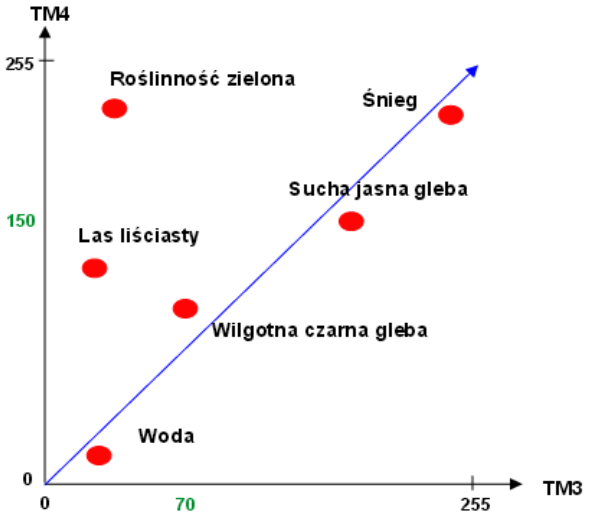
Zakład Hydrologii i Geoinformacji, Instytut Geografii UJK

Teledetekcja środowiska

Konwersatorium: 8
Temat: Cyfrowa klasyfikacja nadzorowana wielospektralnych zobrazowań satelitarnych - wprowadzenie

W ćwiczeniu wykorzystano informacje z preskryptu prof. Katarzyny Osińskiej Skotak – *Przetwarzanie i interpretacja zdjęć satelitarnych*.

Dane teledetekcyjne mogą być analizowane w celu wydobycia istotnych informacji tematycznych. Dane te są więc przekształcane w informację o charakterze tematycznym (np. klasyfikacja terenu pod kątem jego pokrycia i użytkowania). Klasyfikacja wielospektralna jest jedną z częściej stosowanych metod ekstrakcji informacji tematycznej. U podstaw klasyfikacji wielospektralnej leży założenie, że piksele reprezentujące podobne obiekty mają zbliżone właściwości spektralne, czyli w n-wymiarowej przestrzeni spektralnej (zbudowanej z osi reprezentujących poszczególne zakresy spektralne (ryc. 1) grupują się w naturalne skupiska (zwane klastrami) o posobnych właściwościach spektralnych. Definiując określone kryteria można więc rozdzielić grupy pikseli na podstawie ich właściwości spektralnych.



Ryc. 1. Dwuwymiarowa przestrzeń spektralna utworzona z dwóch zakresów spektralnych TM3 i TM4 oraz usytuowanie podstawowych typów pokrycia terenu w tej przestrzeni.

Obecnie, klasyfikacja wielospektralna może być wykonywana przy użyciu różnych algorytmów włączając w to: klasyfikację tzw. twardą z metodą nadzorowaną i nienadzorowaną, klasyfikację miękką wykorzystującą logikę rozmytą (z ang. fuzzy logic), rozwiązania hybrydowe stosowane często pomocniczo, jak również klasyfikacje obiektowe czy eksperckie. Najczęściej stosowanym podejściem do wydobycia informacji tematycznej na podstawie danych wielospektralnych jest klasyfikacja nadzorowana. Do wykonania tego rodzaju klasyfikacji można wykorzystać różnego rodzaju oprogramowanie – w tym przypadku klasyfikację wykonamy w programie SAGA GIS.

W klasyfikacji nadzorowanej, identyfikacja i lokalizacja niektórych typów pokrycia, takich jak: obszary miejskie, grunty rolne, łąki czy lasy są znane a priori (przed faktem) poprzez różnego rodzaju prace terenowe, analizy fotointerpretacyjne (fotografia lotnicza), analizę map lub doświadczenia osobiste. Analiza ogranicza się do zlokalizowania konkretnych miejsc, które dla danych teledetekcyjnych stanowić będą homogeniczne, jednorodne próbki znanych typów pokrycia

terenu. Obszary te zwane są polami treningowymi, ponieważ charakterystyki spektralne tych próbek będą użyte do trenowania algorytmu klasyfikacyjnego do ewentualnego wykonania mapy pokrycia terenu na podstawie obrazów wielospektralnych. Wiele zmiennych parametrów statystycznych (średnia, odchylenie standardowe, macierz kowariancji, macierz korelacji, itp.) jest obliczanych dla każdego z pól treningowych. Każdy piksel zarówno wewnątrz jak i poza obszarem pola treningowego zostaje oceniony i przydzielony do klasy, której jest najbliżej lub, której jest najbardziej prawdopodobnym członkiem. Często o tym typie klasyfikacji mówi się, że jest to klasyfikacja twarda bo każdy piksel jest przydzielany tylko do jednej klasy, podczas gdy sam sposób pozyskiwanych danych powoduje niejednoznaczność określanych klas (na obszarze jednego piksela może znajdować się np. 10% gleby odkrytej, 20% krzewów oraz 70 % lasu). Klasyfikację metodą nadzorowaną można wykonać m.in. przy pomocy algorytmów: prostopadłościanów, minimalnej odległości i największego prawdopodobieństwa.

