

SZEGEDI TUDOMÁNYEGYETEM
Természettudományi és Informatikai Kar
Földtudományok Doktori Iskola
Természeti Földrajzi és Geoinformatikai Tanszék

**Hiperspektrális távérzékelési módszerek alkalmazhatósága
tájvizsgálatokban**

Doktori (PhD) értekezés tézisei

Csendes Bálint

Témavezetők:

Dr. Merényi Erzsébet, egyetemi tanár, Rice University

Dr. habil. Mucsi László, egyetemi docens, SZTE

Szeged

2016

Bevezetés

A távérzékelési technika fejlődése mára lehetővé tette, hogy a kutatók akár országrész méretű területek felszínborítását térképezzék olyan pontossággal, amely az anyagi minőség finomabb eltéréseit is azonosíthatóvá teszi.

A környezeti monitoringozáshoz repülőgépekre, műholdakra telepített hiperspektrális szenzorok a multispektrális rendszereknél legalább egy nagyságrenddel több képi adatot készítenek, miközben a terepi felmérések már nem feltétlenül tudnak lépést tartani az osztályozások kiértékeléséhez kapcsolódó, megnövekedett információ igényvel.

A távérzékelésen alapuló földtudományi kutatások egyik legfontosabb feladata napjainkban ezért az automatizált adatfeldolgozási, illetve képi statisztikai módszerek kifejlesztése, melyek a korlátozott terepi adatgyűjtés mellett is megbízható eredményt nyújtanak.

Doktori dolgozatom témája az AISA hiperspektrális adatok tájvizsgálati alkalmazása. Munkám során magyarországi mintaterületekre jellemző környezetföldrajzi jelenségek példáján kívántam bemutatni a nemzetközi irodalomban elterjedt képfeldolgozási módszereket, továbbá új statisztikai módszerek fejlesztését határoztam el a jellegzetes felszín típusok, mint a túlnedvesedett talajok vagy az ártéri növényfajok spektrális diszkriminációjára.

Célkitűzések

1. A hiperspektrális távérzékelési technikát az 1980-as évek óta használják külföldön, míg Magyarországon rendszeresen az ezredforduló óta alkalmazzák a környezeti felmérésekben. Lehetséges hazai alkalmazásai közül a vegetáció- és a belvíztérképezésben, valamint városi és egyéb mesterséges felszínek elkülönítéséhez kívántam feltárni a hiperspektrális képfelvételezés lehetőségeit és sajátos, adatfeldolgozási és terepi vizsgálati igényeit, amihez a légi felvételek területét részben átfedő LiDAR domborzatmodell adatait is bevontam elemzésembe. A rendelkezésemre álló hiperspektrális adaton megoldást kerestem továbbá a cserjeszint növényeinek térképezésére is.

2. A légköri eredetű, illetve a szenzorhoz köthető zaj gyakran megnehezíti a spektrális minták azonosítását, a terheltebb adatsávokat ezért célszerű kivenni az elemzésből, ennek pontos módjára azonban még nincs általánosan elterjedt megközelítés a távérzékelési kutatásokban. Céлом volt ezért a reflektancia görbék spektrális diszkriminációjának és a redundancia mérséklésének matematikai-statisztikai hátterét elemezni, ismertetni az egyes algoritmusokat, majd alkalmazva azokat megoldást kínálni a hiperspektrális adatok képi zajtartalmának csökkentésére.

3. A felszínborítás nagy pontosságú térképezéséhez ismernünk kell az osztályok belső homogenitását és a képelemek reflektancia értékeinek szórását. Ezek ismeretében lehet az adat normalizációját elvégezni, ugyanakkor az osztályozások eredményességét is befolyásolja a spektrumgörbék eloszlása az egyes térképezési kategóriákon belül. Olyan

homogenitás vizsgálati módszert kívántam tehát kifejleszteni, ami a több száz adatsávot tartalmazó hiperspektrális adatokon is alkalmazható.

