *clear stem* versus *quernadeln*, *clear stem* versus *dense radial* and *clear stem* versus *dense radial plus quernadeln*.

2. This research relates the texture parameters and visual evaluation criteria by starting from the best performance of texture parameters. Then, connecting them to the visual evaluation criteria. The texture parameters which differentiate images from conventional and organic samples can be related to visual evaluation criteria. The single texture analysis parameter can not be used for classification according to organic and conventional samples. Also the selected visual evaluation criteria are uncertainty, e.g., sometimes *quernadeln* is taken place via conventional image meanwhile this character is also occurred through organic image as well. The main reason may be argued from the difference in farming system affected by factors like climate, variety or soil type. For this work it was the aim to connect the texture analysis parameter to visual feature. Therefore, only those features were chosen which can be related to image texture analysis results. Kahl (2006) showed that starting from the picture as the whole, other visual features where selected which are able to classify picture according to organic and conventional samples. In all adjustments not only one single but multiple criteria were applied. Therefore, the present work also shows the limitation of texture analysis by applying single parameter only.

3. Moreover, the texture analysis result in this research can statistically differentiate wheat samples from DOC trial, especially between the organic and conventional samples. The samples from organic and conventional farm pairs (wheat and carrot market) can be also differentiated by using statistical significance (70-80% correct differentiation), additionally, the different varieties can be significantly discriminated. The carrots from University of Kassel can be differentiated over 3 years. There is a stronger significance in varieties than fertilizer (tendency).
5.4 Future work

The texture analysis is important in many applications of computerized image analysis. Because it is able to present much information that contains in an image. In the present work this analysis studied the effect of image parameters, i.e., ROI, color transformation and histogram matching. After that, it was applied to discriminate organic from conventional biocrystallogram image, especially from wheat and carrot sample. The results disclose a valuable meaning that the texture analysis can differentiate most of them. Thereby, the region of interest (ROI) seems to be the strongest effect when compare with the color transformation and histogram matching. Because all of texture parameters present monotonic movement pattern, i.e., even enlarge stronger effect when ROIs change until circle-60 till -80, then is decreasing and reaches the lowest at the edge of image. Anyway, there are still some limitations to apply this method on biocrystallogram picture (e.g., classification of images according to sample origin). Therefore, the future works should be further to study as follows.

1. The color transformation shows uncertainty effect to clarify wheat as well as the extraordinary treatments, e.g., the different nitrate fertilizer levels. Therefore, the elements of colors should be further deeply focused such as the effect of luminance, chroma and hue colors.

2. The calculated texture parameters in this research were computed via the fundamental GLCM at (0, 1) where 0 means direction of pixel pairs on a picture, this means direction $e = 0^\circ$. And 1 refers to distance between the pixel pairs, i.e., one step in the horizontal direction. In order to perceive which exactly the direction as well as distance on an image provides more effective with the texture analysis so they should be also focused by vary several directions and distances.

3. As the GLSH and GLDH calculation shows some effective on biocrystallogram image, hence, the calculating and comparing their histogram should be more deeply study in order to explain how their performance react on an image.

4. The differentiation of all samples in this research are analyzed merely the texture in an image. Anyhow, it would be more benefit when the structure of an image is induced to sort out the quality and quantitative of images as it will provide more understanding how the image patterns are.

5. This research results unveil that the texture analysis can only differentiate organic from conventional wheat and carrot samples. So the classification of them should be further studied (including more years, soil type or varieties).

6. Moreover, it would be more practical when the texture analysis approach is applied to differentiate other common food such as other vegetables, cereal and dairy products etc.
Chapter 6: Abstract

The consumers are becoming more concerned about food quality, especially regarding how, when and where the foods are produced (Haglund et al., 1999; Kahl et al., 2004; Alföldi, et al., 2006). Therefore, during recent years there has been a growing interest in the methods for food quality assessment, especially in the picture-development methods as a complement to traditional chemical analysis of single compounds (Kahl et al., 2006). The biocrystallization as one of the picture-developing method is based on the crystallographic phenomenon that when crystallizing aqueous solutions of dihydrate CuCl$_2$ with adding of organic solutions, originating, e.g., from crop samples, biocrystallograms are generated with reproducible crystal patterns (Kleber & Steinike-Hartung, 1959). Its output is a crystal pattern on glass plates from which different variables (numbers) can be calculated by using image analysis.

However, there is a lack of a standardized evaluation method to quantify the morphological features of the biocrystallogram image. Therefore, the main sakes of this research are (1) to optimize an existing statistical model in order to describe all the effects that contribute to the experiment, (2) to investigate the effect of image parameters on the texture analysis of the biocrystallogram images, i.e., region of interest (ROI), color transformation and histogram matching on samples from the project 020E170/F financed by the Federal Ministry of Food, Agriculture and Consumer Protection(BMELV). The samples are wheat and carrots from controlled field and farm trials, (3) to consider the strongest effect of texture parameter with the visual evaluation criteria that have been developed by a group of researcher (University of Kassel, Germany; Louis Bolk Institute (LBI), Netherlands and Biodynamic Research Association Denmark (BRAD), Denmark) in order to clarify how the relation of the texture parameter and visual characteristics on an image is.

