

대두의 광학적 선별장치 개발을 위한 선별 인자 구명*

Identification of Discrimination Factors for Development of Optical Soybean Sorter

노 상 하*	김 현 릉*	황 인 근*
정회원		정회원
S. H. Noh	H. Y. Kim	I. K. Hwang

ABSTRACT

Spectroscopic analysis of soybean kernels were made in the wavelength range of 400 to 1100 nm to find effective discrimination factors which are required for developing an optical soybean sorter. Soybean samples used for the test were the sound and five classes of the defective kernels such as the immature, discolored(brown and violet), damaged by insect and diseased. Effective discrimination factors to classify the soybean kernels into the sound and the defective were found to be R_{640} , R_{580}/R_{990} , $R_{600}-R_{820}$ and $(R_{590}-R_{820})/R_{990}$ with classification error of less than 4%. Mahalanobis distance was used as a criterion to select significant wavelengths involved in the discrimination factors.

주요 용어(Key Words): 콩(Soybean Seed), 분광반사(Spectral Reflectance), 판별인자(Discrimination Factor)

1. 서 론

콩의 소비량은 해마다 꾸준히 증가하여 96년 현재 국내 소비량은 163만 톤으로 쌀 소비량의 25%에 이르고 있다. 그러나 경작면적의 감소로 인해 국내 생산량은 16만 톤으로 자급률이 10% 정도에 불과하며 거의 대부분을 수입에 의존하고 있다. 콩의 수입 의존도가 높은 이유는 콩의 낮은 수익성과 높은 생산비로 인해 수입콩에 비해 가격 경쟁력이 떨어지기 때문이다. 따라서 최소한의 자급도를 유지하고 국제 경쟁력을 향상시키기 위해서는 기계화를 통한 생산비 절감이 필요하다. 특히 국내에서 생산되는 콩은 수입콩에 비해 품질이 우수하다는 점을 고려할 때,

품질이 우수한 국내콩을 값싼 수입콩과 차별화 함으로써 국제 경쟁력을 향상시키기 위한 수확 후 선별 작업의 기계화가 절실히 요구되고 있다. 그러나 현행 수확 후 선별작업은 단순한 크기 선별을 제외하면 전적으로 인력에 의존하고 있는 실정이다.

미국을 비롯한 선진국에서는 10여년 전부터 영상 처리와 분광반사 특성을 이용한 콩의 선별에 관한 연구가 진행되어 왔다. Paulsen 등(1989)은 칼라 영상 처리를 이용하여 건전립과 피해립을 판별하는 알고리즘을 개발한 결과 98%의 정확도로 3초만에 판별이 가능하며, 피해립 내에서 피해립의 종류를 판별하는 데에는 99%의 정확도로 25초가 소요된다고 보고하였다. Casady 등(1992)은 건전립, 미숙립 및 각

* 이 연구는 한국학술진흥재단에서 지원한 "농산물의 광학적 특성 구명 및 선별센서 개발을 위한 기초연구"로 수행되었음.

* 서울대학교 농업생명과학대학 생물자원공학부

중 피해를 판별하기 위해서 칼라 영상처리를 이용하였는데, 24개의 대표적인 시료에 대해 물체 영상의 RGB로부터 구한 r 및 g 값을 사용한 결과 판별 정확도는 93.9%, 판별시간은 25초 정도 소요된다고 보고하였다. 또한 이들은 판별에 유의한 파장을 선정하여 칼라영상 시스템의 기초 자료로 사용하고자 300~850nm 영역에서 건전립, 미숙립, 각종 피해를 분광반사 특성을 조사한 바 있다.

본 연구는 400~1100nm 파장대역에서 건전립, 미숙립, 갈색립, 충해립, 병해립 등 각종 대두 시료의 분광반사 특성을 분석하여 광센서를 이용한 대두용 선별장치의 선별센서부 개발에 필요한 광학적 판별인자를 제시하고자 하였다.

2. 재료 및 방법

가. 공시재료

96년 전남 해남에서 생산된 대두 중 1차 크기 선별을 거친 것을 경기도 정남농협에서 구입하여 실험 재료로 사용하였다. 콩은 전문가에 의해 건전립, 미숙립, 갈색립, 자색립, 충해립, 병해립의 6종류로 분류하였으며, 판별분석용 시료와 판별분석에서 선정된 판별인자의 판별정확도를 검증하기 위해 시료 종류당 각각 30알을 사용하였다.

