

Estimating the Weight of Ginseng Using an Image Analysis

Seokhoon Jeong[†] · Kuk Won Ko^{††} · Ji-Yeon Lee^{†††} · Jinho Lee^{††††}
Hyeonseok Seo^{†††††} · Sangjoon Lee^{††††††}

ABSTRACT

This study is to estimate proximity without direct measurement of the weight of fresh ginseng. For this work, we developed a ginseng image acquiring instrument and obtained 126 ginseng images using the instrument. Image analysis and parameter extraction process was used C language based Labwindows/CVI development tools and open source library OpenCV. Estimation formula is made by weighing the sample with image analysis of fresh ginseng. We analyzed the correlation between the pixel number and the weight of ginseng using a linear regression approach. It was obtained a strong positive correlation coefficient of 0.9162 with a linearity value.

Keywords : Ginseng Grade Decision Making, Estimating of Weight, Linear Regression Analysis

영상 분석을 이용한 수삼의 중량추정

정 석 훈[†] · 고 국 원^{††} · 이 지 연^{†††} · 이 진 호^{††††} · 서 현 식^{†††††} · 이 상 준^{††††††}

요 약

본 연구의 목적은 수삼 등급을 판정하는 기준 중에서 수삼의 중량을 직접 측정하지 않고 영상 분석으로 가능한 근접하게 추정하는 것이다. 이를 위해 수삼영상 취득장치를 제작하였으며 126개의 수삼샘플을 대상으로 영상을 취득하였다. 각각의 수삼샘플의 중량을 측정하고 영상분석 데이터를 이용하여 중량추정공식을 산출하는데 이용하였다. 영상분석과 파라미터 추출과정에는 C언어 기반의 Labwindows/CVI 개발 툴과 오픈 소스 라이브러리 OpenCV를 이용하였다. 영상분석 과정에서 추출한 파라미터와 중량과의 상관관계를 가장 잘 표현할 수 있는 필터설정 값을 추적하여 적용하였고, 최소제곱법을 사용한 선형 회귀분석으로 선형성을 가지는 0.9162의 강한 양의 상관계수 값을 얻을 수 있었다.

키워드 : 수삼등급판정, 무게추정, 선형회귀 분석

1. 서 론

대한민국은 인삼의 중추국으로 담배인삼공사에서 제조·판매하는 홍삼 및 홍삼 제품은 그 품질 및 신용도에 있어서 세계적으로 인정받고 있다. 홍삼제조용 수삼 수납시의 등급 판정은 전문 감정인에 의하여 1~3등 및 등외로 4등급으로 분류되고 있으며, 각 등급은 수삼의 너두, 동체, 그리고 각부에 대한 체형, 중량, 표피의 색택, 균열여부 등을 감정하여

결정하고 있다[1]. 이러한 분류기준은 계량적 평가단위와 비 계량적 평가단위가 복합적으로 구성되어 훈련받은 품질검사 요원이라 할지라도 관능검사로 각 등급 간의 구분 및 판정은 주관적인 판단에 의존하게 되어 검사요원마다 등급판정에 차이가 있을 수 있으며, 이로 인해 매년 경작인들과의 논란이 되고 있다.

원료 삼 구매 시 등급판정에 필요한 연간 소요비용은 약 10억원 가량으로 연인원 1,530명의 숙련된 검사원이 필요한 실정으로 등급판정에 어려움이 있으므로 단시간 내에 등급을 판정할 수 있는 기계적이고 객관적인 판정 방법이 요구되고 자동등급판정에 관한 연구가 진행되어 왔으나 실용화하지 못하고 있는 실정이다[2, 3].

본 연구에서는 수삼등급 판정기준의 하나인 중량정보를 영상 분석만으로 근접하게 추정할 수 있는지 알아보고 수삼 등급 자동판정의 결정파라미터로 직접적으로 중량을 측정하지 않고, 촬영된 영상데이터에서 예측하여 자동등급판정에 사용 가능성 여부를 판단하고자 한다.

