

# A tematikus osztályozás matematikai leírása

Segédanyag a Távérzékelte felvételek elemzése tantárgyhoz

László István – Fekete István – Dezső Balázs – Csornai Gábor

## Digitális felvétel

Az elméleti tárgyalásban és az informatikai megvalósításnál a többsávú digitális felvételeket egy mátrixként fogjuk fel, amelynek elemei a földfelszín foltjainak felelnek meg, egy elem a legkisebb megfigyelhető felszínadarabot ábrázolja. A mátrixelemek a megfelelő felszínadarabról érkező mérési eredményt, vagyis a sugárzás mért intenzitását mutatják.

A képi adatrendszer jellemzően több sávból áll, így a mátrixelemek maguk is vektorok, melyek elemei a spektrális sávoknak felelnek meg. Egy  $N$  sorból,  $M$  oszlopból álló,  $B$  spektrális sáv adatait tartalmazó felvétel tehát a következőképpen írható le:

$$V = \begin{pmatrix} \vec{v}_{11} & \dots & \vec{v}_{1M} \\ \vdots & \ddots & \vdots \\ \vec{v}_{N1} & \dots & \vec{v}_{NM} \end{pmatrix},$$

ahol  $\vec{v}_{ij} = (v_{ij1}, \dots, v_{ijB})$ . Egy távérzékelte felvétel mátrixának elemei elméletileg folytonos mennyiségeket tartalmaznak, azonban a digitális tárolás során – amikor a mért sugárzási energiát nem tetszőleges pontossággal, hanem kvantálva ábrázoljuk – az intenzitásértékek rendszerint egy egész intervallum értékeit vehetik fel. Vagyis  $v_{ijb} \in [1..H]$ , ahol  $H$  a megkülönböztetett intenzitásértékek száma. Minden sáv esetén egyenlő  $H$  értéket tételezünk fel, ami nem jelenti az általánosság megszorítását.

## Intenzitástér

A távérzékeléses módszerek illusztrálásának szemléletes eszköze az *intenzitástér* (vagy *mérési tér*, angolul *measurement space* vagy *feature space*), amely a hagyományos térábrázolástól eltérő szemléletet valósít meg. A felvételek fent bemutatott ábrázolásánál a „koordinátatengelyek” (a  $V$  mátrix sorai és oszlopai, illetve az egy képpontot reprezentáló  $\vec{v}_{ij}$  vektor indexei, vagyis az elemeinek sorszámai) a térbeli koordinátákat és a spektrális sávokat jelölik, és az elemek intenzitásértékeket tartalmaznak. Az intenzitástér tengelyei viszont az egyes spektrális sávoknak felelnek meg, a koordináták pedig az intenzitásértékeket jelölik. Az intenzitástérben tehát egy pont helye a sugárzása (egyes sávokban mért) intenzitásának felel meg, és nincs kapcsolatban a pont térbeli helyzetével.

Az intenzitástérben egy ponthoz értéként hozzárendelhetjük például a felvételen az adott intenzitású pontok számát. Így a fenti  $V$  mátrix által reprezentált úrfelvétel intenzitástérbeli megfelelője felírható az alábbi  $V'$  mátrixszal, amely tulajdonképpen a kép (minden sávját figyelembe vevő, többdimenziós) hisztogramja.

$$V' \in \mathbb{N}^{\overbrace{H \times \dots \times H}^{B\text{-szer}}}$$

$$\forall (h_1, \dots, h_B) \in [1..H]^B : V'_{(h_1, \dots, h_B)} = \sum_{i=1}^N \sum_{j=1}^M \chi((v_{ij1} = h_1) \wedge \dots \wedge (v_{ijB} = h_B))$$

## A tematikus osztályozás bemenőadatai

A tematikus osztályozásnál a kiindulási képi adatrendszer minden pixeljéhez egy célkategóriát (osztályt) rendelünk a következő adatok alapján:

1. A célkategóriák halmaza:

$$\omega_1, \dots, \omega_{K_F},$$

ahol  $K_F$  a célkategóriák száma. (Az  $F$  indexet a később ismertetésre kerülő teljes osztályozási folyamat leírásával való egységesség kedvéért használjuk.)

2. Az egyes osztályokhoz tartozó pixelek valószínűségi eloszlása:

$$p(\vec{v}|\omega_1), \dots, p(\vec{v}|\omega_{K_F})$$

Vagyis  $p(\vec{v}|\omega_k)$  adja meg a  $\vec{v}$  intenzitásvektorok sűrűségfüggvényét, amennyiben  $\vec{v}$  a  $k$ . osztály eleme.