나. 분광반사 스펙트럼 측정

시료의 분광반사 스펙트럼은 서울대학교 농업과학공동기기센터에 있는 분광광도계(Model 6500, NIRSystems Inc., USA)를 사용하여 400~1100nm의 파장 영역에서 2nm 간격으로 측정하였다. 단립 상태로 콩의 분광반사 스펙트럼을 측정하기 위하여 그림 1에서 보는 바와 같은 샘플홀더를 제작하였다. 무광택 흑색 페인트로 도포(塗布)된 두께 0.5mm의 스테인리스 철판에 지름 6mm의 구멍을 뚫어 입사광이 조사되도록 하였다. 이 때, 구멍 주위에 흑색 고무스펀지를 부착하여 반사광의 누광과 불필요한 광의 유입을 막았으며 클램프로 콩을 고정시킨 상태에서 스

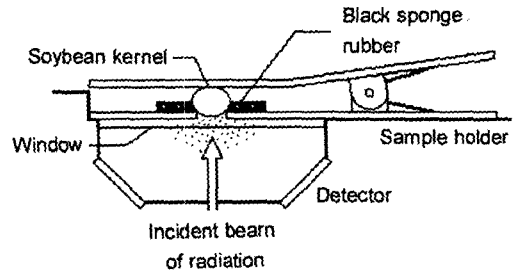


Fig. 1 Sample holder for reflectance measurements on a single kernel.

펙트럼을 측정하였다.

샘플홀더 내에 콩을 놓는 방향과 위치를 달리하면서 스펙트럼의 재현성(reproducibility)을 시험하였다. 건전립 시료집단에서 임의의 한 알을 선택하여 반사 스펙트럼을 30회 반복 측정한 다음 각 파장에서 반사도의 표준편차의 평균을 조사한 결과 건전립 시료 전체(30알)의 평균 표준편차의 1/4 정도였다. 따라서 샘플홀더에 의한 시료의 자세 변화는 무시할 수 있을 것으로 판단되었다.

다. 판별분석 방법

1) 판별인자

본 연구에서는 장치구현의 간소화를 위해 판별인자로 단일파장에서의 반사도(R_{λ}), 두 파장에서의 반사도의 차($R_{\lambda_1} - R_{\lambda_2}$) 및 비($R_{\lambda_1}/R_{\lambda_2}$), 세 파장에서의 반사도의 정규화된 차(normalized difference) ($(R_{\lambda_1} - R_{\lambda_2})/R_{\lambda_3}$)로 제한하였다. 반사도의 차는 스펙트럼의 모양을 강조하고 기준선(baseline)의 이동을 제거하는 장점이 있으며, 비는 무차원수로서 광학 선별시스템에서 조명의 영향을 감소시키는 이점이 있다. 판별 분석은 콩을 건전립과 불량립으로 구분하기 위해 수행하였으며 불량립끼리의 판별 분석은 하지 않았다.

2) 파장 선택 방법

판별 분석은 미지의 개체에 클래스(class)를 부여하기 위해 판별인자 $x^T = [x_1, x_2, \dots, x_n]$ 가 다변량 정규 분포를 따른다는 가정하에 미지의 개체와 각

클래스 평균과의 마할라노비스 거리(Mahalanobis distance)를 계산하여 가장 가까운 클래스에 귀속시키는 것이다. 미지의 개체와 i 번째 클래스 평균과의 마할라노비스 거리 D^2 은 식 1과 같이 정의된다.

$$D^2 = (x - \bar{x}_i)^T \Sigma_i^{-1} (x - \bar{x}_i) \dots\dots\dots (1)$$

여기서,

- x = 미지의 개체 x 의 위치 벡터
- \bar{x}_i = i 번째 클래스의 평균의 위치 벡터
- Σ_i = i 번째 클래스의 분산-공분산 행렬

판별인자에 포함된 파장은 원칙적으로 다음 식 (2)의 값을 최소로 하는 파장으로 선택하였으나, 단순한 수치적인 결과에 의한 파장의 선택을 피하고 실용화를 고려한 보다 안정적인 파장을 선택하기 위해 그래프와 분포도를 그려 유의한 파장을 선택하였다.

$$\text{Criterion}(\lambda) = \sum_{i=1}^P \sum_{j \neq i} \frac{1}{D_{ij}^2(\lambda)} \dots\dots\dots (2)$$

여기서,

$D_{ij}^2(\lambda)$ = 특정파장 λ 에서 i 번째와 j 번째 클래스 평균간의 마할라노비스 거리

P = 판별하고자 하는 클래스의 수

3. 결과 및 고찰

가. 시료별 반사도 스펙트럼의 특성

건전립과 불량립 즉, 미숙립, 갈색립, 자색립, 충해립, 병해립 시료집단의 400~1100nm 파장영역에서의 평균 반사도 스펙트럼은 그림 2에서 보는 바와 같다. 건전립의 반사도는 가시광선 영역에서 불량립보다 높게 나타났다. 미숙립은 엽록소에 의한 흡광으로 반사도가 급격히 떨어지는 670nm 부근이, 미숙립을 제외한 나머지 불량립은 500~700nm의 비교적 넓은 파장영역이 판별에 적합함을 알 수 있다.

