

근적외선 분광법을 이용한 콩과 이물질의 판별

임종국 · 강석원* · 이강진 · 모창연 · 손재용
농촌진흥청 국립농업과학원 농업공학부

Identification of Foreign Objects in Soybeans Using Near-infrared Spectroscopy

Jong-Guk Lim, Sukwon Kang*, Kangjin Lee, Changyeon Mo, and Jaeyong Son

National Academy of Agricultural Science, RDA

Abstract

The objective of this research was to classify intact soybeans and foreign objects using near-infrared (NIR) spectroscopy. Intact soybeans and foreign objects were scanned using a NIR spectrometer equipped with scanning monochromator. NIR spectra of intact soybeans and foreign objects in the wavelength range from 900 to 1800 nm were collected. The classification of intact soybeans and foreign objects were conducted by using partial least-square discriminant analysis (PLS-DA) and soft independent modelling of class analogy (SIMCA) multivariate methods. Various types of data pretreatments were tested to develop the classification models. Intact soybeans and foreign objects were successfully classified by the PLS-DA prediction model with mean normalization pretreatment. These results showed the potential of NIR spectroscopy combined with multivariate analysis as a method for classifying intact soybeans and foreign objects.

Key words: NIR, Soybeans, Foreign objects, PLS-DA, SIMCA

서 론

콩은 단백질과 지방함량이 풍부하며 예로부터 우리의 생활에 폭넓게 이용되어 왔으며 최근에는 이소플라본, 올리고당, 사포닌 등과 같은 다양한 기능성 성분이 포함된 건강 식품으로 주목받고 있다(Kim & Kang, 2009). 일반적으로 식용으로서 콩의 용도는 밥을 지을 때 넣어 먹는 밥밑콩, 발아식품인 콩나물, 두부, 두유, 콩기름 등과 같은 가공식품, 된장, 간장, 고추장 등의 발효식품 등으로 분류할 수 있다(Kim et al., 1995). 이러한 식용콩의 소비 증가로 콩 식품 가공업체에서는 양질의 콩을 확보하여 품질을 유지하려고 노력하고 있다.

포장에서 수확된 콩에는 주원료인 콩 이외에 돌, 콩각지, 검불 등의 다양한 이물질이 혼입되어 있으며 깨진 농약병의 파편이나 플라스틱 조각 등도 혼입될 수 있다. Jong et al.(2007)은 시중에 판매되고 있는 브랜드 콩 제품을 수집하여 외관품질을 조사한 결과, 전체 100 개의 제품 중 36 개

의 제품에서 이종 곡립, 돌, 천 등의 이물질이 검출되었다고 보고하였다. 최근 식품에서의 이물질 검출에 대한 관심이 않는 언론보도로 소비자들의 불신이 커져 제조업체의 매출 감소로 이어지는 경우가 많아 이물질 검출에 대한 기술 개발 요구도가 점점 커지고 있다.

콩이나 밀과 같은 곡물류에서 이물질을 선별하는 장치는 이물질의 종류에 따라 다양한 원리의 선별기술이 적용되고 있으며 이 중에서 근적외선 분광특성을 이용한 광학식 선별기는 이물질 선별과 같은 외관품질 뿐만 아니라 내부품질도 동시에 선별할 수 있는 장점이 있어 이와 관련된 장치개발에 관한 연구도 활발해지고 있다(Delwiche, 2003; Armstrong, 2006; Tønning et al., 2009; Dowell et al., 2010).

근적외선 분광법을 이용하여 이물질을 판별할 때 이용되는 방법으로는 부분 최소 제곱 회귀-판별 분석법(Partial Least-Square Discriminant Analysis, PLS-DA)과 SIMCA(Soft Independent Modeling of Class Analogy) 등의 수치적인 모델을 개발하여 미지의 시료를 예측하는 분석방법이 많이 이용되고 있다(Roggo et al., 2003; Kasemsumran et al., 2007; Alexandrakakis et al., 2008).

본 연구에서는 근적외선 스펙트럼 측정장치를 이용하여 정상 콩에 혼입될 가능성이 있는 이물질들을 판별할 수 있는 예측 모델을 개발하고 개발된 모델의 성능과 판별정확도를 검증하고자 하였다.

Corresponding author: Sukwon Kang, Researcher, National Academy of Agricultural Science, RDA, Suwon-city, Gyeonggi-do 441-100, Republic of Korea.