* 이 논문(저서)은 2014년 교육부와 한국연구재단의 지역혁신창의인력 양성사업(NRF-2014H1C1A1066998), KGC 인삼연구원의 지원을 받아 수행된 연구임.

† 비 회 원 : 선문대학교 정보통신공학과 박사과정

†† 비 회 원 : 선문대학교 기계ICT융합공학부 부교수

††† 비 회 원 : 선문대학교 자동화에너지연구소 연구원

†††† 비 회 원 : KGC인삼공사 과장

††††† 비 회 원 : KGC인삼공사 대리

†††††† 정 회 원 : 선문대학교 기계ICT융합공학부 조교수

Manuscript Received: April 29, 2016

Accepted: May 17, 2016

* Corresponding Author: Sangjoon Lee(mcp94lee@sunmoon.ac.kr)

2. 재료 및 방법

2.1 수삼영상 취득장치

수삼영상 취득 시 고려해야 할 점은 객체의 모든 부분을 검출할 수 있는 카메라 각도와 취득된 영상의 균일한 품질과 균일한 영상인식결과를 얻기 위해 일정한 조명이 필요하다. 이를 위해 외부 조명을 완전히 차단할 수 있는 폐쇄형 영상 취득장치를 개발하였다. 개발한 장치는 5W급 LED조명을 포함한 1920×1080화소(pixel)의 해상도의 Full HD급 카메라와 영상을 취득 프로그램을 탑재한 mini PC 및 모니터를 장착하였다. 객체의 크기는 (500×500×500mm)의 충분한 크기로 제작되어 6년근 수삼의 세근까지도 측정 가능하도록 설계 하였다. 영상취득과 분석에 사용된 프로그램은 C언어 기반의 LabWindows/CVI 2013 개발 툴과 영상처리용 오픈소스 라이브러리 OpenCV(3.0)을 사용하였다.

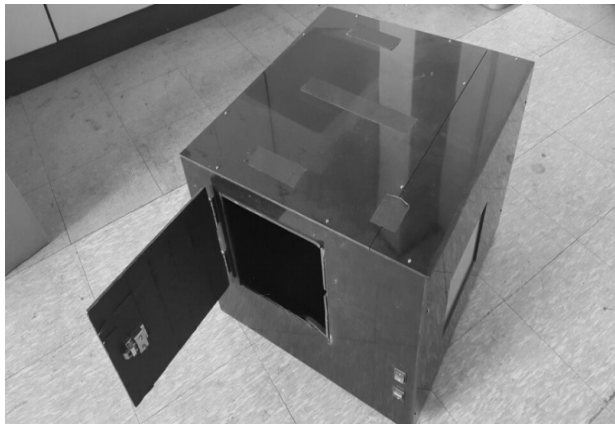


Fig. 1. Developed image acquisition system

2.2 수삼시료

실험에서 사용한 시료는 등급에 관계없이 216개의 6년근 수삼을 임의로 선정하였으며, 중량을 실측하고 수삼영상 취득장치의 상단에서 촬영한 이미지를 사용하였다. 측정된 시료의 중량은 28.4g부터 191.3g까지(평균 96.4g) 분포한다.

2.3 수삼 파라미터 추출

이미지에서 수삼의 중량과 상관관계를 유추하기 위해서 1920×1080 화소의 컬러 원본 이미지를 명암의 차이에 의하여 256 수준(8bit), 즉 0(흑색)에서 255(백색)의 gray 수준의 값으로 변환하고 이후 gray 이미지를 필터 처리하여 중량정보와 가장 밀접한 연관이 있는 파라미터를 추출하는 과정을 수행한다. 변환된 gray 이미지는 0부터 255사이의 값을 가지는 픽셀개수를 계수하여 분포도를 나타내는 히스토그램(Histogram) 분석 과정을 거치고, 수삼이미지에서 배경영역을 제거하여 몸체와 구분하기 위하여 이진화(binartization)를 한다. 이진화된 이미지는 수삼의 몸체 이외에도 배경부분에도 밝은 조명의 영향에 따라 부분적으로 남아있게 되는데 이러한 부분과 미세한 수삼의 중량에 크게 영향을 미치지

않을 것으로 예상되는 미세한 잔뿌리영역을 제거하기위한 블록필터 처리를 하여 남아있는 영역의 블록수를 계수하여 파라미터로 사용하였다.