나. 판별분석 결과

1) 반사도 $R_{\lambda i}$ 를 이용한 판별 분석

건전립과 불량립을 판별하는 데 가장 유의한 단일 파장을 선정하기 위해서 각 파장별 Criterion을 구한 결과 그림 2에 표시한 바와 같이 나타났다. 그림에서 알 수 있듯이 Criterion은 580~650nm 파장영역에서 작게 나타났으며, 분석 결과 640nm에서 판별정도

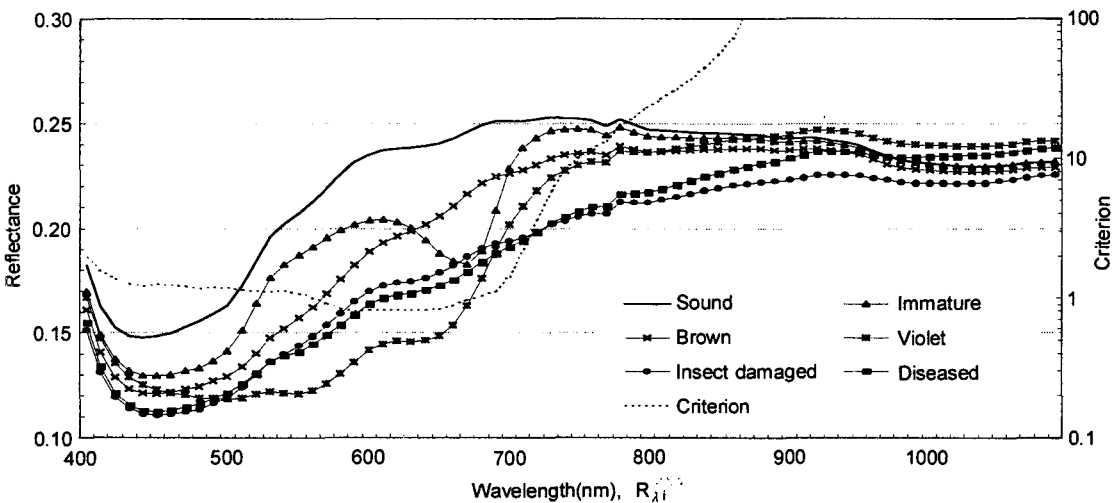


Fig. 2 Mean reflectances of the sound and the defective sample sets and the Criterion(λ) computed with $R_{\lambda i}$

가 가장 우수한 것으로 판명되었다.

2) 반사도의 차 ($R_{\lambda_i} - R_{\lambda_j}$)를 이용한 판별 분석

판별에 가장 유의한 인자를 선정하기 위해 먼저 두 파장 λ_i, λ_j 의 모든 조합에 대해 $\text{Criterion}(\lambda_i, \lambda_j)$ 을 구하고, 그 역수의 분포를 조사하였다(그림 3). Criterion 값은 범위가 커서 편의상 역수인 $1/\text{Criterion}$ 을 사용하였으며, 분포도에서 X축은 λ_i 를, Y축은 λ_j 를 의미한다. 분포도는 $Y = X$ 의 직선에 대해서 대칭이며 $1/\text{Criterion}$ 이 1 이상인 파장의 조합

은 그림에서 보는 바와 같이 3영역이고 그 중에서도 $(\lambda_i, \lambda_j) = (400, 600)$ 부근에서 최대값을 나타내었다. 그러나 $1/\text{Criterion}$ 이 최대는 아니지만 보다 광범위한 파장 영역에서 비교적 큰 값을 갖는 $(\lambda_i, \lambda_j) = (600, 800)$ 부근의 파장을 잠정적인 유의 파장으로 선택하였다. 이와 같이 판별에 적합한 대략적인 파장의 조합이 결정되면 그림 4와 같은 그래프를 통해 최종적인 파장을 선정하였다. 즉, 그림 3을 통해 $\lambda_i = 600\text{nm}$ 를 잠정적인 첫 번째 파장으로 선정하고 λ_j 에 따른 판별인자 $R_{600} - R_{\lambda_j}$ 에 대한 $1/\text{Criterion}$ 의 변