Tel: +82-31-290-1903; Fax: +82-31-290-1900,

E-mail : skang@korea.kr

Received October 18, 2010; revised February 28, 2011; accepted April 18, 2011

재료 및 방법

공시재료

본 연구에 사용된 콩은 2009년 대구에서 생산된 것으로 농촌진흥청 국립식량과학원 두류유지작물과에서 보유하고 있는 유전자원용 콩을 지원받아 사용하였다. 실험에 사용한 품종은 대원콩(Daewonkong), 단백콩(Danbaegkong), 태광콩(Taekwangkong), 풍산나물콩(Pungsannamulkong), 다울콩(Daolkong), 큰울콩(Keunolkong)으로 각 품종별로 10 립씩 총 60 립의 건전한 정상 콩을 선별하였다(Fig. 1). 또한, 콩에 혼입될 가능성이 있는 이물질로는 작은 돌(small stone), 깨진 유리조각(broken glass), 합성수지(synthetic resin)의 3종을 준비하였으며 돌은 콩과 색상이 유사한 베이지색(beige color)의 돌과 다양한 색상(variegated color)이 섞여있는 돌을 각각 10 점씩 준비하였고 깨진 유리 조각도 색상별로 갈색(brown)과 투명(clear)의 두 가지로 각각 10 점씩 만들어 사용하였다(Fig. 2). 구형상의 합성수지는 흰색(white) 계통의 폴리에틸렌(Polyethylene)과 반투명(translucent) 계통의 폴리프로필렌(Polypropylene)을 각각 10 점씩 준비하여 총 60 점의 이물질을 준비하였다.

근적외선 스펙트럼 획득

근적외선 스펙트럼 측정장치는 텅스텐 할로겐램프(DL-250TH, Dongwoo Optron Co., Korea), 단색화장치(MonoRa

150i, Dongwoo Optron Co. Korea), 전원공급부 등으로 구성되었으며 스펙트럼은 외부에서 유입되는 광의 영향을 배제하기 위하여 측정 시료를 빛이 차단된 실험장치 내부의 고정 지그에 부착시켜 반사광의 세기를 측정하여 사용하였다. 여기에 이용된 근적외선의 파장대역은 900 nm에서 1800 nm의 구간으로서 기본 파장간격(gap size)은 10 nm로 설정하여 분석하였다. 정상 콩의 스펙트럼 측정은 종피(seed coat) 부위와 배꼽(hilum) 부위를 각각 1 회씩 측정하여 총 120 개의 스펙트럼을 획득하였고 이물질의 스펙트럼 획득은 최초 측정한 부위와 방향을 반대로 하여 각각 1 회씩 측정하여 총 120 개의 스펙트럼을 획득하였다.

이물질 판별예측모델 개발

이물질의 판별을 위한 예측 모델은 PLS-DA와 SIMCA를 사용하였으며 개발된 모델에 대한 판별 정확도(Correct Classification Rate, %)를 구하였다(Kasemsumran et al., 2007). PLS-DA와 SIMCA 판별 모델 개발을 위해 사용된 전체 240 개의 스펙트럼은 두 그룹으로 분류하였는데 모델 개발용 Calibration group으로 168 개(정상 콩 84 개, 이물질 84 개), 모델 검증용 prediction group으로 나머지 72 개(정상 콩 36 개, 이물질 36 개)의 스펙트럼 데이터를 각 시료별로 무작위로 추출하여 그룹화 하였다.

부분최소제곱회귀(Partial Least Square Regression, PLSR) 모델에서 종속변수(dependent variables)가 실제 측정된 농도

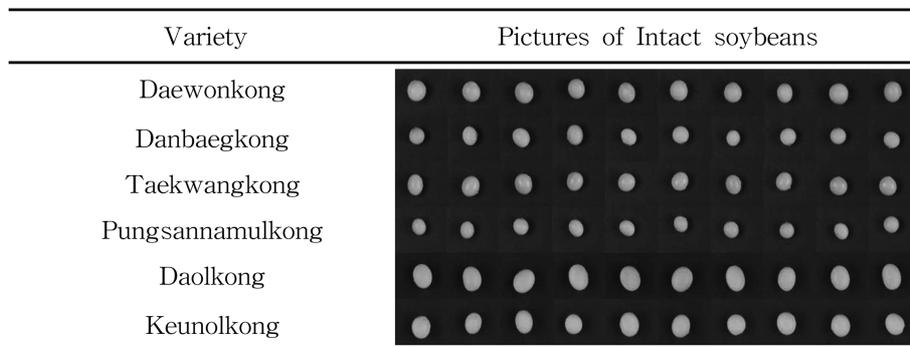


Fig. 1. Pictures of intact soybeans used for an experiment.