Fig. 2는 파라미터 추출과정인 원본 이미지 로딩(Fig. 2a), Gray 이미지로 변환(Fig. 2b), 영상 이진화(Fig. 2c), 블록필터(Fig. 2d)의 순서대로 보인다.

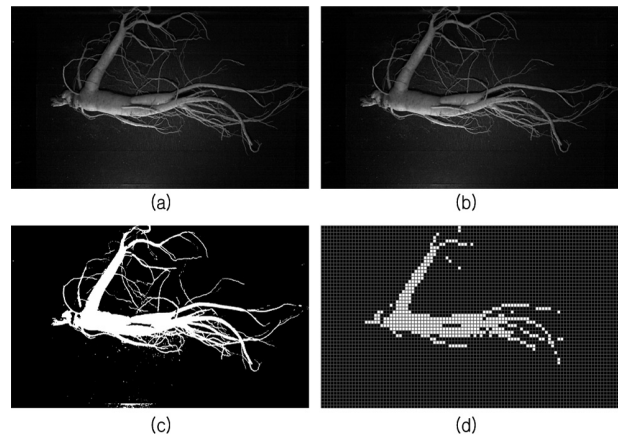


Fig. 2. Parameter extraction process (a) Original Image, (b) Gray Image, (c) Binarization Image, (d) Block Filtered Image

1) 이진화(Binarization)

영상처리에서 이진화는 어떤 주어진 임계값(threshold)보다 밝은 픽셀들은 모두 흰색으로, 그렇지 않은 픽셀들은 모두 검은색으로 바꾸는 것을 지칭한다. 이진화 방법에는 전역 고정 이진화(Global fixed thresholding), 지역 가변 이진화(Locally adaptive thresholding), 히스테레시스(Hysteresis thresholding) 등 여러 가지 방법이 있으나, 본 연구에서는 개발한 수삼영상 취득장치를 통하여 조명과 카메라 위치 등이 제한된 동일한 환경에서 영상을취득한 이미지를 사용하였기 때문에 고정된 단일 임계값을 사용하는 전역 고정 이진화 방법을 이용하였다.

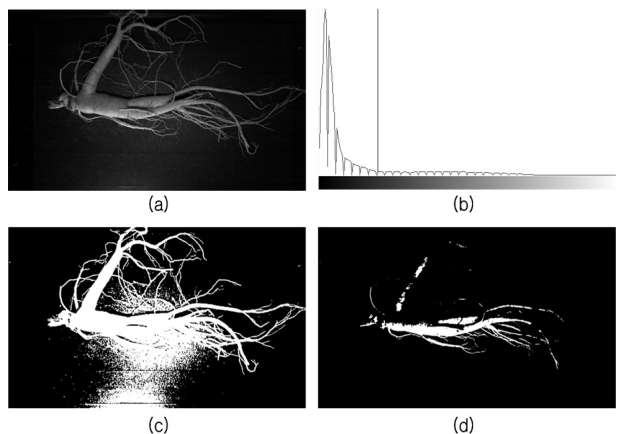


Fig. 3. Binarization result of the threshold setting change (a) Gray Image (b) Histogram (c) Binarization Threshold(30), (d) Binarization Threshold(100)

Fig. 3에서 임계값 설정변화에 따른 이진화 결과를 보인다. Gray 이미지(Fig. 3a)에서 픽셀들의 밝기값 히스토그램(Fig. 3b)을 살펴보면, 0부터 255 레벨에 고르게 분포하며 수삼과 배경과의 영역을 정확히 나눌 수 없어서 이진화는 매우 간단한 연산 이지만 임계값을 설정하는 것은 쉽지 않은 문제이기도 하다. 임계값을 너무 낮게 설정하면(Fig. 3c, 임계값 30 적용) 조명에 의한 배경의 밝은 영역이 부각되기도 하고, 너무 높게 설정하면(Fig. 3d, 임계값 100 적용) 수삼 몸통부분에서 상대적으로 어두운 부분까지도 검정으로 바꾸어 이미지 손실이 생기게 된다.

