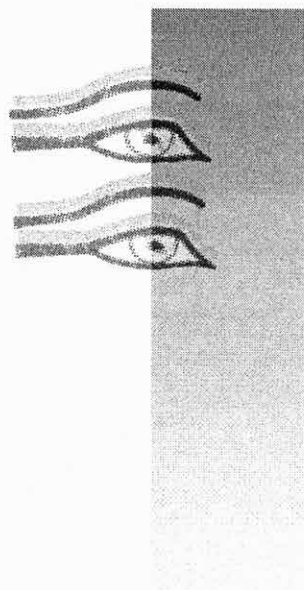


# Karakterfelismerés raszteres térképeken



A Sintagma projekt

PLIHÁL Katalin – JANCSÓ Tamás

Tájékoztatósi szempontból a kartográfiai dokumentumok elektronikus elérhetősége, ill. visszakereshetősége éppolyan fontos, mint a monográfiáké, időszaki kiadványoké, tankönyveké stb. A KSZ/5 Földrajzi nevek mint adatbázisrekordok hozzáférési pontjai c. szabályzat 2005-ös megjelenése után sok szempontot mérlegelő, mélyreható elemzések jelentek meg e folyóirat hasábjain a szabályzat gondolatmenetéről, megoldásairól, s ennek kapcsán a földrajzinév-használat problémáiról, ellentmondásairól.\*

A szkennelt kartográfiai dokumentumok minőségi használatát, olvashatóságát az optikai karakterfelismerő programok (OCR) segítik, de a térképeken lévő feliratok sokfélesége komoly nehézségek elé állítja a programozókat.

Az Országos Széchényi Könyvtár Térképtára a közelmúltban kezdte meg a gyűjteményében található régi térképek szkennelését. A létrejött térképarchívum jól használható segédeszköze lett a térkép kutatóknak, történészeknek, etnográfusoknak, ill. az egyéb tudományterületek kutató-

inak. A térképeken való keresés hatékonyságát nagyságrendekkel lehet növelni a térképekhez kapcsolt helyesnévtárakkal. Ezek létrehozásához szükség van a karakterek automatizált felismerésére. Bizonyos térképek digitálisan csak raszteres formában érhetők el, ill. gyakori, hogy térképek interneten történő publikálása is csak raszteres formában valósítható meg.

Írásunkban bemutatjuk a Sintagma projektet\*\*, mely raszteres térképek feliratainak annotációját (térképi információk adatbázisba szervezését), lekérdezését, megjelenítését oldja meg integrált webes környezetben. Külön foglalkozunk azzal, hogy milyen követelményeknek kell a különböző térképtípusoknak megfelelniük ahhoz, hogy a kifejlesztett karakterfelismerő program hatékonyan le tudja rövidíteni az adatbevitelhez szükséges időt.

A térképeken lévő feliratokat tartalmazó adatbázis központi elemét a jelkulcsok jelentik. A jelkulcsok szerint tárolt adatokat integráltan az Objektum adattábla tartalmazza (1. ábra).