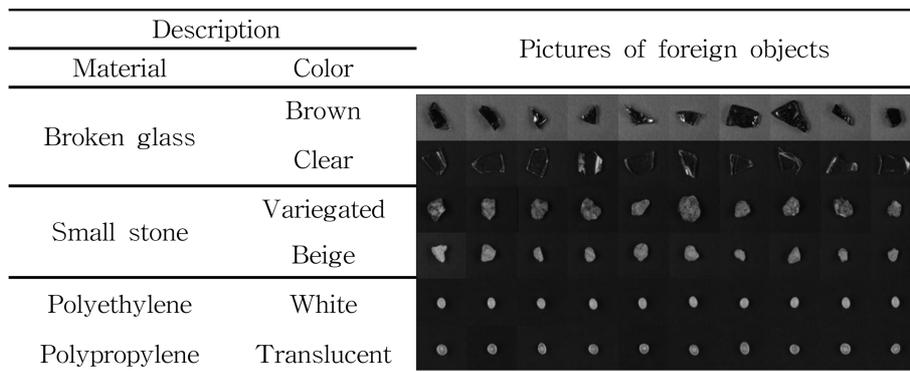


Fig. 2. Pictures of foreign materials used for an experiment.

정보를 이용하여 독립변수(independent variables)인 스펙트럼 데이터와의 선형 회귀모델을 개발하는(Lee et al., 1998; Son et al., 2009) 것과는 달리 PLS-DA는 종속변수로 농도 값 대신에 판별하고자 하는 그룹들을 가상 변수(dummy variables)로 지정하여 회귀 모델을 개발한다(Alexandrakis et al., 2008). 본 연구에서 PLS-DA 모델 개발을 위해 이용된 168 개 스펙트럼의 calibration group에서 84 개의 정상 콩 스펙트럼 'CLASS I'은 임의의 가상 변수 값 '0'으로 하고 나머지 84 개의 이물질에 대한 스펙트럼 'CLASS II'는 임의의 가상 변수 값 '1'로 종속변수를 지정하여 판별 모델을 개발하였으며 미지의 72 개 스펙트럼인 prediction group은 개발된 모델을 적용하여 콩과 이물질 판별을 예측하였다.

SIMCA는 판별하고자 하는 개별 클래스의 특성을 나타내는 주성분(Principal Component, PC)을 주성분분석(Principal Component Analysis, PCA)에 의해 추출하고 개발된 각각의 PCA 모델들은 미지의 새로운 시료와의 유사성을 비교하여 각각의 클래스를 판별하는 패턴 인식 기법이다(Wold, 1976; Maesschalck et al., 1999; Roggo et al., 2003). SIMCA를 적용한 판별 원리는 각 클래스가 선형구조로 모델화되고 한 개의 개체가 고정된 확률 값에서 이미 모델화된 특정 클래스(class)속에 포함될 수 있는 지를 임계거리(critical distance)를 이용하여 결정하는 것이다. 본 연구에서 SIMCA 판별 분석을 하기 위해 168 개의 calibration group에 대한 PCA 모델을 개발한 후 다시 CLASS I과 CLASS II의 개별 PCA 모델들로 추출하여 이 두 클래스 PCA 모델을 prediction group에 적용하여 진행하였다.

원본 스펙트럼 데이터는 불규칙한 이물질 형상에 의한 광산란을 보정하고 우수한 판별예측모델 개발을 위해 다양한 수학적 방법에 의해 전처리(pretreatment)를 적용하게 된다(Hwang, 2000). 본 실험에 사용된 3종의 정규화(Normalization) 방법으로는 스펙트럼의 최대값 정규화(Maximum Normalization), 평균값 정규화(Mean Normalization), 일정 범위값 정규화(Range Normalization)을 이용하였으며, 스펙트럼 측정시 광 산란의 영향을 제거할 목적으로 SNV(Standard Normal Variate)와 MSC(Mean Scattering Correction)를 사용하였다. 또한 스펙트럼의 최소값을 동일하게 보정하기 위한 Baseline과 측정센서 자체의 노이즈(noise)를 제거하기 위해 평활화(Smoothing) 방법을 적용하였으며, 스펙트럼을 측정할 때 변화되는 조명이나 온도 등의 주변 환경요인으로 인한 스펙트럼 기준선의 이동을 보정하기 위해 파장간격을 50 nm로 하여 1차 미분(1st Derivatives)과 2차 미분(2nd Derivatives)의 전처리를 이용하였다.

획득된 스펙트럼을 이용한 모델 개발, 검증 및 데이터의 전처리는 다변량 데이터 분석 소프트웨어(Unscrambler v9.2, Camo co., Norway)를 이용하였다.