2) 블록필터(Block Filtering)

블록필터는 1920×1080 화소의 영상에서 일정크기를 갖는 블록으로 영역을 나누고 해당 블록에서 흰색픽셀의 분포가 전체의 일정 임계값(%) 이상이면 흰색으로 표시하고 이하면 검정색으로 표시하는 필터이다. 필터를 사용하는 목적은 미세한 잔뿌리 영역을 제거하여 중량의 비중을 많이 차지하는 몸통영역을 나타내기 위함이다.

Fig. 4에서 임계값을 60%로 고정시키고 블록필터 크기로 가로×세로의 블록크기를 20×20, 40×40, 60×60으로 증가시키면서 관찰한 결과를 보인다. 20×20의 블록크기(Fig. 4a)를 설정했을 경우 몸체의 형태를 나타내주면서 세근을 제거하는 효과를 나타내지만 블록크기가 커질수록 이미지에서 계단현상과 해상도 저하로 인한 손실이 나타나게 되므로, 적절한 블록크기와 임계값을 설정하는 것이 관건이다.

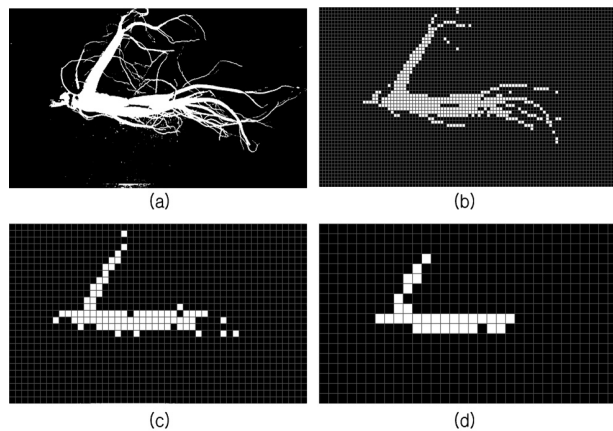


Fig. 4. Block filtering result of the setting change
 (a) Binarization Image (b) Block size 20×20
 (c) Block size 40×40 (d) Block size 60×60

3) 상관관계 분석

수삼의 중량과 수삼 이미지 픽셀값 사이의 관련성을 알아보기 위해 상관관계 분석을 실시하였다. 수삼의 중량과 상관관계를 유추하기 위하여 실제 측정된 수삼의 중량과 블록필터를 거친 후 남아있는 블록개수와의 관계를 최소제곱법(least square method)을 사용한 선형 회귀분석(Linear regression analysis)으로 가장 높은 상관계수를 (correlation coefficient) 가지도록 하는 이진화 임계값 및 블록필터의 가

로, 세로, 임계값의 파라미터를 찾는 과정을 수행한다. 먼저 이진화 임계값은 41~50으로 변화를 주면서 상관계수의 값을 관찰하여 가장 높은 상관계수를 가지는 임계값을 확정하고, 다음으로 블록필터의 설정 값은 블록크기 (가로×세로)는 동일하게 11~20픽셀로, 임계값을 86%~95%로 각각 변화를 주면서 상관계수의 변화를 관찰하였다. 본 실험에서 사용한 상관계수를 이용하여 모상관계수에 대해 유의수준 0.05로 t 분포 양측 검정을 실시하였다. 모집단에서 수삼이미지의 픽셀값과 무게의 모상관계수에 대한 검정은 모집단에서 두 변수 사이의 상관관계가 있음을 판정함과 동시에 상관계수가 1에 가까울수록 강한 양의 상관관계가 있음을 시사한다.