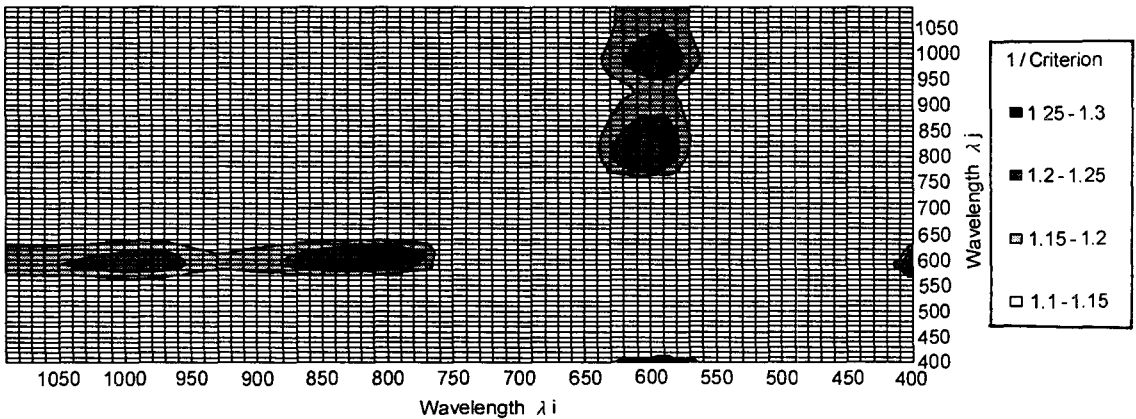


Fig. 3 Distribution of $1/\text{Criterion}(\lambda_i, \lambda_j)$ values computed with the differences of reflectances, $R_{\lambda_i} - R_{\lambda_j}$ depending on the wavelength combination of λ_i and λ_j .

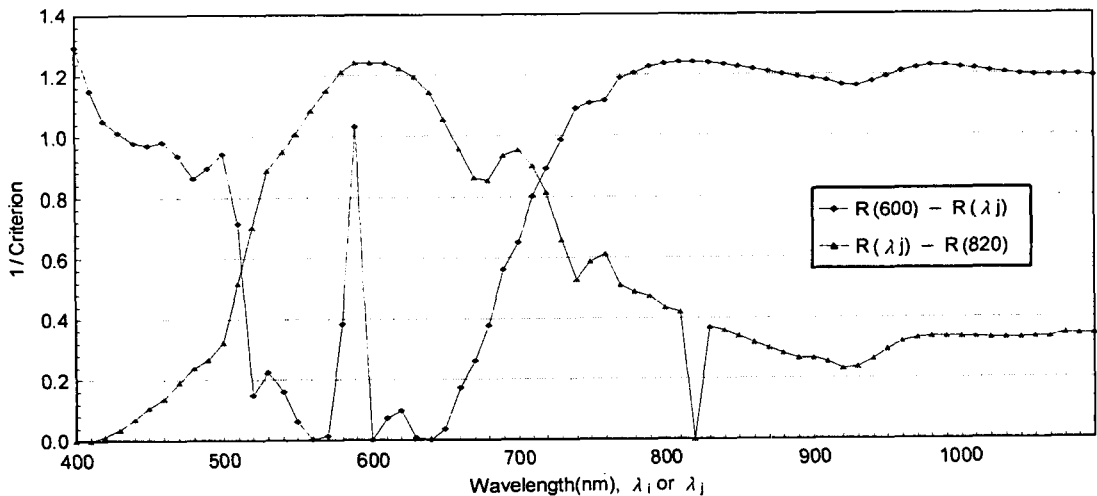


Fig. 4 Variation of $1/\text{Criterion}$ values computed with $R_{600} - R_{\lambda_j}$ and $R_{\lambda_i} - R_{820}$ as the λ_i and λ_j change.