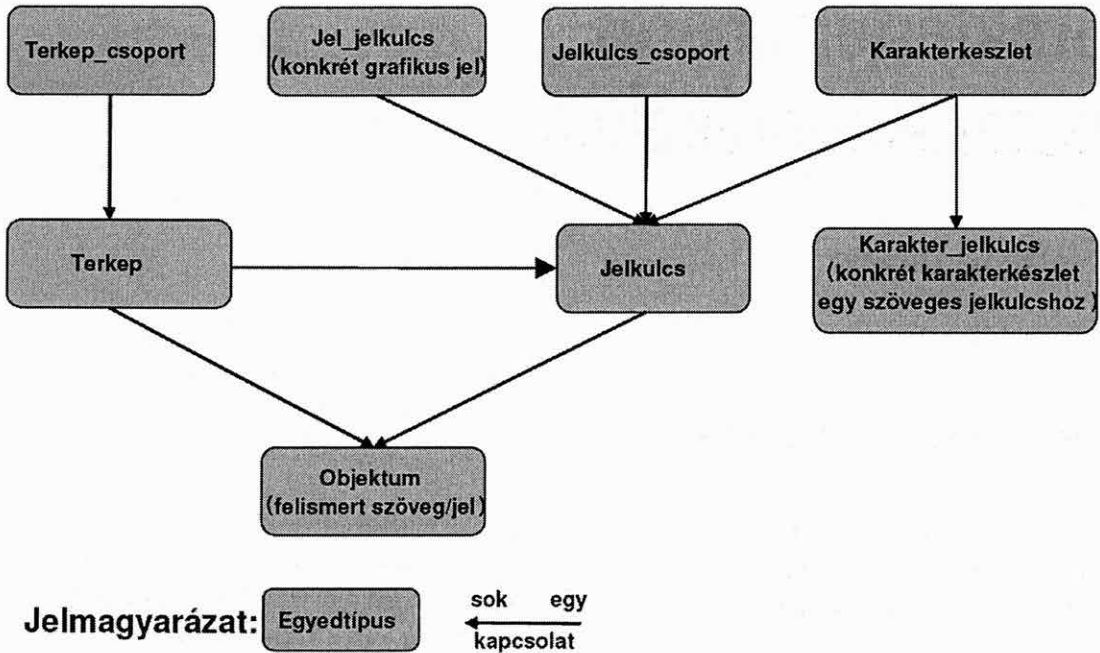
\* BERKE Barnabásné: Egy új magyar szabályzat és kitekintés a nemzetközi szabályzatalkotási munkálatokra. In: Könyvtári Figyelő, 15. (51.) évf. 2005. 4. sz. 761–764. p.; UNGVÁRY Rudolf – PÁSZTI László: A földrajzi nevek mint az adatbázisrekordok hozzáférési pontjai. Egy könyvtári szabályzat elvei

és megoldásai. In: Könyvtári Figyelő, 15. (51.) évf. 2005. 4. sz. 765–790. p.; FARAGÓ Imre: A magyar földrajzinév-használat. In: Könyvtári Figyelő, 15. (51.) 2005. 4. sz. 791–815. p.

\*\* A Sintagma projekt száma: NKFP-2/052/2004, kidolgozója Dr. Jancsó Tamás főiskolai docens.

1. ábra

Térképi adatbázis felépítése



Az adatbázis feltöltésének meggyorsítása érdekében a projekten belül kidolgoztunk egy speciális karakterfelismerő algoritmust. A felismerési algoritmus kiválasztása volt az első feladat. Ehhez több algoritmust kellett részletesen kidolgoznunk, hogy látható legyen, melyik hozza a leghatékonyabb eredményt. A feladatot nehezítette, hogy a legtöbb OCR-algortmussal szemben itt fel kellett készíteni az algoritmust arra is, hogy a háttér színes és jelentősen „zajos” is. A térképeken sokféle betűtípus fordul elő, különböző színben és méretben, amelyek kezelése eleve nehéz feladat elé állítja a karakterfelismerő szoftvert, de az igazi problémát az okozza, hogy a szöveg bármilyen irányban, ill. szögben, sőt pl. egy folyó vagy hegylanc esetében ívesen is szerepelhet a térképen.

## Felismerési algoritmus

Elsőként egy euklideszi távolságon alapuló pixel alapú illesztést dolgoztunk ki. Erről gyorsan kiderült, hogy ebben a formában lassú és sok

benne a hibalehetőség, de mindenképpen szükséges a karakterfelismerésnél az ellenőrzéshez. A sebesség gyorsítása egyszerűbb feladat volt. Ehhez csak azt kellett megoldani, hogy minden műveletet előre elvégezzünk és letároljuk a memóriában (pl. a letárolt mintakép forgatását, nagyítását, és „torzítását” stb.). A hatékonyság javításához, a képet „túlmintavételeztük”, majd a többi letárolt mintához hasonlóan méreteztük, és nem euklideszi, hanem tangens távolságmértést alkalmaztunk.

További algoritmusokat is készítettünk. Ilyen volt pl. a Hough transzformáción alapuló geometriai felismerés. A kivágott képen Hough transzformáció történik, hogy megállapítsuk a betűt alkotó jellegzetes szakaszokat és azok tulajdonságait. Ez a megoldás azonban túl sok szakaszt eredményezett, melyek között voltak hamisak is. Az is előfordult, hogy a hamis szakaszok domináltak, vagyis nem valós képet mutattak a betűről.