3. 결과 및 고찰

Gray 이미지에서 이진화 과정에서의 임계값은 블록필터 설정값을 고정시키고(Block size : 10×10, Threshold : 90%) 이진화 임계값을 41~50까지 변화시키면서 선형회귀분석에 의한 상관계수의 값을 관찰하여 상관관계가 가장 큰 43 레벨의 임계값으로 결정하였고(Table 1), 이후 Table 2와 Table 3에서 블록필터의 가로×세로와 임계값은 이진화 임계값을 43으로 고정시키고 블록필터의 크기를 11×11 부터 20×20 픽셀로 증가시키고, 임계값은 86%부터 95%까지 증가시키면서 선형회귀분석에 의한 상관계수의 값이 가장 큰 설정값으로 블록크기 18×18 픽셀, 임계값 90%의 설정값을 선정하였다.

결과는 이진화 임계값은 43이고, 18×18 크기와 90%의 임계값을 가지는 블록필터를 적용하여 출력된 픽셀개수와 수삼 중량과의 상관계수는 0.9162로 최대값을 가진다. 이 때 추정된 회귀식은 $y = 2.517x + 45.82$ 이며 추정된 회귀식에 대한 분산분석을 실시하여 추정된 회귀식이 유의성에 대해 검정하였다. 유의수준 0.05로 F검정을 실시하여 p-value= 4.4652E-51로 유의하다고 판단하였으며 결정계수값 $r^2 = 0.839$ 는 자료의 83%가 추정된 회귀식으로 설명되어짐을 의미한다. 결정계수가 83%인 이유는 자료 분석에 앞서 산포도를 그려본 결과 특이값들이 관측되었는데 그러한 특이값의 영향을 받았으리라 추정된다.

Fig. 5는 126샘플의 수삼 중량과 최적의 설정값으로 이진화와 블록필터를 수행하여 출력된 픽셀개수와의 관계를 나타내는 분포도이다. 그림에서 특이값을 보이는 a, b 점에서의 수삼과 같은 경우 대해 따로 분류하고 에러를 발생시키는 원인을 분석하였다.

- 1) 수삼 중량에 비해 픽셀개수가 상대적으로 많이 나오는 경우(Fig. 5에서 a점): 수삼의 세근이 많고 밀집된 분포로 상대적으로 많은 부피를 차지한다.
- 2) 수삼 중량에 비해 픽셀개수가 상대적으로 적게 나오는 경우(Fig. 5에서 b점): 수삼에 흠이나 이물질로 인하여 몸체가 어둡게 나오거나 수삼의 지근이나 세근등이 몸체를 덮어서 그림자를 발생시켜서 상대적으로 적은 부피를 차지한다.

Table 1. Experimental results of the binarization threshold change

Binarization Threshold	Block Filter			Correlation coefficient
	width	height	threshold	
41	10	10	90	0.9057
42	10	10	90	0.9085
43	10	10	90	0.9089
44	10	10	90	0.9087
45	10	10	90	0.9087
46	10	10	90	0.9057
47	10	10	90	0.9057
48	10	10	90	0.9043
49	10	10	90	0.9018
50	10	10	90	0.9019

Table 2. Experimental results of the block filter size change

Binarization Threshold	Block Filter			Correlation coefficient
	width	height	threshold	
43	11	11	90	0.9110
43	12	12	90	0.9112
43	13	13	90	0.9129
43	14	14	90	0.9117
43	15	15	90	0.9125
43	16	16	90	0.9158
43	17	17	90	0.9141
43	18	18	90	0.9162
43	19	19	90	0.9133
43	20	20	90	0.9089

Table 3. Experimental results of the block filter threshold change

Binarization Threshold	Block Filter			Correlation coefficient
	width	height	threshold	
43	18	18	86	0.9146
43	18	18	87	0.9147
43	18	18	88	0.9161
43	18	18	89	0.9152
43	18	18	90	0.9162
43	18	18	91	0.9159
43	18	18	92	0.9144
43	18	18	93	0.9144
43	18	18	94	0.9132
43	18	18	95	0.9114