3. Az egyes osztályok előre ismert (a priori) előfordulási valószínűsége a teljes képen:

$$p(\omega_1), \dots, p(\omega_{K_F})$$

Tehát  $p(\omega_k)$  annak a – más információtól független – valószínűsége, hogy a kép egy pixele a  $k$ . osztályba tartozik. Ez az összetevő opcionális: nem mindig ismert, illetve rendszerint csak közelítő érték áll rendelkezésre pl. egy korábbi felmérésből az adott területre. Amennyiben nem állnak rendelkezésre, vagy nem kívánjuk használni az előzetes valószínűségeket, értéküket azonosnak ( $1/K_F$ ) tekintjük. Az osztályozás eredménye kevésbé függ az osztályok előre ismert valószínűségétől, mint a sűrűségfüggvényüktől.

4. A téves osztályozásoknak a felhasználó által meghatározott veszteségi értékei: jelöljük  $\lambda(\omega_k, \omega_l)$ -vel (vagy röviden  $\lambda_{kl}$ -l) azt az elemi veszteséget, amelyet egy  $\omega_l$ -beli pixel  $\omega_k$  osztályba történő sorolása okoz. ( $k, l \in [1..K_F]$ ) Ez az összetevő szintén opcionális; hiánya esetén a helyes osztályozás ( $\lambda_{kk}$ ) veszteségét 0-nak, míg minden helytelen osztályozás ( $\lambda_{kl}$ ,  $k \neq l$ ) veszteségét egységnyinek tekintjük.

## Az osztályozás eredménye

Célunk, hogy meghatározzuk a kiindulási képi adatrendszer pixeljeinek az osztálybesorolását, vagyis a következő mátrixot:

$$C^{(F)} = \begin{pmatrix} c_{11}^{(F)} & \dots & c_{1M}^{(F)} \\ \vdots & \ddots & \vdots \\ c_{N1}^{(F)} & \dots & c_{NM}^{(F)} \end{pmatrix} \quad c_{ij}^{(F)} \in [1..K_F],$$

ahol  $K_F$  az osztályok száma.

## Diszkriminánsfüggvény

Ehhez a fenti négy feltétel mellett keresünk egy  $g$  relációt, amely megadja a kapcsolatot a felvétel pixelértékei és az osztályok között. A  $g_k : [1..H]^B \rightarrow \mathbb{R}$  ( $k \in [1..K_F]$ ) ún. diszkriminánsfüggvények halmaza alapján egy  $\vec{v}$  vektort abba az  $\omega_k$  osztályba sorolunk, amelyre  $g_k(\vec{v})$  maximális. Amennyiben tehát a  $g_k$  diszkriminánsfüggvények ismertek, a  $g$  reláció a következőképpen írható fel:

$$g = \{(\vec{v}, k) \mid \vec{v} \in [1..H]^B, k \in [1..K_F], g_k(\vec{v}) = \max_{l=1}^{K_F} g_l(\vec{v})\}$$

Ez a reláció azért nem függvény, mert az osztályok találkozásánál ugyanarra az intenzitásvektorra több  $g_k$  függvény is adhat azonos értéket. Ezekben az esetekben a hozzárendelést másik szabállyal kell megtenni.

## A maximum likelihood- és a Bayes-osztályozás

Az előző szakaszban bevezetett terminológiát felhasználva az ún. *maximum-likelihood* döntési szabály diszkriminánsfüggvényei:  $g_k(\vec{v}) = p(\vec{v}|\omega_k)p(\omega_k)$ . Vagyis a maximum-likelihood-szabály abba az  $\omega_k$  osztályba sorol egy  $\vec{v}$  intenzitásvektort, amelynél  $p(\vec{v}|\omega_k)p(\omega_k) \geq p(\vec{v}|\omega_l)p(\omega_l)$ , minden  $l \in [1..K_F]$ -re.