결과 및 고찰

정상 콩과 이물질의 분광특성

Fig. 3은 정상 콩과 이물질의 전체 스펙트럼으로서 대략 1050 nm와 1200 nm의 두 파장대역에서 가장 강하게 반사광의 세기가 측정되었으며 1400 nm 영역에서는 모든 시료의 세기가 급격히 감소한 후에 다시 약간 증가하는 경향이 있었는데 이는 센서의 특성에 기인되는 것으로 판단되며 1700 nm 파장대역을 지나서는 반사광 에너지가 소멸되는 것을 알 수 있었다. Fig. 3a-Fig. 3f에서 보듯이 품종이 동일하고 단립의 크기가 비슷한 정상 콩일지라도 반사스펙트럼의 세기가 차이 나는 것을 알 수 있으며, 이는 콩의 단백질, 지방 등과 같은 내부 성분이 스펙트럼에 반영된 결과라고 판단할 수 있다(Lee et al., 2001). Fig. 3g-Fig. 3i에 나타난 바와 같이 이물질에 있어서는 각기 다른 시료의 물성과 형상으로 인해 정상 콩에 비하여 스펙트럼의 세기와 형태가 다양하게 나타났다.

PLS-DA를 이용한 이물질 판별

불규칙하면서도 유사한 경향이 있는 정상 콩과 이물질의 스펙트럼에서 정상 콩과 이물질로 구분할 수 있는 판별예측 모델을 개발하기 위해 먼저 PLS-DA를 적용하였다. 개발된 PLS-DA 모델의 정밀도는 결정계수(Coefficient of Determination, R^2)와 모델작성시의 표준오차(Standard Error of Calibration, SEC), 모델검증을 위한 예측치의 표준오차(Standard Error of Prediction, SEP)로 각각 나타내었으며 판별 정확도는 Fig. 4와 5의 결과에서 보듯이 정상 콩 스펙트럼 CLASS I과 이물질 스펙트럼 CLASS II의 판별기준인 상수 값 '0'과 '1'의 중간 값 0.5를 기준으로 하여 대부분의 예측치의 분포가 나누어지기 때문에 예측결과 값이 0.5보다 낮으면 정상콩에 해당하는 CLASS I로 0.5보다 높으면 이물질에 해당하는 CLASS II로 판단하였다(Alexandrakis et al., 2008). 따라서 정상 콩 스펙트럼과 이물질 스펙트럼의 예측 결과 값이 각각의 판별기준인 0과 1에 근접할수록 PLS-DA 판별 모델의 성능은 우수하다고 볼 수 있다. 정상 콩과 이물질의 스펙트럼 데이터를 다양한 방법으로 전처리하여 개발된 PLS-DA 모델의 결과는 Table 1에 나타난 바와 같다.

아무런 전처리를 하지 않은 스펙트럼 데이터를 이용하여 개발한 판별예측모델에서는(Fig. 4a) 앞에서 설명한 CLASS I과 CLASS II를 모두 충족시켜 100%의 판별율을 보여주었으며 SNV 전처리를 적용한 모델의 경우 정상 콩은 100% 판별하였지만 이물질인 CLASS II의 84 개 예측치 중 2 개가 정상 콩의 판별기준인 0.5 이하로 예측(false positives)되어 97.6%의 판별율을 나타내고 있음을 알 수 있었다(Fig. 4b). 판별예측모델 개발에서 PLS-DA의 결과는 Table 1에서