화를 나타내는 그래프를 그려 이 값이 가장 크고, 그 주변값의 변화가 작은 파장인 820nm를 λ_j 로 선정하였다. 다음은 같은 방법으로 판별인자 $R_{\lambda_i} - R_{820}$ 에 대한 $1/Criterion$ 의 그래프를 그려 λ_i 를 결정하였다. 이와 같은 과정을 반복함으로써 수렴하는 특정 두 파장을 최종적인 파장으로 선택하였다. 한편 분석과정에서 선택되는 두 개의 파장이 서로 인접할 경우는 분석에서 제외시켰다. 그 이유는 $R(600) - R(\lambda_j)$ 그래프에서 알 수 있듯 λ_j 가 580nm 일 때에는 주변의 $1/Criterion$ 이 급격히 변하므로 실제 선별센서를 구성할 경우 약간의 피크파장의 변화가 선별결과에 매우 민감한 영향을 미칠 수 있기 때문이다. 그림 3과 4를 근거로 분석한 파장의 조합중에서 600nm와 820nm에서의 반사도의 차를 이용할 경우 $1/Criterion$ 이 가장 크고 안정적이었다.

3) 비 $R_{\lambda_i}/R_{\lambda_j}$ 를 이용한 판별 분석

두 파장에서의 반사도의 비 $R_{\lambda_i}/R_{\lambda_j}$ 에 포함된 유의한 두 파장을 선택하는 방법은 차를 이용한 판별 분석에서 설명한 것과 유사하다. 그림 5는 두 파장의 모든 조합에 대해 $1/Criterion$ 이 1 이상인 영역을 나타낸 것으로, X축은 분모의 파장 λ_j 를, Y축은 분자의 파장 λ_i 를 의미한다. 그림에서 $1/Criterion$ 의

값은 분모의 파장보다는 분자의 파장에 더 민감하게 변하며, 분자의 파장이 600nm 부근일 때 $1/Criterion$ 이 최대값을 나타냄을 알 수 있다.

그림 5에서 대략적으로 결정한 분자의 파장 영역 내에서 λ_i 를 변화시키고, 각 λ_i 에서 분모의 파장 λ_j 의 변화에 따른 $1/Criterion$ 값의 변화를 분석한 결과 판별에 가장 적합한 두 개의 파장은 $\lambda_i=580nm$, $\lambda_j=990nm$ 로 판명되었다. 그림 6은 분자의 파장으로 580nm를 사용 하였을 때 $1/Criterion$ 이 가장 큰 분모의 파장은 990nm이며, 반대로 분모의 파장을 990nm로 하였을 때는 분자의 파장이 580nm 일 때 $1/Criterion$ 이 최대값으로 나타남을 보여주고 있다. 이 분석결과에 의하면 분자의 파장을 580nm로 선택할 경우 분모의 파장은 비교적 광범위한 영역에서 선정 될 수 있음을 알 수 있다.

4) 정규화된 차 $(R_{\lambda_i} - R_{\lambda_j})/R_{\lambda_k}$ 를 이용한 판별 분석

정규화된 반사도의 차 $(R_{\lambda_i} - R_{\lambda_j})/R_{\lambda_k}$ 에 포함되는 세 파장 중에서 λ_i 는 식 3을 통해 잠정적으로 결정하였다. 즉, 주어진 λ_i 에 대해 20nm 간격으로 λ_j 와 λ_k 의 모든 조합에 대해 판별인자 $(R_{\lambda_i} - R_{\lambda_j})/R_{\lambda_k}$ 에 의해 계산되는 $Criterion$ 의 합을 구하고, 이들의

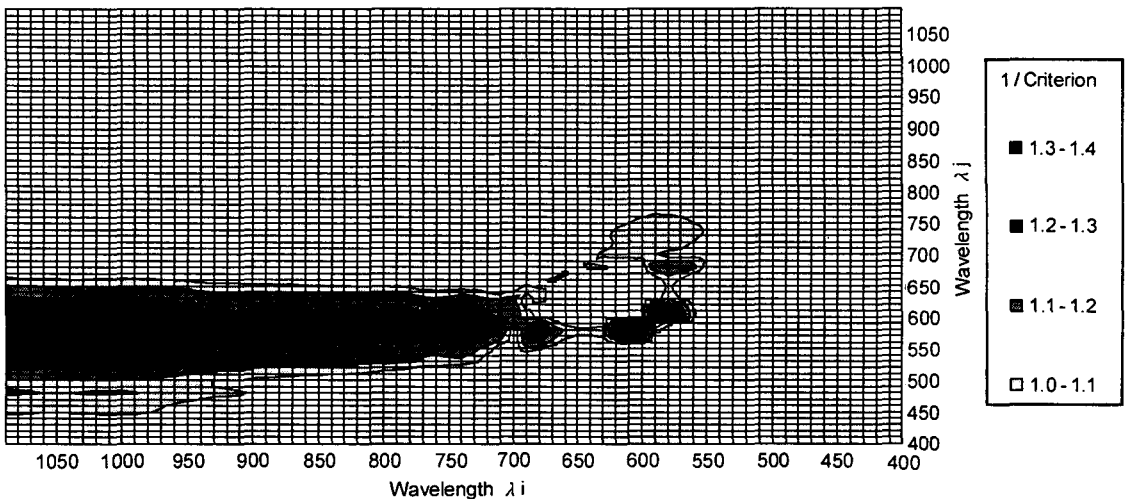


Fig. 5 Distribution of $1/Criterion$ values computed with the ratios of reflectances, $R_{\lambda_i}/R_{\lambda_j}$ depending on the wavelength combination of λ_i and λ_j .