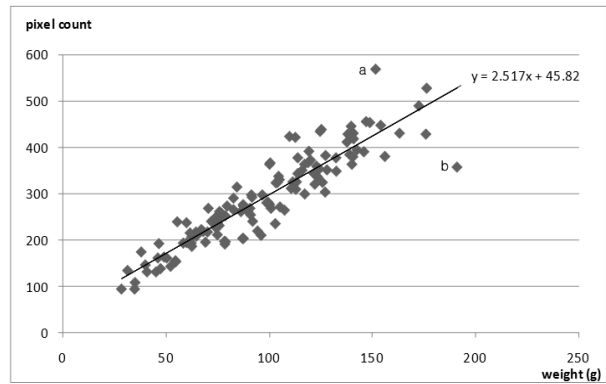


Fig. 5. Distribution of the relationship between the weight of Ginseng and the number of pixels

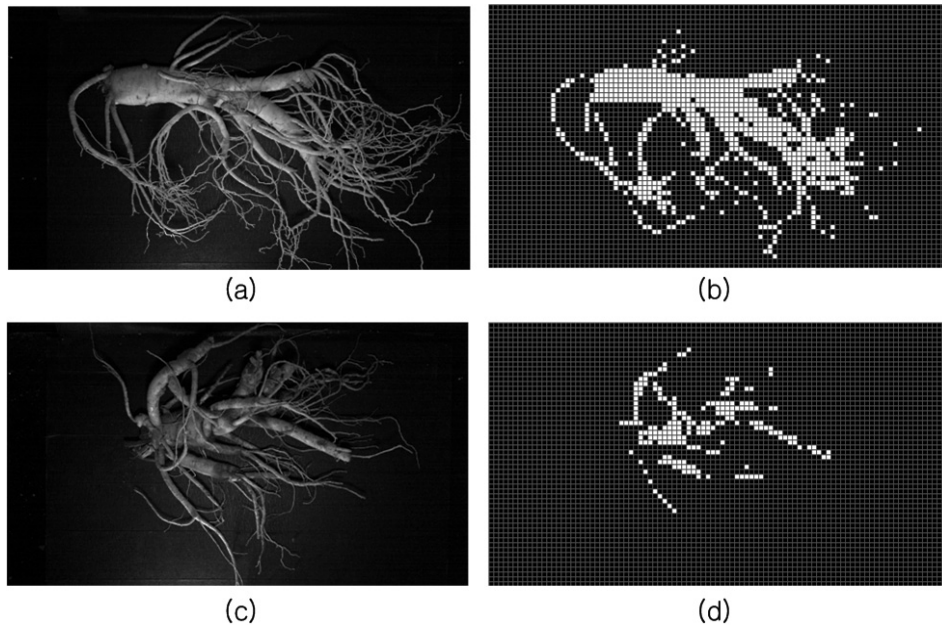


Fig. 6. Examples of the cause of an error (a) Gray image of the error case 1, (b) Filtered image of the error case 1, (c) Gray image of the error case 2, (d) Filtered image of the error case 2

Fig. 6은 에러발생의 두 가지 예를 보이며, 수삼의 세근이 많고 밀집된 분포를 가지는 경우의 원본 이미지(Fig. 6a)와 필터이미지(Fig. 6b)를 보면 수삼의 세근들이 대다수 제거되지 않고 남아있어서 출력되는 픽셀개수가 상대적으로 크게 나온다. 수삼에 흙이나 이물질이 첨가된 경우의 원본 이미지(Fig. 6c)와 필터이미지(Fig. 6d)를 보면 수삼 몸체에서 흙이나 이물질로 어둡게 나타난 부분이 필터후 제거되어 출력되는 픽셀개수가 상대적으로 작게 나온다.