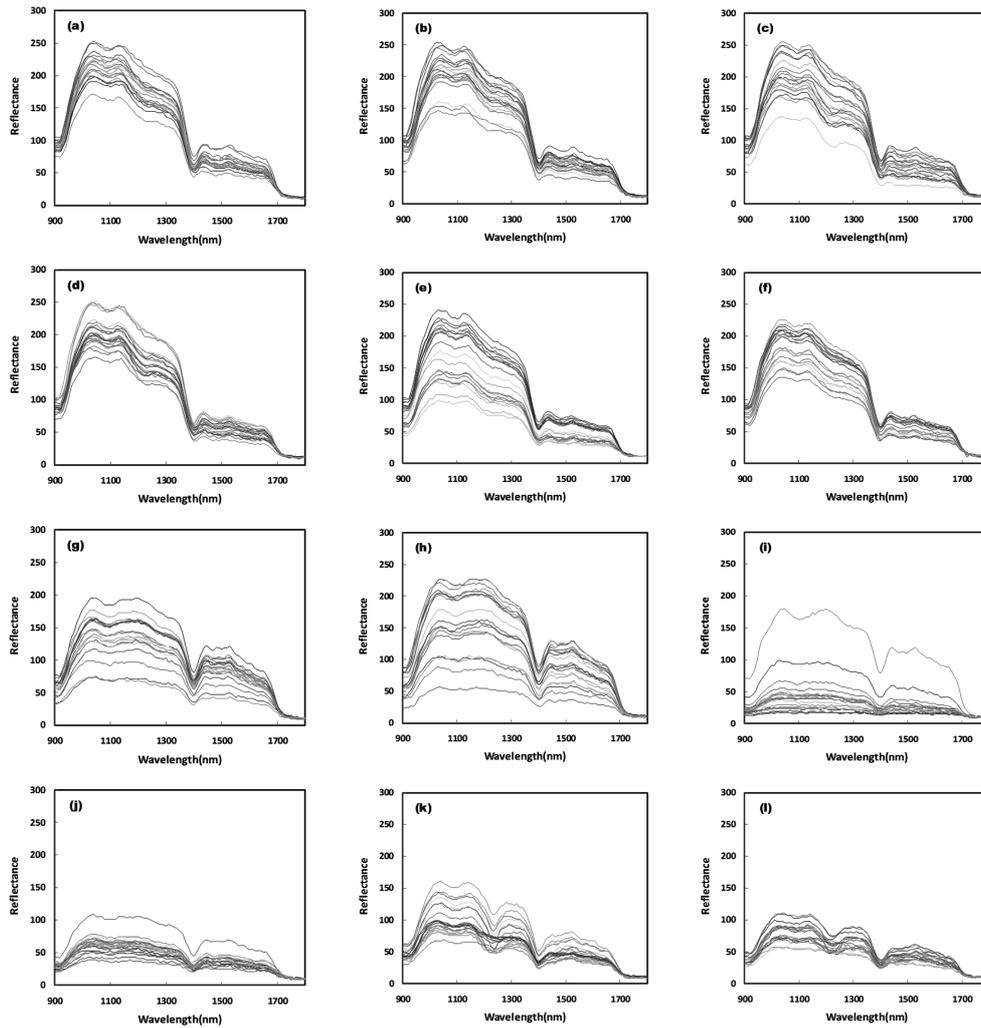


Fig. 3. Original reflectance spectra of Intact soybeans and foreign objects in the wavelength range from 900 to 1800 nm: (a) Daewonkong; (b) Danbaengkong; (c) Taekwangkong; (d) Pungsannamulkong; (e) Daolkong; (f) Keunolkong; (g) Broken glass(brown); (h) Broken glass(clear); (i) Small stone(variegated color); (j) Small stone(beige color); (k) Polyethylene; and (l) Polypropylene.

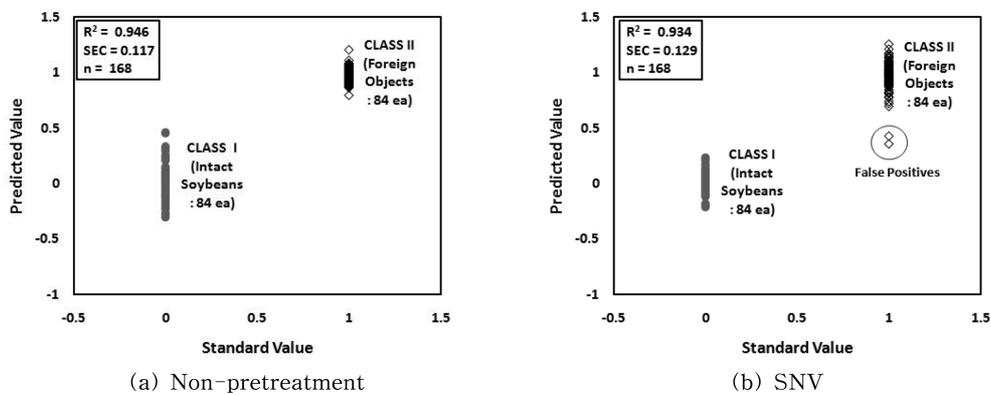


Fig. 4. PLS-DA calibration results for intact soybeans and foreign objects discrimination with non-pretreatment and SNV pretreatment.

Table 1. Calibration and prediction results of intact soybeans and foreign objects using PLS-DA model with various pretreatments in the wavelength range from 900 to 1800 nm.

Data pretreatment	Calibration result (sample : 168 ea)				Prediction result (sample : 72 ea)				PC ⁷⁾
	R ²	SEC	CCR ¹⁾ (%)		R ²	SEP	CCR (%)		
			IS ²⁾	FO ³⁾			IS	FO	
Non-Pretreatment	0.946	0.117	100.0	100.0	0.936	0.124	97.2	97.2	3
Mean	0.963	0.097	100.0	100.0	0.945	0.115	100.0	100.0	4
Nor ⁴⁾ Maximum	0.961	0.100	100.0	100.0	0.946	0.115	100.0	97.2	4
Range	0.967	0.091	100.0	100.0	0.943	0.117	100.0	91.6	4
SNV ⁵⁾	0.934	0.129	100.0	94.4	0.941	0.117	100.0	91.6	3
MSC ⁶⁾	0.961	0.098	100.0	100.0	0.945	0.115	100.0	91.6	5
Baseline	0.947	0.116	100.0	100.0	0.932	0.126	97.2	97.2	3
Smoothing (50 nm gap)	0.945	0.118	100.0	100.0	0.929	0.132	94.4	97.2	3
1st Derivate (50 nm gap)	0.950	0.112	100.0	100.0	0.941	0.120	94.4	100.0	3
2nd Derivate (90 nm gap)	0.938	0.125	100.0	100.0	0.929	0.127	100.0	97.2	4

¹⁾CCR : Correct Classification Rate.