4. A tanítóterületek spektrális értékei között gyakran átfedések vannak, melyek összezavarhatják a képosztályozási algoritmusokat, egyúttal viszont fontos információval szolgálnak a vizsgált felszínekről. A térképezésekhez használt osztályok elkülöníthetőségét ezért úgynevezett spektrális szeparabilitási számításokkal mértem fel a transzformált hiperspektrális adatmintákon, majd az így kapott értékeket összevettem a félreosztályozások mértékével.

5. Célkitűzéseim között szerepelt továbbá az egyes irányított osztályozási algoritmusok összehasonlítása pontosság, megbízhatóság és az igénybevett számítási kapacitás szerint. Kritikai elemzésemet a Tápai-réten felvett mintaterületek lehető legszélesebb körét bevonva végeztem, mivel fontosnak tartottam, hogy olyan, sajátos követelményű távérzékeléses vizsgálatban elemezzem a módszerek hatékonyságát, ahol a mintapontok terepi gyűjtése időben vagy térben korlátozott, az osztályok nagy száma pedig tovább nehezíti a felszín típusok azonosítását.

6. A képosztályozások több megközelítési módszerét ismeri a távérzékelés irodalma, az adatok interpretációjára irányítatlan, valamint paraméteres és nem-paraméteres irányított klasszifikációkat is alkalmazhatunk, hatékonyságuk viszont függ az olyan korlátozó tényezőktől, mint a tanítópontok száma, a felvételek zajossága vagy a részletes terepi információ hiánya. Választ kívántam adni arra a kérdésre, hogy milyen távérzékeléses tájvizsgálat esetén érdemes irányítatlan osztályozást

használni és mikor célszerű tanítóterületeket alkalmazni. A két megközelítés relatív előnyeit és hátrányait magyarországi képfelvételek segítségével vetettem össze, a vegetációval borított felszínek mellett a belvizes mintaterületen is, figyelembe véve az algoritmusok elméleti háttéréből fakadó különbségeket és az elemzett földrajzi jelenségek, környezeti folyamatok térbeli mintázatát.

Eredmények

1.

A hiperspektrális légi felvételek 2010. szeptember 22-én, a vegetációs periódus végén készültek, geometriai pontosságuk 1,5 méter, ezért különösen alkalmasnak bizonyultak a növényzet részletes térképezésére. A 4,29-6,28 nanométeres spektrális felbontás mellett lehetővé vált a Tápai-réten található vegetációs felszínek elkülönítése növényfajonként és a spektrumkönyvtárak összeállítása, amit a mintaterület élőhelyeinek és mezőgazdasági parcelláinak teljes körű felmérése és azonosítása, valamint a tanítópontok kijelölése előzött meg. Spektrális szögeltérési számításokkal (SAM) kimutattam az invazív gyalogakác (*Amorpha fruticosa*) állományainak jelenlétét a telepített nemesnyár liget cserjeszintjében, eredményeim alapján a vizsgált erdőfoltot lefedő képelemek 1,5%-a rendelkezett a gyalogakác átlagspektrumával 0,03 radiánnál kisebb szög bezáró reflektancia értékkel és további 10,9 százalékkal mutatott 0,03-0,06 rad közötti spektrális szögeltérést mind a gyalogakác, mind a nemesnyár átlagos spektrumgörbéjével. A hiperspektrális adaton megbízható pontossággal sikerült elkülöníteni és térképezni továbbá a mesterséges

felszíneket és a belvízzel érintett területeket is, ahol a nedves talajfelszínnek osztálya egybeesett a LiDAR modell domborzati adatai alapján összeállított összegyülekezési térkép (Flow Accumulation Map) magasabb értékű képelemeivel. Az építkezésekkel átmozgatott talajfoltok osztályozása szerényebb pontosságú eredményt adott.

2.

