

INSTYTUT AGROFIZYKI  
IM. BOHDANA DOBRZAŃSKIEGO  
POLSKIEJ AKADEMII NAUK

Anna Siedliska

**OKREŚLENIE WYBRANYCH CECH  
JAKOŚCIOWYCH OWOCÓW W OPARCIU  
O ANALIZĘ HIPERSPEKTRALNĄ ORAZ METODY  
SZTUCZNEJ INTELIGENCJI**

DETERMINATION OF SELECTED QUALITY FEATURES OF FRUITS BASED  
ON HYPERSPECTRAL ANALYSIS AND METHODS OF ARTIFICIAL INTELLIGENCE

Rozprawa doktorska

Doctoral thesis

Rozprawa doktorska przygotowana pod kierunkiem

Promotora: dra hab. Piotra Baranowskiego, prof. IA PAN

oraz

Promotora pomocniczego: dr Moniki Zubik

© Anna Siedliska

Lublin 2018

## Streszczenie

Wysoka jakość owoców to jeden z podstawowych warunków odniesienia sukcesu na rynku krajowym i międzynarodowym. Jest ona określana na podstawie analizy wartości odżywczych związanych ze składem chemicznym (np. zawartość cukrów, witamin, minerałów, aminokwasów, enzymów, przeciwutleniaczy), właściwości sensorycznych (tj. wygląd zewnętrzny, smak, aromat, tekstura), czy występowania defektów wewnętrznych i zewnętrznych.

Obrazowanie hiperspektralne jest dynamicznie rozwijającą się techniką rejestracji obrazu w szerokim zakresie spektrum, z jednoczesną rejestracją charakterystyk spektralnych dla poszczególnych pikseli obrazu. Istotnymi cechami techniki obrazowań spektralnych, wskazującymi na jej przydatność do określania cech jakościowych próbek heterogenicznych takich jak owoce, są bezinwazyjność, efektywność analizy, szeroki zakres badanego spektrum (od ultrafioletu po podczerwień) oraz olbrzymia ilość informacji uzyskanych z każdego zobrazowania (rozdzielczość widmowa wynosi kilka nanometrów). Powyższe cechy sprawiają, że systemy obrazowania hiperspektralnego zyskują coraz większe uznanie jako szybka metoda oceny bezpieczeństwa i jakości szerokiej gamy produktów rolnych, w tym również owoców.

W niniejszej pracy wykazano przydatność metody obrazowania hiperspektralnego oraz metod sztucznej inteligencji do określania wybranych cech jakościowych owoców. W zależności od konkretnego zastosowania, wykorzystano zakresy spektralne VNIR i SWIR oraz tryb odbiciowy lub transmisyjny pozyskiwania danych spektralnych. Przeanalizowano różne metody wstępnego przetwarzania danych hiperspektralnych w celu wyselekcjonowania charakterystycznych pasm absorpcyjnych lub transmisyjnych, pozwalających klasyfikować badane cechy jakościowe.

W badaniach przetestowano również kilkanaście metod eksploracji danych, oszacowując ich przydatność do klasyfikacji takich cech jakościowych, jak: różnice odmianowe, zawartość ekstraktu, występowanie uszkodzeń mechanicznych, niejednorodności tkanek oraz choroby wywołane przez patogeny grzybowe. Zastosowane klasyfikatory zbudowane zostały przy wykorzystaniu niewielkiej (optymalnej) liczby długości fali, dla których charakterystyki spektralne były najlepiej skorelowane z badanymi cechami jakościowymi. Dla poszczególnych problemów klasyfikacyjnych wyselekcjonowano najefektywniejsze metody klasyfikacji nadzorowanej.

Modele klasyfikacji uszkodzeń mechanicznych (obić) pozwoliły uzyskać zdolność predykcyjną na poziomie 95%. Skuteczność detekcji pestek i ich fragmentów w świeżych owocach wiśni wynosiła 95%, a w przypadku owoców mrożonych – 83,3%. Równie wysoką

zdolność predykcyjną (97%) uzyskano dla modeli służących do klasyfikacji owoców truskawek porażonych grzybami *Botrytis cinerea* oraz *Colletotrichum acutatum*.

**Słowa kluczowe:** obrazowanie hiperspektralne, jakość owoców, metody uczenia maszynowego, modele klasyfikacji nadzorowanej

## Abstract

High quality of fruits is one of the basic conditions for success on the domestic and international market. It is determined on the basis of analysis of nutritional values related to the chemical composition (e.g. content of sugars, vitamins, minerals, amino acids, enzymes, antioxidants), sensory properties (i.e. external appearance, taste, aroma, texture) or the occurrence of internal and external defects.

Hyperspectral imaging is a dynamically developing technique of image recording in a wide range of the spectrum, with simultaneous registration of spectral characteristics for individual pixels of the image. The essential features of the spectral imaging technique, indicating its usefulness for determining the qualitative characteristics of heterogeneous samples such as fruits, are noninvasiveness, analytical efficiency, a wide range of the spectrum examined (from ultraviolet to infrared) and a huge amount of information obtained from each scan (spectral resolution is several nanometers). The above features make the hyperspectral imaging systems more and more appreciated as a quick method of assessing the safety and quality of a wide range of agricultural products, including fruit.

This paper demonstrates the usefulness of the hyperspectral imaging method and artificial intelligence methods for determining selected fruit qualities. Depending on the specific application, VNIR and SWIR spectral ranges as well as the reflection or transmission modes of spectral data acquisition were used. Different methods of hyperspectral data preprocessing were analyzed in order to select characteristic absorption or transmission bands that allow to classify the examined qualitative features.

Several methods of data mining have been tested in the study, estimating their suitability for classifying such qualitative features as: varietal differences, soluble solid content, occurrence of mechanical damages, tissue inhomogeneities and diseases caused by fungal pathogens. The applied classifiers were constructed using a small (optimal) number of wavelengths, for which the spectral characteristics were most precisely correlated with the examined qualitative features. For each classification problem, the most effective supervised classification methods were selected.

The models of mechanical damage classification (bruises) allowed to obtain the prediction capability at the level of 95%. The effectiveness of detection of pits and their fragments in fresh cherry fruit was 95%, and in the case of frozen fruits – 83.3%. Equally high prediction accuracy (97%) was obtained for the models used to classify strawberries inoculated with *Botrytis cinerea* and *Colletotrichum acutatum* fungi.

**Key words:** hyperspectral imaging, fruit quality, machine learning methods, supervised classification models