²⁾IS : Intact Soybeans.

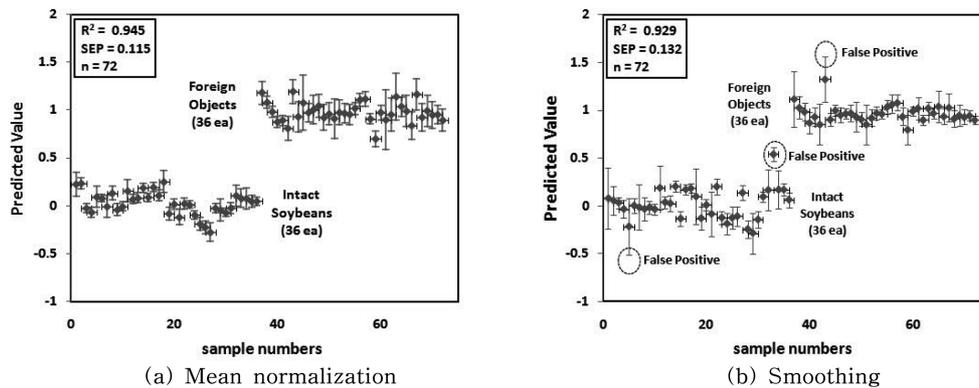
³⁾FO : Foreign Objects.

⁴⁾Nor : Normalization.

⁵⁾SNV : Standard Normal Variate.

⁶⁾MSC : Mean Scattering Correction.

⁷⁾PC : Principal Component

**Fig. 5. PLS-DA prediction results for intact soybeans and foreign objects discrimination with mean normalization and smoothing pretreatment.**

알 수 있듯이 개발된 모든 모델들은 정상 콩에 대해 100%의 판별율을 보였으며 이물질에 대해서도 SNV 전처리 방법을 제외하고는 모두 완벽한 판별 결과를 보여주었다. 하지만 미지 시료에 대한 스펙트럼 데이터인 72 개의 prediction group에 대한 예측 모델의 검증에 있어서는 평균값 정규화 전처리 방법을 적용한 모델만이 정상 콩 36 립과 이물질 36 개를 100% 판별하였으며 최대값 정규화, 일정 범위값 정규화, SNV, MSC, baseline, 2차 미분의 전처리 방법을 적용한 PLS-DA 모델에 대해서는 정상 콩은 100%의 판별율을 나타내었지만 이물질에서는 상대적으로 판별율이 낮게 나타났다.

가장 완벽한 판별율을 나타낸 평균값 정규화 전처리의 PLS-DA 모델의 결정계수(R²)는 0.945, 예측치의 표준오차

(SEP)는 0.115이었다(Fig. 5a). 평활화 전처리된 PLS-DA 모델의 결과는 Fig. 5b에 나타난 바와 같으며, 정상 콩에 대한 예측 값 중 하나가 이물질로 판단되었고 오차를 포함한 예측 값을 적용했을 경우에는 정상 콩에서는 2 개, 이물질에서는 1개가 판별기준 범위를 벗어났음을 알 수 있었다.

Table 1에서 100% 판별율을 보여주었던 평균값 정규화의 전처리 방법을 적용한 PLS-DA 모델에서의 PC 1과 PC 2의 점수(scores)를 나타내면 Fig. 6과 같다. 여기서 Fig. 6a는 calibration group, Fig. 6b는 prediction group 시료에 대한 결과 값으로서 PC 1과 PC 2 두 개의 주성분에 대한 scores plot에서 보듯이 정상 콩과 이물질의 스펙트럼이 PC 1 축의 0을 기준으로 서로 잘 구분되어 군집화되어 있으며 정상 콩 스펙트럼에 해당하는 CLASS I이 이물질 스펙트럼의