A zajos sávok eltávolításánál elsősorban a képelemek értékei között fennálló területi autokorrelációt vettem figyelembe, melyet a Geary-féle C képlettel számítottam ki, míg a különböző fényintenzitás okozta anomáliákat normalizációval szűrtem ki. A redundancia további csökkentésére főkomponens transzformációt hajtottam végre, a főbb felszíntípusokat pedig az első két, albedó hatástól mentes főkomponens sáv spektrális terében, scatter plot diagramon is megjelenítettem.

3.

Az osztályok spektrális homogenitását kétfajta megközelítéssel is vizsgáltam, egyrészt a reflektancia értékek szórását számítottam ki minden ötvenedik spektrális csatornára, másrészt box-plot ábrán mutattam be az adatok eloszlását. Utóbbihoz egy olyan módszert fejlesztettem ki, ami hiperspektrális felvételekre is alkalmazható: a homogenitást az osztályátlagtól való spektrális szögeltérés alapján határoztam meg. A radiánban kiszámított értékek közül a maximális, a minimális és az átlagos szögeltérést, valamint a harmadik és az első kvartilis különbségét ábrázoltam a box-plot diagramon.

4.

A tanítóterületek szeparabilitását Jeffries-Matusita és Transformed Divergence algoritmussal is kiszámítottam. A 0-tól 2-ig terjedő skálán az alacsonyabb értékek jelzik a jelentősebb mértékű spektrális átfedést, a 2-es szeparabilitási index pedig a teljes elkülöníthetőséget. A vizsgálathoz az első öt transzformált főkomponens sávot használtam, mivel az AISA adatok a felmért mintapontoknál jóval nagyobb számú spektrális csatornát tartalmaznak.

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	TD
A	0	2	2	2	2	2	2	2	1,9	1,19	2	2	2	2	2	2	2	2	A
B	2	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	B
C	2	2	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	C
D	2	2	2	0	1,93	2	2	2	2	2	2	2	2	2	2	2	2	2	D
E	2	2	2	1,6	0	2	2	2	2	2	2	2	2	2	2	2	2	2	E
F	1,99	1,85	1,83	2	2	0	2	2	2	2	2	2	2	2	2	2	2	2	F
G	1,78	2	2	2	2	2	0	2	1,92	1,99	1,96	1,99	2	2	2	2	2	2	G
H	1,81	2	2	2	2	2	2	0	2	1,99	2	2	2	2	2	2	2	2	H
J	1,64	2	2	2	2	2	1,8	2	0	1,54	1,04	0,8	2	2	2	2	2	2	J
K	1	2	2	2	2	2	1,89	1,87	1,45	0	1,89	1,94	2	2	2	2	2	2	K
M	1,85	2	2	2	2	2	1,85	2	0,87	1,56	0	1,29	2	2	2	2	2	2	M
N	1,79	2	2	2	2	2	1,92	2	0,7	1,82	1,05	0	2	2	2	2	2	2	N
U	1,9	2	2	2	2	2	1,98	2	1,7	1,98	1,93	1,92	0	2	2	2	2	2	U
V	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	2	2	2	V
W	2	2	2	2	2	2	1,99	2	2	2	1,96	2	2	2	0	2	2	2	W
X	1,98	2	2	2	2	2	2	2	2	1,99	2	2	2	2	2	0	1,96	1,68	X
Y	1,95	2	2	2	2	2	2	2	2	1,98	2	2	2	2	2	1,89	0	1,11	Y
Z	1,97	2	2	2	2	2	2	2	2	1,99	2	2	2	2	2	1,56	1,02	0	Z
JM	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	

JM	Jeffries-Matusita index	TD	Transformed Divergence index
A	Amerikai köris	K	Fehér fűz
B	Aszfalt	M	Szürke nyár
C	Beton	N	Tölgy
D	Bolygatatlan talajfelszín	U	Vízben álló növényzet
E	Bolygatott talajfelszín	V	Nyílt vízfelület
F	Piros tetőcserép	W	Lucerna
G	Gyalogakác	X	Kukorica
H	Gyepfoltok	Y	Fokhagyma
J	Nemesnyaras	Z	Cukorrépa