Etapy tworzenia mapy pokrycia terenu

Główne etapy tworzenia mapy pokrycia terenu z wykorzystaniem metody klasyfikacji nadzorowanej na podstawie danych satelitarnych powinny obejmować:

1. Rozpoznanie natury klasyfikowanego problemu:
   a) zdefiniowanie obszaru zainteresowania;
   b) identyfikacja klas istotnych dla klasyfikowanego problemu;
2. Pozyskanie danych teledetekcyjnych i rozpoznanie obszaru badań:
   a) wybór danych teledetekcyjnych: rozważenie systemu teledetekcyjnego (rozdzielczość przestrzenna, spektralna, czasowa i radiometryczna); analiza środowiska badanego obszaru (cykle fenologiczne, atmosfera, wilgotność gleby, itp.);
3. Przetwarzanie cyfrowe danych teledetekcyjnych dla uzyskania informacji tematycznej:
   a) korekcja radiometryczna;
   b) korekcja geometryczna – czasami konieczne jest jej wykonanie przed procesem klasyfikacji, ale zalecane jest, aby wykonywać ją po wykonaniu klasyfikacji ze względu na konieczność ponownego próbkowania obrazu, która zmienia wartości radiometryczne;
   c) wybór metody klasyfikacji i/lub algorytmów (prostopadłościanów, minimalnej odległości, największego prawdopodobieństwa itd.);
   d)określenie wzorców klas - digitalizacja na ekranie poligonów reprezentujących pola treningowe na podstawie wiedzy a priori;
   e) obliczenie sygnatur, czyli wydobycie informacji statystycznej (wartość minimalna, maksymalna, wartość średnia, odchylenie standardowe, histogram jednowymiarowy, macierz wariancyjno – kowariancyjna, macierz korelacji międzykanałowej itp.), dla pól treningowych dla każdego zakresu spektralnego;
   f) analiza sygnatur pól treningowych pod względem ich jednorodności i rozłączności (analiza parametrów statystycznych, histogramów jednowymiarowych, wykresów dwuwymiarowych - skaterogramów);
   g) wybór najbardziej odpowiednich zakresów spektralnych (np. dywergencja, korelacja międzykanałowa);
   h) klasyfikacja - wydobycie informacji tematycznej na podstawie zdefiniowanych pól treningowych;
4. Ocena dokładności klasyfikacji:
   a) określenie danych testowych (a priori lub poprzez wybór losowy);
   b) określenie dokładności statystycznej sklasyfikowanej mapy (dokładność producenta, dokładność użytkownika, błędy z tytułu pominięcia i nadmiaru);

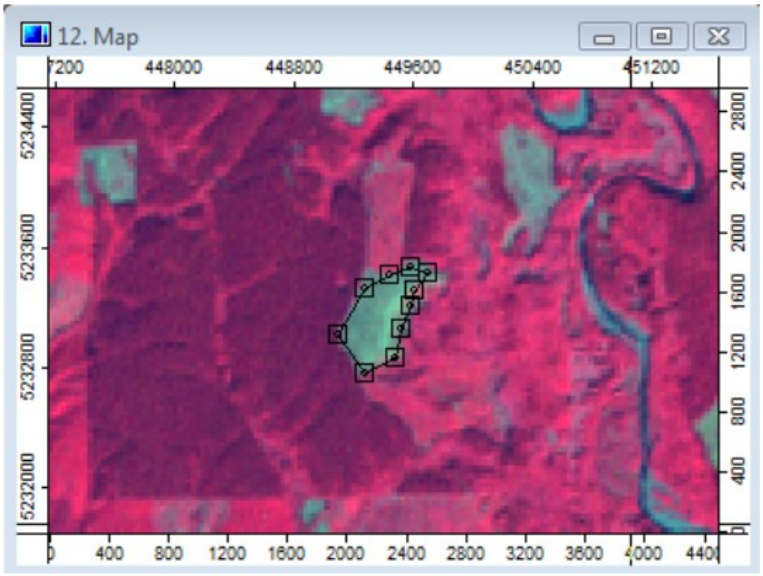
5. Dystrybucja wyników klasyfikacji (produkty cyfrowe, produkty analogowe – wydruki, raporty błędów, itp.).

Ćwiczenie do wykonania:

Zwektoryzować do formatu *.shp pola testowe, które będą stanowiły pola treningowe dla algorytmu klasyfikacji nadzorowanej na wskazanym przez prowadzącego obszarze.

Dane do ćwiczenia znajdują się na stronie internetowej www.ujk.edu.pl/zhig

Aby stworzyć poligony testowe w programie QGIS należy edytować plik o nazwie *super_clas_ts.shp* dorysowując do niego kolejne poligony. Plik ten ma już odpowiedni układ współrzędnych oraz strukturę tabeli atrybutów. Pola testowe wektoryzować trzeba na obszarach, których zagospodarowanie jest nam znane i jesteśmy w stanie przypisać mu jedną z poniżej przestawionych klas. Aby sprecyzować zagospodarowanie wybranego terenu można posłużyć się kanałem ósmym (czyli zdjęciem czarno-białym o wyższej rozdzielczości) oraz kompozycjami barwnymi i wzmocnieniami spektralnymi (szczególnie NDVI opisanymi w ćw. 5), na których łatwiej można ocenić rodzaje pokrycia/użytkowania terenu (ryc. 2).



Ryc. 2. Przykład prawidłowo zlokalizowanego pola testowego dla klasy 'łąki i pastwiska' na podstawie kompozycji barwnej wygenerowanej z wykorzystaniem widma podczerwieni (opisane w ćw. 5).

Podczas wektoryzacji należy poszukiwać następujących typów pokrycia/użytkowania terenu:
1. las liściasty
2. las iglasty
3. las mieszany
4. łąki i pastwiska
5. pola uprawne
6. gleby odkryte suche
7. gleby odkryte wilgotne
8. tereny zabudowane
9. odkryte skały

10. wody powierzchniowe

Wynikiem klasyfikacji nadzorowanej, którą przeprowadzimy na kolejnych ćwiczeniach będzie mapa sklasyfikowana zgodnie z powyższymi klasami poligonów testowych.