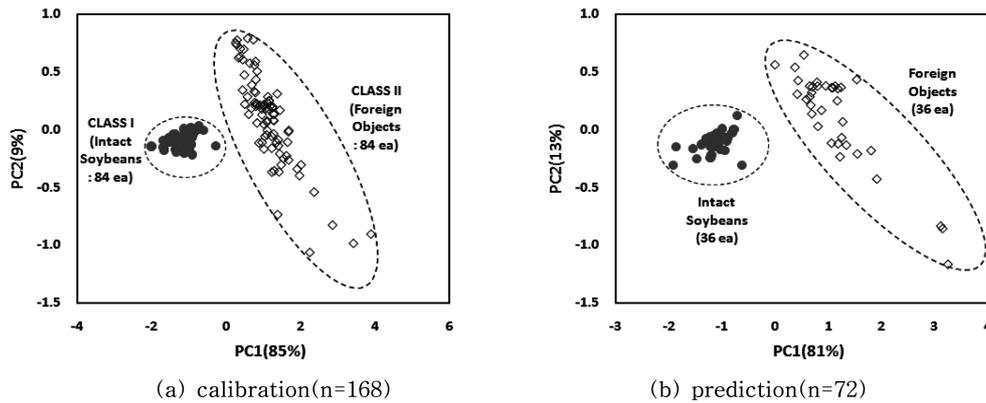


Fig. 6. PLS-DA scores plots(PC1 × PC2) for intact soybeans and foreign objects discrimination with mean normalization pretreatment.

CLASS II에 비하여 더 밀집되어 있었으며 이러한 scores plot을 이용하면 정상 콩과 이물질의 구분이 가능함을 알 수 있었다.

SIMCA를 이용한 이물질 판별

정상 콩 스펙트럼과 이물질 스펙트럼의 168 개를 활용하여 SIMCA 모델을 개발하였으며 미지의 스펙트럼 72 개에 대해서 모델을 검증하였다. 모델 개발용 스펙트럼 168 개 중에서 84 립의 정상 콩 스펙트럼 CLASS I에 대한 PCA 모델과 84 개의 이물질 스펙트럼 CLASS II에 대한 PCA 모델을 전처리 방법별로 적용하였으며 검증도 동일한 전처리를 적용하여 판별 정확도를 확인하였다. SIMCA를 이용한 판별 결과는 Table 2에 나타낸 바와 같으며 전체적인 판별 정확도는 PLS-DA과 비교할 때 낮게 나타났다. PLS-DA에서 100% 판별율을 나타냈던 평균값 정규화를 적용한

전처리는 SIMCA에서는 93.1%의 판별율을 나타내었으며 최대값 정규화 및 일정 범위값 정규화 전처리 방법을 적용한 모델이 정상 콩과 이물질의 평균 판별율에서 94.4%로 가장 양호한 결과를 나타내었다. 전체적으로 양호한 판별율을 보여주었던 PLS-DA와는 달리 SIMCA의 판별결과는 어떠한 전처리에서도 완벽한 판별율을 얻지 못함을 알 수 있었으며, 현재까지 결과로는 이물질의 판별에는 PLS-DA의 방법이 더 양호한 것으로 판단되었다. 이는 SIMCA 모델의 예측 판별 값이 콩과 이물질에 대해 개발된 각각의 선형화된 모델들이 공간상에서 서로 중첩되는 부분이 발생하여 판별율이 저조하였으며 이에 비해 PLS-DA 모델로 개발된 예측 모델은 콩과 이물질의 공간상에서 중첩되는 부분이 존재하지 않았기에 판별율이 높게 나타났다.

요 약

본 연구에서는 정상 콩과 이에 혼입되는 이물질을 판별하기 위해 900 nm에서 1800 nm의 파장대역에서 단색화장치가 장착된 근적외선 분광장치를 이용하여 획득된 콩과 이물질의 반사 스펙트럼의 세기를 이용하여 각각의 판별예측모델을 개발하고 그 성능과 판별정확도를 검증해보았다. 정상 콩 60 립과 이물질 60 점을 각각 2 회 반복하여 측정된 총 240 개의 반사스펙트럼에 대해서 모델 개발용인 calibration group으로 168 개를, 나머지 72 개는 개발된 모델을 예측하는 prediction group으로 나누어 사용하였다. 획득된 스펙트럼은 광원의 불안정함, 시료의 크기와 형태에서 기인되는 여러 변이들을 최소화하기 위해 다양한 수학적 전처리를 적용하였으며 판별예측모델의 개발을 위해 PLS-DA와 SIMCA 방법을 사용하여 모델의 예측 성능과 판별율을 검토하였다.

PLS-DA에서 모델 개발에 사용된 84 개의 정상 콩 스펙트럼 CLASS I은 적용된 모든 전처리에서 100%의 판별율을 보여주었으며 이물질 스펙트럼 CLASS II에서도 SNV 전처리를 제외하고는 모두 100% 이물질로 판별하여 분류

Table 2. SIMCA classification results with various pretreatments in the wavelength range from 900 to 1800 nm.