A következő algoritmus a topológiai információk kinyerésén alapult, melyhez először el kel-

lett készíteni a betű vázát, vagyis az egy pixel vastagságú középvonalát, mely megadta a betű „nyomvonalát”. Ehhez először a Stentiford algoritmust használtuk, ami viszont nem mindig eredményezett „szép” képet. A valósághoz közelebb álló betűképet adott a másodiknak kipróbált Zhang-Suen algoritmus, főleg miután néhány ötletet a Stentiford algoritmusból is átvettünk, és használtuk a „Holt Staircase Removal” nevű algoritmust az álcsomópontok kiszűrésére. Lényeges kérdés volt még, hogy milyen jellemzőket (feliratokat, ikonként ábrázolt grafikus jeleket) és milyen módszerrel vonjunk ki. Lehetőségek:

- FFT-vel (a jelfeldolgozásban alkalmazott szűrési eljárással)

Az első próbálkozás a körvonal koordinátáinak 2D FFT analízise volt. Ehhez csak a körvonalat alkotó pontokat kellett előtte extrapolálni a második hatványára. Az eredmény egy viszonylag tömör méret és forgatás érzéketlen „kivonat” lett. Felismerésnél ezeket a kivonatokot kellett összehasonlítani. A tapasztalat azt mutatta azonban, hogy viszonylag kis eltérések a betűben már nagyobb hibákat okoztak, vagyis meglehetősen nagy mintahalmaz kellett akár egyetlen betű megjelenítéséhez is.

- Súlyfüggvény

Az FFT függvény helyett súlyfüggvényekkel működött a program, de sajnos ezekről hamar kiderült, hogy nem hozták meg a megfelelő eredményeket.

- Vektorizálással

A legjobb eredményt végül a váz vektorgrafikus átalakítása hozta, melyben a kapott körvonalaikat a minimális szükséges számú szakaszokra és pontokra bontotta a program, és ezen szakaszok és pontok jellemzőit tárolta le. Emellett az előzetes szelektálás érdekében a vázban lévő csomó- és végpontokat is felismerte, illetve a belső kontúrok számát és elhelyezkedését is kezelni tudta a program.

A végső, empirikusan legjobb eredményt adó algoritmusban a tangens távolságot mérő algoritmus és a betű vázának vektorizálásán alapuló algoritmus ötvözése valósult meg.

Megjegyzendő, hogy az automatikus karakterfelismerés csak a következő feltételek betartásával működtethető gazdaságosan ráfordítási idő-költség tekintetében:

1. A szkennelt térkép minősége legyen annyira jó, hogy a karaktereket alkotó pixelmátrixok legalább 8x8 pixelből álljanak. Ennél kisebb pixelszám esetén a betűk jellemzőit leíró információk megbízhatósága rohamosan romlik. Ennek megfelelően a szkennelés átlagos felbontása 600 DPI fölött legyen.

2. A szavakat alkotó karakterek nem érhetnek egymáshoz, közöttük legalább egy pixel távolságnak lennie kell, ellenkező esetben a betanítás, illetve az automatizált karakterfelismerés értelmét veszti.