Ez a döntési szabály minden egyes  $\vec{v}$  intenzitásvektorhoz azt az osztályt rendeli, amelyikbe a legnagyobb valószínűséggel tartozik, mivel egy adott  $\vec{v}$  vektorra  $p(\vec{v}|\omega_k)p(\omega_k)$  maximalizálása ekvivalens  $p(\omega_k|\vec{v})$  maximalizálásával. A feltételes valószínűség definíciója szerint ugyanis

$$p(\omega_k|\vec{v})p(\vec{v}) = p(\vec{v} \cap \omega_k) = p(\vec{v}|\omega_k)p(\omega_k),$$

és ezt felhasználva

$$p(\omega_k|\vec{v}) = p(\vec{v} \cap \omega_k)/p(\vec{v}) = p(\vec{v}|\omega_k)p(\omega_k)/p(\vec{v}).$$

Amennyiben veszteségi értékek is rendelkezésre állnak, használhatjuk az osztályozáshoz az ún. *Bayes-döntést*. Egy osztályozási módszer *Bayes-optimalis*, ha az átlagos veszteséget minimalizálja. A diszkriminánsfüggvények alakja  $g_k(\vec{v}) = -L_{\vec{v}}(k)$ , ahol  $L_{\vec{v}}(k)$ -val egy adott  $\vec{v}$  vektor  $\omega_k$  osztályba történő sorolásából származó átlagos veszteséget jelöljük:

$$L_{\vec{v}}(k) = \sum_{l=1}^{K_F} \lambda_{kl} p(\omega_l|\vec{v})$$

Behelyettesítéssel belátható, hogy a maximum-likelihood-osztályozás a Bayes-osztályozás speciális esete az elemi veszteségmátrixot használva.

$$(\lambda_{kl} = \chi(k \neq l); k, l \in [1..K_F])$$

## Clusterezés

Tekintsünk egy  $V$  képet, amely  $N$  sorból és  $M$  oszlopból áll, és tegyük fel, hogy a clusterek száma  $K_f$ . Ekkor a képpontok clusterbesorolását a következő mátrix határozza meg:

$$C^{(f)} = \begin{pmatrix} c_{11}^{(f)} & \dots & c_{1M}^{(f)} \\ \vdots & \ddots & \vdots \\ c_{N1}^{(f)} & \dots & c_{NM}^{(f)} \end{pmatrix} \quad c_{ij}^{(f)} \in [1..K_f]$$

(Az  $f$  indexet a később ismertetésre kerülő teljes osztályozási folyamat leírásával való egységesség kedvéért használjuk.)

Az egyik gyakran alkalmazott clusterezési kritérium a négyzetes hiba összegét méri („sum of squared error”, SSE). Jelöljük  $\mu_k^{(f)}$ -fel a  $k$ . cluster átlagát, melynek értéke értelemszerűen a következő (osztás alatt skalárral való komponensenkénti osztást értve):

$$\vec{\mu}_k^{(f)} = \frac{\sum_{i=1}^N \sum_{j=1}^M \chi(c_{ij}^{(f)} = k) \vec{v}_{ij}}{\sum_{i=1}^N \sum_{j=1}^M \chi(c_{ij}^{(f)} = k)}$$

Ekkor a négyzetes hiba összege a következőképpen számítható egy  $V$  felvételre,  $C^{(f)}$  clusterbesorolás mellett:

$$\text{SSE}^{(f)} = \sum_{k=1}^{K_f} \sum_{i=1}^N \sum_{j=1}^M \chi(c_{ij}^{(f)} = k) \|\vec{v}_{ij} - \vec{\mu}_k^{(f)}\|^2 = \sum_{i=1}^N \sum_{j=1}^M \|\vec{v}_{ij} - \vec{\mu}_{c_{ij}^{(f)}}^{(f)}\|^2$$

Ez az érték tehát a pontok és a nekik megfelelő clusterközéppontok (clusterátlagok) közötti távolság összege, a kép minden pontjára összesítve.

Az ISODATA a fenti SSE mennyiséget minimalizáló, iteratív clusterezőeljárás.

## Referenciaadatok

A clusterezés felügyelet nélküli folyamata után az alosztályok finomításának felügyelt folyamatában már felhasználjuk a referenciaadatokat, melyeket az inputkép méretének megfelelő tematikus raszteres kép formájában használunk fel:

$$R = \begin{pmatrix} r_{11} & \dots & r_{1M} \\ \vdots & \ddots & \vdots \\ r_{N1} & \dots & r_{NM} \end{pmatrix} \quad r_{ij} \in [0..K_F],$$

A mátrixelemek értéke az osztályok sorszámán kívül 0 is lehet, ami azt jelenti, hogy az adott helyen nem áll rendelkezésre referenciaadat. Két részre osztva használjuk: a tanuló-referenciaadatokra ( $R^{(\alpha)}$ ) és a teszt-referenciaadatokra