The refined statistical model was accomplished by using a lme model with repeated measurements via crossed effects, programmed in R (version 2.1.0). The validity of the F and P values is checked against the SAS program. While getting from the ANOVA the same F values, the P values are bigger in R because of the more conservative approach. The refined model is calculating more significant P values.
The optimization of the image analysis is dealing with the following parameters: ROI (Region of Interest which is the area around the geometrical center), color transformation (calculation of the 1 dimensional gray level value out of the three dimensional color information of the scanned picture, which is necessary for the texture analysis), histogram matching (normalization of the histogram of the picture to enhance the contrast and to minimize the errors from lighting conditions).

The samples were wheat from DOC trial with 4 field replicates for the years 2003 and 2005, "market samples" (organic and conventional neighbors with the same variety) for 2004 and 2005, carrot where the samples were obtained from the University of Kassel (2 varieties, 2 nitrogen treatments) for the years 2004, 2005, 2006 and "market samples" of carrot for the years 2004 and 2005.

The criterion for the optimization was repeatability of the differentiation of the samples over the different harvest (years). For different samples different ROIs were found, which reflect the different pictures. The best color transformation that shows efficiently differentiation is relied on gray scale, i.e., equal color transformation. The second dimension of the color transformation only appeared in some years for the effect of color wavelength (hue) for carrot treated with different nitrate fertilizer levels. The best histogram matching is the Gaussian distribution.

The approach was to find a connection between the variables from textural image analysis with the different visual criteria. The relation between the texture parameters and visual evaluation criteria was limited to the carrot samples, especially, as it could be well differentiated by the texture analysis. It was possible to connect groups of variables of the texture analysis with groups of criteria from the visual evaluation. These selected variables were able to differentiate the samples but not able to classify the samples according to the treatment. Contrarily, in case of visual criteria which describe the picture as a whole there is a classification in 80% of the sample cases possible. Herewith, it clearly can find the limits of the single variable approach of the image analysis (texture analysis).
ZUSAMMENFASSUNG


Um die Standardisierung der Auswertung zu verbessern, wird in dieser Arbeit (1) das statistische Modell für die Auswertung der computergestützten „image analysis“ Ergebnisse optimiert, (2) die Parameter der „image analysis“ hinsichtlich einer Optimierung der Differenzierung von Weizen- oder Möhrenproben aus dem BMELV Projekt 020E170/F hinsichtlich Anbau und Sorten untersucht und (3) untersucht wie sich die Ergebnisse der „image analysis“ in den Kriterien der visuellen Beurteilung, wiederfinden welche innerhalb einer Forschungsgruppe bestehend aus: Universität Kassel, Louis-Bolk-Instituut (NL), Biodynamic Research Association Danmark (DK) entwickelt wurden.


Bei der Optimierung der „image analysis“ Parameter geht es um folgende Größen: ROI (Region Of Interest ist Fläche um das geometrische Zentrum des Bildes), „color transformation“ (Umrechnung der dreidimensionalen Farbinformation des Bildes auf die eindimensionalen „Grauwerte“, die von der Texturanalyse nur verarbeitet werden können), „histogramm matching“ (Normierung des Histogramms des Bildes, um den Kontrast zu erhöhen und Beleuchtungsfehler zu minimieren).

Als Beurteilungskriterium wurde die Wiederholbarkeit der Differenzierbarkeit der Proben über die vorhandenen Jahre genommen. Für die verschiedenen Probenarten wurden pro Fragestellung verschiedene ROIs gefunden, in denen sich die Verschiedenheit der Bilder abbilden. Für „color transformation“ und „histogramm matching“ wurden für alle Proben dieselben Einstellungen gefunden, die mit den Ausgangseinstellungen übereinstimmten.


Die Untersuchung der Verbindung der „image analysis“ Ergebnisse mit den visuellen Kriterien wurde auf die Möhrenproben beschränkt und insbesondere auf solche, die sich stark signifikant unterschieden. Es wurde versucht anhand der einzelnen „image analysis“ Variablen und der Art und Weise, wie sie berechnet werden eine Verbindung zu den
Chapter 7: References


Andersen, J-O.2005. Ability to resist degradation from stressful conditions as a quality indicator-A newly developed microwave heating test for biocrystallization. Proceeding of the 1st scientific FQH conference, Frick, Switzerland.


Andersen, J-O., and Busscher, N. 2001. Effect of the different levels of mineral N and light intensity on the pictomorphological properties of barley (Hordeum sativum,


Mahony, M.O.1986. Sensory evaluation of food, statistical methods and procedures. New York :Marcel Dekker.