1. és 2. táblázat: A vizsgált osztályok spektrális szeparabilitási értékei

Az eredményül kapott indexértékek alapján a Jeffries-Matusita-féle megközelítés érzékenyebb az osztályok spektrális átfedéseire, mint a Transformed Divergence módszer (1. és 2. táblázat), mindazonáltal mindkét szeperabilitási vizsgálat előre jelezte, mely osztályok esetén várható magasabb arányú félreosztályozás. Többek között az amerikai kőris (A, *Fraxinus pennsylvanica*) és a szürkenyár (M, *Populus x canscens*) példája is jól mutatta a kapcsolatot az alacsony (1,5 alatti) szeperabilitási számok és a szerényebb klasszifikációs pontosság között.

5.

A Tápai-rét mozaikos, természetes, adventív és mezőgazdasági növényekkel, valamint szénhidrogén-bányászathoz használt betonplatformokkal tarkított területeinek térképezésére olyan képosztályozási módszereket alkalmaztam, melyek kisszámú mintaponttal is megbízható eredményt adnak. Az SAM eljárás szerényebb pontosságot nyújtott (64%), ezért további klasszifikációkat végeztem, melyek közül a nem-paraméteres Support Vector Machine (SVM) gépi tanulási algoritmus esetén volt legkisebb a félreosztályozás mértéke (97%-os teljes pontosság, 0,97-os kappa index). A parametrikus eljárások közül korábbi tanulmányok eredményei alapján a Maximum Likelihood módszert alkalmaztam, amihez viszont a bemeneti adatsávok számát a rendelkezésre álló mintapontoknak megfelelően kellett korlátozni. A főkomponens transzformáció során végbement információ-veszteség ellenére a teljes pontosság ebben az esetben is 90% fölötti értéket mutatott. Figyelembe véve a tanítópontok korlátozott számát, az irányított képosztályozásra jellemző túlilleszkedést kiküszöbölendő egy úgynevezett hibrid SOM (Self-Organizing Map) rendszerben kombináltam az irányítatlan klaszterezést az irányított gépi tanulással.

P.	Producer's Accuracy	U.	User's Accuracy
A.		A.	
A	Amerikai kőris	K	Fehér fűz
B	Aszfalt	M	Szürke nyár
C	Beton	N	Tölgy
D	Bolygatatlan talajfelszín	U	Vízben álló növényzet
E	Bolygatott talajfelszín	V	Nyílt vízfelület
F	Piros tetőcserép	W	Lucerna
G	Gyalogakác	X	Kukorica
H	Gyepfoltok	Y	Fokhagyma
J	Nemesnyaras	Z	Cukorrépa

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	U.A.
A	9																		1
B		14						1											93%
C		1	15																94%
D				10	1														91%
E				5	14														74%
F						14													1
G							13			1									93%
H							2	14	1										82%
J									9	1	2								75%
K	2									14	1	1	1						74%
M	1						1		3	1	12	1							63%
N	1								2			12							80%
U	1												14						93%
V														17					1
W															15				1
X																13	1		93%
Y																1	13	3	77%
Z																1	1	12	86%
P.	64	93	1	67	93	1	81	93	60	93	80	75	93	1	1	87	87	80	86
A.	%	%		%	%		%	%	%	%	%	%	%			%	%	%	%

3. és 4. táblázat: Az SOM osztályozás tévesztési mátrixa

A mesterséges neurális hálózaton alapuló eljárás 86 százalékos teljes pontosságot nyújtott az első pontkiosztási halmazon, a keresztvalidációs halmazok átlaga pedig 82,4%, jelentősebb mértékű félreosztályozás csupán az amerikai kőris (A, *Fraxinus pennsylvanica*) és a nemesnyár (J, *Populus sp.*) esetében volt megfigyelhető (3. és 4. táblázat). Az irányított, nem-paraméteres módszerek (az SAM és az SVM) lényegesen kevesebb

számítási kapacitást igényeltek mind memóriahasználat, mind feldolgozási idő szempontjából, mint az iteratív, irányítatlan osztályozási algoritmusok, köztük az SOM klaszterezési eljárása. Az említetteken túl további nehézségeket és időbeli csúszást okozhatnak az adattranszformációs és – konverziós műveletek.