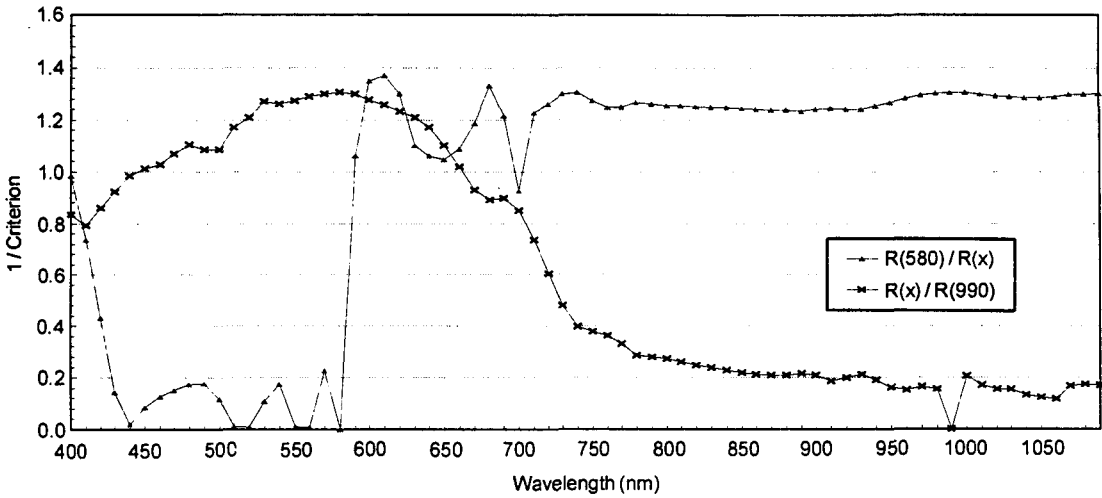


Fig. 6 Change of 1/Criterion values computed with R_{580}/R_{λ_j} and R_{λ_i}/R_{990} as the λ_i and λ_j change.

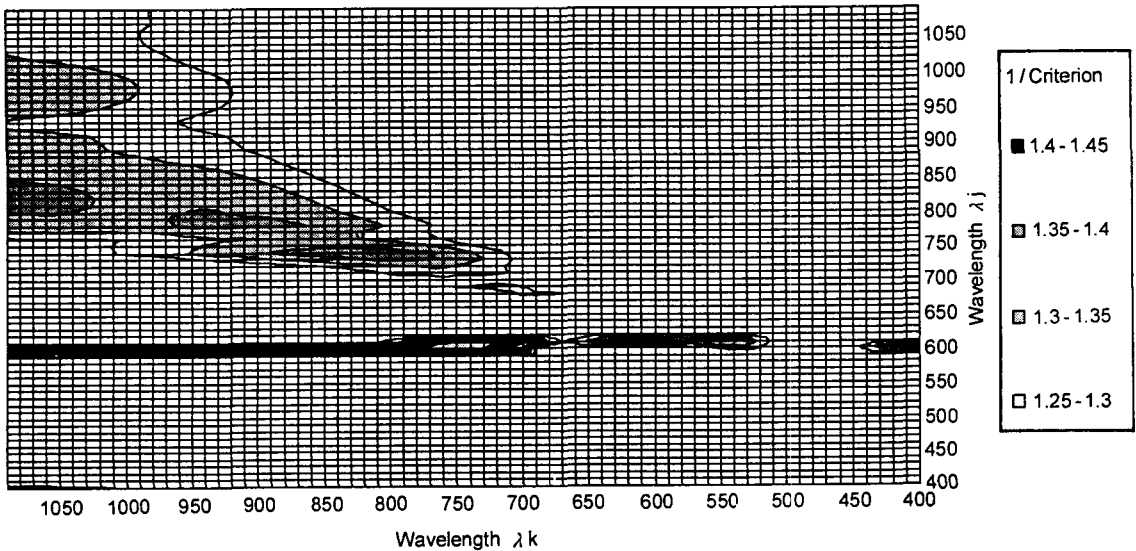


Fig. 7 Distribution of 1/Criterion values computed with the normalized differences of reflectances, $(R_{580} - R_{\lambda_j})/R_{\lambda_k}$ depending on the wavelength combination of λ_i and λ_k .