$(R^{(\omega)})$ . A két mátrix mérete is megegyezik  $R$ -ével. Azt az ideális esetben teljesülő tulajdonságot, hogy a tanulóadatok és a tesztadatok pontosan a referenciaadatok halmazának diszjunkt felbontását adják, a következőképpen írhatjuk le:

$$\forall (i, j) \in [1..N] \times [1..M] : r_{ij} \neq 0 \iff (r_{ij}^{(\alpha)} = r_{ij} \wedge r_{ij}^{(\omega)} = 0) \vee (r_{ij}^{(\alpha)} = 0 \wedge r_{ij}^{(\omega)} = r_{ij})$$

## Az alosztályok finomítása

A kép pixeleinek alosztálybesorását, amely az alosztályok finomításának lépéseiben alakult ki, a clusterbesoroláshoz hasonlóan ábrázolhatjuk:

$$C^{(f)} = \begin{pmatrix} c_{11}^{(f)} & \dots & c_{1M}^{(f)} \\ \vdots & \ddots & \vdots \\ c_{N1}^{(f)} & \dots & c_{NM}^{(f)} \end{pmatrix} \quad c_{ij}^{(f)} \in [1..K_f],$$

ahol  $K_f$  az alosztályok száma az  $f$ . fázisban, és  $f$  értéke  $G+1$ -től (a clusterezés utolsó lépését követő lépéstől)  $F-1$ -ig (a tematikus osztálybesorolást megelőző végső alosztálybesorolásig) terjed.

## Kontingenciamátrix

Az alosztályok finomításának lépései attól függenek, hogy a clusterek és a későbbi iterációk spektrális alosztályai milyen összefüggést mutatnak a referenciaadatokkal. Ennek vizsgálatára az automatikus gépi és a szakértői feldolgozás során is hasznos eszköz a *kontingenciamátrix*. Soraiban a clusterek, illetve az alosztályok találhatók, oszlopaiban a referenciaadatok tematikus kategóriái. (Az alosztályok finomításánál a referenciaadatok közül a tanulóadatokat használjuk.) Egy eleme az alosztály és a referenciaadatok egy kategóriájának metszetébe eső pixelek számát (illetve az általuk lefedett területet) tartalmazza:

$$T^{(f)} \in \mathbb{N}^{[1..K_f] \times [1..K_F]} \quad f \in [G..F-1]$$

$$t_{kl}^{(f)} = \sum_{i=1}^N \sum_{j=1}^M \chi(c_{ij}^{(f)} = k \wedge r_{ij}^{(\alpha)} = l)$$

## A végső osztálybesorolás

A finomítás eredményét akkor tekintjük megfelelőnek, amikor már minden alosztályhoz egyértelműen hozzá tudunk rendelni egy tematikus kategóriát. Ezt az egyes iterációkban kiszámolt kontingenciamátrixok, illetve az alosztályok eloszlásának vizsgálata alapján döntjük el. A végeredmény ismeretében tehát arra a – célként kitűzött – feltételre alapozva végezzük az osztályozást, hogy a tematikus kategóriák normális eloszlású alosztályok kompozíciójaként állnak

össze. Minden képpont besorolása két lépésben végezhető. Először egy alosztályba soroljuk be maximum likelihood-döntéssel, az alosztály paraméterei által meghatározott normális eloszlás sűrűségfüggvénye alapján. A második lépésben az alosztálysorszám által egyértelműen meghatározott tematikus kategóriához soroljuk.

Tekintsük a teljes osztályozási folyamat  $F$ ., utolsó lépésének a tematikus kategóriák hozzárendelését a pixelekhez, az ezt megelőző  $(F - 1)$ . lépését pedig az alosztálybesorolás végeredményének. Az előző bekezdés értelmében megadható egy  $\mathbb{A} : [1..K_{F-1}] \rightarrow [1..K_F]$  függvény az alosztályok végső halmaza és az osztályok (tematikus kategóriák) között, amelynek segítségével a képpontok osztálybesorolását megkaphatjuk a végső alosztálybesorolásból:

$$\forall(i, j) \in [1..N] \times [1..M] : c_{ij}^{(F)} = \mathbb{A}(c_{ij}^{(F-1)}).$$

Általában nincs ilyen megfeleltetés a  $c_{ij}^{(f)}$  és a  $c_{ij}^{(f-1)}$  értékek között, ha  $f < F$ .