4. 결 론

본 연구에서는 수삼등급 판정기준의 하나인 중량정보를 영상 분석만으로 근접하게 유추하기 위하여 영상 이진화 임계값과 블러필터의 최적화된 설정값을 찾는 방법으로 중량과의 상관관계 확인하였다. 이 과정에서 수삼의 중량과 필터된 영상에서의 픽셀개수와의 관계가 선형성을 가지는 높은 상관관계(상관계수 0.9162)를 확인할 수 있었다. 수삼은 다양한 직경과 비중을 가지는 뇌두, 동체, 지근, 세근 등이 복잡하게 결합되어 시료의 상태에 따른 평면 영상만으로 수삼의 중량을 추정한다는 것이 어려움이 있으나 특이값을 보이는 수삼에 대해 따로 분류하고 분석하였고 수삼의 세근이 많고 밀집된 분포를 가지는 경우와 흙이나 이물질로 오염된 경우들로 확인하였다. 향후 이러한 문제를 보완하여 몸체와 세근을 좀더 세밀하게 분류하여 동체부분과 지근 및 세근부분의 면적에 가중치를 부과하는 형태로 필터를 개선하면 결정계수를 좀더 높일 수 있으며, 영상분석을 이용한 수삼의 중량을 추정할 값으로 수삼등급을 자동분류에 하나의 신뢰성 있는 파라미터로 사용이 가능할 것이다.

References

[1] C. M. Jeong and J. S. Shin, "Comparison of Grade of Raw and Red Ginseng on each Factor of Quality in Korean and American Ginseng" *Korean Journal of Medicinal Crop Science*, Vol.14, No.4, pp.229-233, 2006.

[2] J. Y. Kang, M. G. Lee, and Y. T. Kim, "Automatic Decision-Making on the Grade of 6-Year-Old Fresh Ginseng (*Panax ginseng* C.A. Meyer) by an Image Analyzer 1. Shape and Weight Analyses according to the Grade of Fresh Ginseng," *Journal of Ginseng Research*, Vol.20, Issue 1, pp.65-71, 1996.

[3] J. Y. Kang and M. G. Lee, "Automatic Decision-Making on the Grade of 6-Year-Old Fresh Ginseng (*Panax ginseng* C.A. Meyer) by an Image Analyzer 1. Shape and Weight Analyses according to the Grade of Fresh Ginseng," *Journal of Ginseng Research*, Vol.26, Issue 1, pp.6-9, 2002.



정 석 훈

e-mail : stellong2@gmail.com
 2002년 명지대학교 전자공학과(학사)
 2004년 명지대학교 전자공학과(석사)
 2015년~현 재 선문대학교 정보통신공학과 박사과정
 관심분야: 의료신호처리, 패턴인식, 영상처리



고 국 원

e-mail : kuks2309@sunmoon.ac.kr
 1992년 서울대학교 농업기계학과(학사)
 1994년 KAIST 정밀공학과(석사)
 2001년 KAIST 기계공학(박사)
 2003년~현 재 선문대학교 기계ICT융합 공학부 부교수
 관심분야: 3차원 측정시스템, 영상처리, 인공지능



이 지 연

e-mail : jylee2930@gmail.com
 1993년 이화여자대학교 통계학과(학사)
 2010년 인제대학교 교육대학원 수학교육(석사)
 2014년 인제대학교 컴퓨터응용과학(박사)
 2014년~현 재 선문대학교 자동화에너지 연구소 연구원

관심분야: 수치해석, 영상처리, 통계처리



이 진 호

e-mail : jhhunter7@kgc.co.kr
 1996년 충북대학교 농학과(학사)
 1996년 KT&G 입사
 2010년~현 재 KGC인삼공사 과장
 관심분야: 인삼 및 생약재 재배관리



서 현 석

e-mail : 20109115@kgc.co.kr
2006년 충북대학교 연초학과(학사)
2007년 KT&G 입사
2010년~현 재 KGC인삼공사 대리
관심분야: 인삼재배



이 상 준

e-mail : mcp94lee@sunmoon.ac.kr
2001년 명지대학교 제어계측공학과(학사)
2005년 명지대학교 전자공학과(석사)
2011년 연세대학교 전기전자공학부(박사)
2014년~현 재 선문대학교 기계ICT융합
공학부 조교수

관심분야: 임베디드시스템, 의료신호처리, 패턴인식, 영상처리