합이 최소가 되는 파장 λ_i 를 잠정적인 λ_i 로 결정하였다.

$$\min_{\lambda_i} \sum_{\lambda_j} \sum_{\lambda_k} \text{Criterion}(\lambda_i, \lambda_j, \lambda_k) \dots\dots\dots (3)$$

그리고 두 번째 파장 λ_j 와 세 번째 파장 λ_k 는 λ_i 를 고정시킨 상태에서 차를 이용한 판별 분석에서 설명한 것과 같은 방법으로 구한 다음, 이차적으로

이들 λ_j 와 λ_k 를 고정시키고 판별인자 $(R_{\lambda_i} - R_{\lambda_j})/R_{\lambda_k}$ 의 λ_i' 에 따른 Criterion을 관찰하여 가장 적절한 파장 λ_i' 를 구하였다. 이 때 λ_i' 와 λ_i 가 같지 않으면 λ_i' 를 λ_i 로 설정한 다음 λ_j 와 λ_k 를 다시 결정하였다. 이와 같은 과정을 λ_i' 와 λ_i 가 같을 때까지 반복함으로써 최종적인 $\lambda_i, \lambda_j, \lambda_k$ 를 결정하였다.

그림 7은 Criterion의 합이 최소가 되는 $\lambda_i =$

590nm 일 때 λ_j 와 λ_k 의 모든 파장 조합에 대해 1/Criterion의 분포를 나타낸 것이다. 이 결과에 의하면 두 번째 파장 λ_j 가 600nm 부근의 값을 가질 때 1/Criterion은 최대를 나타내지만 이 파장은 이미 선정된 590nm와 서로 인접하기 때문에 분석과정에 제외시키고 나머지 파장중에서 선정할 결과 (λ_j, λ_k) = (820nm, 1090nm) 일 때 1/Criterion이 최대로 나타났다. 일반적으로 광학선별에 사용되는 광센서의 파장별 감도는 가시광영역에서 멀어질수록 급격히 감소하기 때문에 이 파장을 제외할 경우 최종적으로 990nm가 제3의 파장으로 선정되었다. 그림 8은 최종적으로 선택된 세 파장의 조합에 대한 1/Criterion의 분포를 나타낸 것으로, 파장선정을 위한 기초자료로

사용할 수 있을 것이다.

5) 각 판별인자의 판별정확도 검증

이상의 판별분석 결과 선정된 각 판별인자를 이용하여 미지의 대두 시료(종류별 30알)를 대상으로 건전립인지 아닌지를 판정한 결과 각 판별인자 및 시료별 판별오차는 표 1과 같다. 이 표에 의하면 판별인자 R_{640} 의 경우 미숙립, 갈색립, 자색립, 충해립, 병해립 등의 비정상립이 건전립으로 오판되는 경우가 있었으며, 그 외 판별인자의 경우 충해립이 건전립으로 오판되는 비율이 가장 큰 것으로 나타났다. 판별오차, 선별센서의 구성비용 등을 감안할 때 대두 선별용 광학센서는 R_{580}/R_{990} 을 판별인자로 사용하는