## A pixelbesorolást megadó C mátrix jelentése

A pixelek besorolását megadó  $C$  mátrix értékei tehát jelenthetnek clustersorszámot, alosztálysorszámot vagy osztálysorszámot attól függően, hogy a teljes folyamat melyik fázisát hajtjuk éppen végre. Az értékek jelentését a következő táblázat foglalja össze.

clustersorszámok			alosztálysorszámok			osztálysorszámok		
1.	...	$G$ .	$G + 1$ .	...	$F - 1$ .	$F$ .		

Tehát az első  $G$  lépés a clusterezés iterációinak felel meg, ezt követi – rendszerint kevés, 2-3 lépésben – az alosztályok finomítása, és az egységes jelölésrendszerbe illeszkedve a végső osztálybesorolást is önálló,  $F$ . sorszámú lépésnek tekintjük.

## Az osztályozási eredmény pontosságvizsgálata

A *tévesztési mátrix* megadja az osztályozás eredményeként kapott és a referenciaadatok által meghatározott tematikus osztályok viszonyát a teljes vizsgált területre összesítve. Szerkezete és számításának módja megegyezik a kontingenciamátrixéval, azzal a különbséggel, hogy jellemzően az utolsó fázisban, a végső tematikus osztályokat tartalmazó  $C^{(F)}$  mátrixra számítjuk ki, és az összehasonlítás alapját a teszt-referenciaadatok képezik. A sorai az osztályozás eredményeként kapott tematikus osztályoknak felelnek meg, oszlopai a teszt-referenciaadatok tematikus kategóriáinak. Egy eleme az osztály és a referenciaadatok egy kategóriájának metszetébe eső pixelek számát tartalmazza:

$$T^{(F)} \in \mathbb{N}^{[1..K_F] \times [1..K_F]}$$

$$t_{kl}^{(F)} = \sum_{i=1}^N \sum_{j=1}^M \chi(c_{ij}^{(F)} = k \wedge r_{ij}^{(\omega)} = l)$$

## Szegmensalapú osztályozás

Szegmensalapú eset: kétszeres leképezés, vagyis minden képponthoz tartozik egy szegmenssorszám, és az változhat fázisonként, hogy egy szegmens pontjai melyik clusterhez/alosztályhoz tartoznak, de egy fázisban egy szegmens minden pontja ugyanahhoz a clusterhez/alosztályhoz tartozik.

A kép pixeljeinek szegmensbesorolása:

$$Z = \begin{pmatrix} z_{11} & \dots & z_{1M} \\ \vdots & \ddots & \vdots \\ z_{N1} & \dots & z_{NM} \end{pmatrix} \quad z_{ij} \in [1..S],$$

ahol  $S$  a szegmensek száma.

A szegmenssorszám megfeleltetése a cluster-/alosztálysorszámnak az  $f$ . fázisban ( $f \in [1..F]$ ):

$$\mathbb{S}^{(f)} : [1..S] \rightarrow [1..K_f], \text{ vagyis } \forall s \in [1..S] : \mathbb{S}^{(f)}(s) \in [1..K_f]$$

Tehát a kép pixeljeire nézve

$$\forall (i, j) \in [1..N] \times [1..M] : c_{ij}^{(f)} = \mathbb{S}^{(f)}(z_{ij}), \text{ ezért } \mathbb{S}^{(f)}(z_{ij}) \in [1..K_f].$$

A szegmentálást tekinthetjük úgy, mint az osztályozás „nulladik” fázisát:

$$Z = „C^{(0)}” = \begin{pmatrix} c_{11}^{(0)} & \dots & c_{1M}^{(0)} \\ \vdots & \ddots & \vdots \\ c_{N1}^{(0)} & \dots & c_{NM}^{(0)} \end{pmatrix} \quad c_{ij}^{(0)} \in [1..S] = „[1..K_0]”$$

Ezzel a jelöléssel minden  $f$  fázisra ( $f \in [1..F]$ ):

$$\mathbb{S}^{(f)}(c_{ij}^{(0)}) = c_{ij}^{(f)},$$

de általában nincs olyan  $\mathbb{C}$  függvény, amelyik az egymás utáni fázisok alosztály-/clustersorszámait feleltetné meg egymásnak, vagyis amelyre minden  $f \in [2..F-1]$  fázisban igaz lenne, hogy

$$\mathbb{C}^{(f)}(c_{ij}^{(f-1)}) = c_{ij}^{(f)}.$$