Data pretreatment	SIMCA classification results (CCR ¹⁾ %		
	IS ²⁾	FO ³⁾	Average
Non-Pretreatment	91.7	88.9	90.3
Mean	94.4	91.7	93.1
Nor ⁴⁾ Maximum	97.2	91.7	94.5
Range	97.2	91.7	94.5
Standard Normal Variate	94.4	91.7	93.1
Mean Scattering Correction	94.4	91.7	93.1
Baseline	91.7	83.37	87.5
Smoothing (50 nm gap)	96.1	88.9	87.5
1st Derivate (50 nm gap)	97.2	86.1	91.7
2nd Derivate (90 nm gap)	97.2	88.9	93.1

¹⁾CCR : Correct classification rate.

²⁾IS : Intact soybeans.

³⁾FO : Foreign objects.

⁴⁾Nor : Normalization.

하였다. 개발된 PLS-DA의 모델에 대한 prediction group의 검증에 있어서는 평균값 정규화 전처리 방법이 정상 콩과 이물질에서 100% 판별율을 보여주었다.

SIMCA를 이용한 이물질 판별예측모델 개발은 PLS-DA와 비교할 때 상대적으로 저조한 판별율 결과를 나타냈으며 최대값 정규화와 일정 범위값 정규화의 전처리 방법을 적용한 모델이 평균 판별율 94.4%로 다소 양호한 결과를 보여주었다.

따라서 콩에 혼입되어 있는 이물질을 판별하는 시스템을 개발하는 데 있어서 근적외선 분광장치를 이용하여 획득한 반사도 스펙트럼은 PLS-DA로 판별예측모델을 개발하고 최적의 전처리 방법을 적용한다면 콩과 이물질의 선별시에 보다 나은 판별율을 얻을 수 있을 것이다.

참고문헌

- Alexandrakis D, Downey G Scannell AGM. 2008. Detection and identification of bacteria in an isolated system with near-infrared spectroscopy and multivariate analysis. *J. Agric. Food Chem.* 56: 3431-3437.
- Armstrong PR. 2006. Rapid single-kernel NIR measurement of grain and oil-seed attributes. *Appl. Eng. Agric.* 22(5): 767-772.
- Delwiche SR. 2003. Classification of scab-and other mold-damaged wheat kernels by near-infrared reflectance spectroscopy. *Trans. ASAE.* 46: 731-738.
- Dowell FE, Maghirang EB, Jayaraman V. 2010. Measurement grain and insect characteristics using NIR laser array technology. *Appl. Eng. Agric.* 26: 165-169.
- Hwang IG. 2000. Development of on-line apple(Fuji) sorting system by the soluble solid and acid contents using VIS/NIR spectroscopy. Ph. D. Thesis. Seoul National Usiniversity.
- Jong SK, Woo SH, Kim HS. 2007. Evaluation of external quality of brand soybeans. *Korean J. Crop Sci.* 52(3): 239-248.
- Kasemsumran S, Thanapase W, Kiatsoonthon A. 2007. Feasibility of near-infrared spectroscopy to detect and to quantify adulterants in cow milk. *Anal. Sci.* 23: 907-910.
- Kim JI, Kang MJ. 2009. Soybean health benefits of application in food. *Korean Soybean Digest.* 26: 1-9.
- Kim YH, Kim SD, Hong EH. 1995. Present status and perspectives of soybean breeding program for high seed quality in korea. *Korean Soybean Digest.* 15: 1-20.
- Lee HB, Choi BR, Kang CS, Kim YH and Choi YJ. 2001. Determination of seed protein and oil concentration in kidney bean by near infrared spectroscopic analysis. *Korean J. Crop Sci.* 46: 248-252.
- Lee KJ, Hruschka WR, Abbott JA, Noh SH, Park BS. 1998. Predicting the soluble solids of apples by near infrared spectroscopy(II) - PLS and ANN models. *J. Korean Soc. Agr. Mach.* 23: 571-582.
- Maesschalck RD, Candolfi A, Massart DL and Heuerding S. 1999. Decision criteria for soft independent modeling of class analogy applied to near infrared data. *Chemometr. Intell. Lab.* 47: 65-77.
- Roggo Y, Duponchel L, Ruckebusch C, Huvenne JP. 2003. Statistical tests for comparison of quantitative and qualitative models developed with near infrared spectra data. *J. Mol. Struct.* 654: 253-262.
- Son JR, Lee KJ, Kang SW, Yang GM, Seo YW. 2009. Development of prediction model for sugar content of strawberry using NIR spectroscopy. *Food Eng. Prog.* 13: 297-301.
- Tonning E, Thybo AK, Pedersen L, Munck L, Hansen A, Togersen FA, Engelsen SB and Norgaard L. 2009. Bulk functionality diversification by unsupervised single-kernel near-infrared(SKNIR) sorting of wheat. *Cereal Chem.* 86: 706-713.
- Wold S. 1976. Pattern recognition by means of disjoint principal components models. *Pattern Recog.* 8: 127-139.