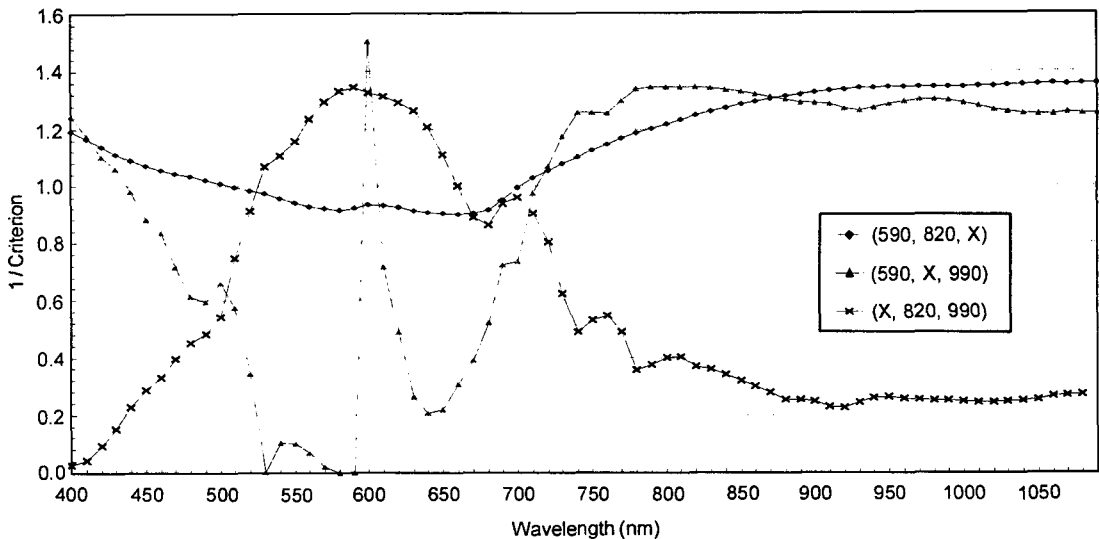


Fig. 8 Change of 1/Criterion values computed with $(R_{590} - R_{\lambda_j})/R_{\lambda_k}$, $(R_{\lambda_i} - R_{820})/R_{\lambda_k}$, $(R_{\lambda_i} - R_{\lambda_j})/R_{990}$ as the λ_j , λ_i and λ_k change.

Table 1 Classification results by each discrimination factor indicating number of kernels misclassified, the sound soybean kernels into the defective and vice versa

Factor	No. of kernels misclassified among 30 kernels/class						Error (%)
	Sound	Immature	Brown	Violet	Insect d.	Diseased	
R_{640}	0	2	1	2	1	1	3.9
$R_{600} - R_{820}$	1	1	0	0	4	0	3.3
R_{580}/R_{990}	0	1	0	0	4	0	2.8
$(R_{590} - R_{820})/R_{990}$	0	1	0	0	4	0	2.8

것이 가장 바람직 할 것으로 판단된다.

4. 요약 및 결론

본 연구는 광센서나 카메라를 이용한 콩의 외부품질 판정시스템의 선별센서부 개발을 위한 기초 연구로서 콩의 분광반사 특성을 이용하여 선별에 유의한 판별인자와 파장을 제시하기 위해서 수행되었다.

1차 정선된 콩을 건전립과 미숙립, 갈색립, 자색립, 충해립 및 병해립으로 분류하고, 이들 시료를 대상으로 400~1100nm 파장영역에서 반사 스펙트럼을 측정된 후 건전립과 비정상립을 판별하기 위해 마할라노비스 거리를 이용하여 판별 분석을 수행하였다. 판별인자로는 단일파장에서의 반사도($R_{\lambda i}$), 두 파장에서의 반사도의 차($R_{\lambda i} - R_{\lambda j}$)와 비($R_{\lambda i}/R_{\lambda j}$), 세 파장에서의 반사도의 정규화된 차($(R_{\lambda i} - R_{\lambda j})/R_{\lambda k}$)를 사용하였다. 본 연구의 결과를 요약하면 다음과 같다.

1) 400~1100nm 파장영역에서 각 시료별 반사스펙트럼을 제시하고, 이를 이용하여 각 판별인자에 포함되는 파장을 선정하는 방법을 제시하였다.

2) 콩을 선별하기 위해 R_{640} , $R_{600} - R_{820}$, R_{580}/R_{990} 및 $(R_{590} - R_{820})/R_{990}$ 을 판별인자로 사용할 경우 선별오차는 4% 미만으로 분석되었으며, 선별센서 구

현의 용이성과 판별오차를 고려할 경우 R_{580}/R_{990} 를 판별인자로 사용하는 것이 가장 바람직한 것으로 판단되었다.

참 고 문 헌

1. 김현룡. 1998. 분광 반사 특성을 이용한 대두의 외관품질 판별인자 구명. 서울대학교 석사학위논문.
2. Casady W. W., M. R. Paulsen, J. F. Reid and J. B. Sinclair. 1992. A trainable algorithm for inspection of soybean seed quality. Trans. of the ASAE 35(6): 2027-2034.
3. Casady W. W., M. R. Paulsen and J. B. Sinclair. 1993. Optical properties of damaged soybean seeds. Trans. of the ASAE 36(3):943-947.
4. Gunasekaran S., T. M. Cooper and A. G. Berlage. 1988. Soybean seed coat and cotyledon crack detection by image processing. J. Agri. Eng. Res. 41:139-148.
5. Paulsen M. R., W. D. Wigger, J. B. Litchfield and J. B. Sinclair. 1989. Computer image analysis for detection of maize and soybean kernel quality factors. J. Agri. Eng. Res. 43:93-101.