6.

A belvízborítás kiterjedése jelentős idő- és térbeli változékonyságot mutat, a gyakran csak néhány centiméter mélységű vízfoltok és a túlnedvesedett talajfelszínnek pontos kijelölése ezért nem minden esetben lehetséges. A 2010-2013 közötti évek belvizes időszakában azonban sikerült tanítóterületeket meghatározni mind a túlnedvesedett, mind a belvíz által nem érintett, száraz talajfelszínekről, amely mintapontokat a hiperspektrális adatok irányított osztályozásához használtam fel. Az SAM módszer 84, míg a bináris spektrumkódolás 90%-os pontossággal választotta szét a száraz és a túlnedvesedett talajokat, ugyanakkor világosan megmutatkoztak a tanító pontos klasszifikációk korlátai, ezért ISODATA alapú klaszterezést is végeztem a képkivágaton, melynek eredménytérképe a 100%-os pontosságon túl a belvizesedés térbeli mintázatát is jobban követi. A vegetációs fajok térképezésénél az irányított osztályozások adtak jobb eredményt, mivel ezek az algoritmusok, bár hajlamosak a túlilleszkedésre (overfitting), a gyakran terepen is nehezen elkülöníthető növények közti finom spektrális különbségekre érzékenyebbek, különösen akkor, ha az adat zajjal terhelt.

A tézishez kapcsolódó publikációk

Csendes, B. (2012): Detection of invasive plants on the flood plain of river Tisza, using hyperspectral airborne imagery. In: Tomislav Malvić, János Geiger, & Marko Cvetković (eds.): Conference book “*Geomathematics as geoscience*”: 4th Croatian-Hungarian and 15th Hungarian geomathematical congress, pp. 187-194.

Csendes, B. (2013): Invazív növények spektrális tulajdonságainak vizsgálata légifotókon. In: Gábor Keresztes (ed.): *Tavaszi Szél, 2013: Spring wind*, pp. 408-414.

Csendes, B. & Mucsi, L. (2016): Inland excess water mapping using hyperspectral imagery. *Geographica Pannonica* Vol. 20 Issue 4, in print.

Tobak, Z., **Csendes, B.**, Henits, L., van Leeuwen, B., & Mucsi, L. (2012): A városi felszín spektrális tulajdonságainak vizsgálata légifelvételek alapján. In: Diána Nyári (ed.): *Kockázat - Konfliktus - Kihívás: A VI. Magyar Földrajzi Konferencia, a MERIEXWA nyitókonferencia és a Geográfus Doktoranduszok Országos Konferenciájának Tanulmánykötete*, pp. 1088-1097.

Tobak, Z., **Csendes, B.**, Henits, L., van Leeuwen, B., Szatmári, J., & Mucsi, L. (2012): Városi felszínek spektrális tulajdonságainak vizsgálata légifelvételek alapján. In: József Lóki (ed.): *Az elmélet és gyakorlat találkozása a térinformatikában III.* - Térinformatikai Konferencia és Szakkiállítás, Debrecen, 2012.05.24.-2012.05.25. pp. 413-420.

Tobak, Z., **Csendes, B.**, Henits, L., van Leeuwen, B., & Mucsi, L. (2013): Légifelvételek spektrális és térbeli információtartalmának felhasználása a városi felszínborítás térképezésében. In: József Lóki (ed.): *Az elmélet és a gyakorlat találkozása a térinformatikában IV.* - Térinformatika Konferencia és Szakkiállítás, Debrecen, 2013.05.23.-2013.05.24. pp. 441-450.