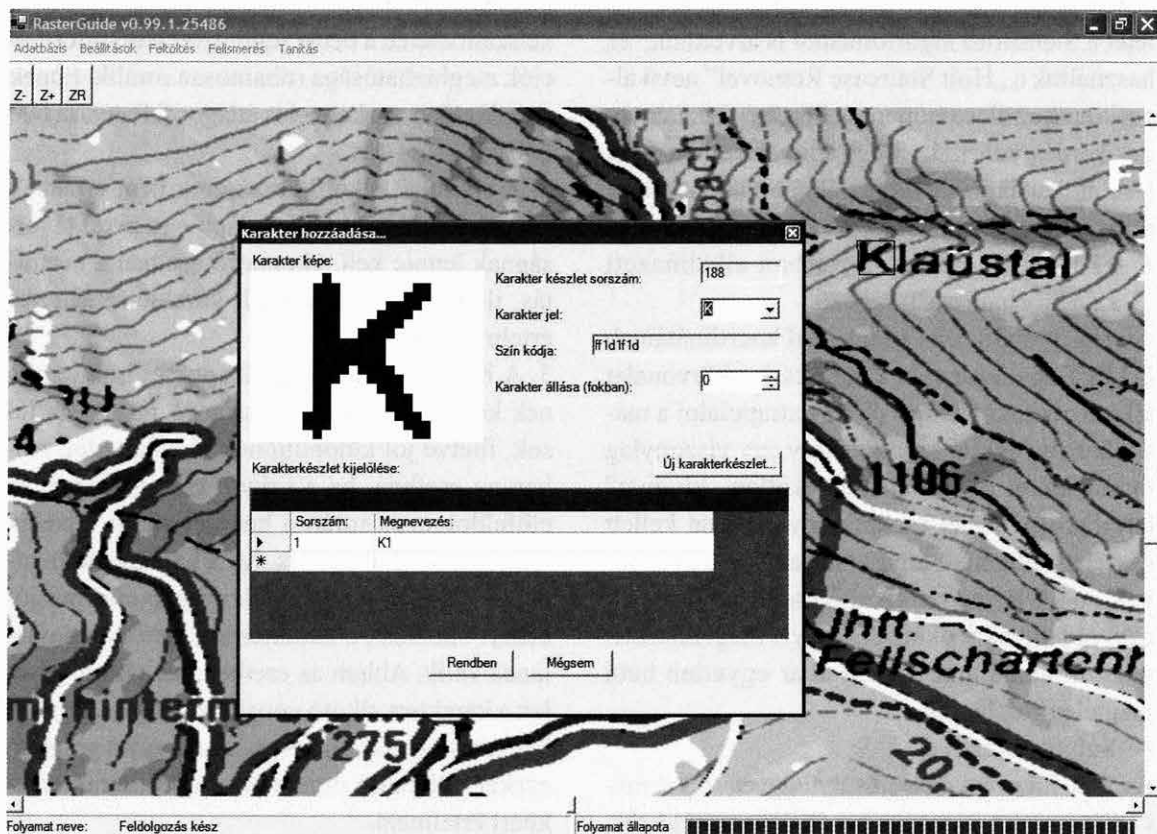
3. A betűket alkotó színek spektrálisan legyenek közel egymáshoz és számuk ne legyen túl sok, illetve jól különüljön el a háttértől. Abban az esetben, ha a színek száma magas, az előfeldolgozás során a betűket alkotó pixelek száma lecsökken, kiesik, így a karakterekből hiányzó részek a topológiai leírásban információhiányt okoznak, a karakterfelismerés bizonytalanává válik. Abban az esetben, ha a kieső pixelek a karaktert alkotó vonalakat megszakítják, a betanítás nem lesz sikeres, hiszen így a program ezeket különálló objektumokként (karakterekként) értelmezi.

A fenti feltételekből következik, hogy az algoritmus 1920 előtti térképek esetében csak korlátozottan alkalmazható, hiszen jóval nagyobb az olyan térképek száma, melyeken a karakterek összekötéssel kapcsolódnak. Ilyen típusú feliratok felismerése teljesen más megközelítést kíván, mely ennek a projektnek nem feladata.

A térképi annotációt támogató adatbeviteli program modulja a következő funkciókkal bír:

- karakterkészletek betanítása térképtípusok szerint,
- karakterek automatizált felismerése a betanítás után,
- grafikus jelkulcsi elemek felismerése minta-illesztéssel,

- mdb típusú adatbázis felépítése, melyben az „Objektum” adattábla tartalmazza a felismert térképi elemeket.

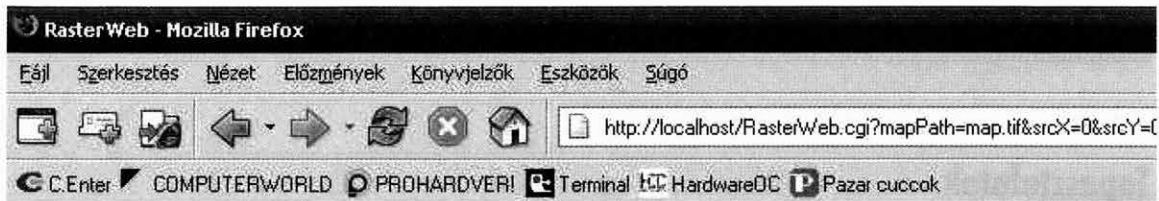


### 1. kép

*Karakterek automatizált felismerése a Rasterguide program-modul segítségével*

Az „Objektum” adattábla a hozzá kapcsolódó térképpel szolgáltatja az adatokat a Sintagma rendszer számára. Az adattábla lekérdezése során a teljes rekordkészlet szűrése megtörténik. Az eredménytábla rekordjainak megfelelő térképi részleteket „Map-server” funkcióban,

kliens-szerver architektúrában a Map-server program-modul szolgáltatja, illetve továbbítja az interneten keresztül, a kliens program pedig integrált Sintagma környezetben a felhasználó számára megjeleníti azt.



Kiszolgáló szoftver verziószám: 1.0.2586.26847

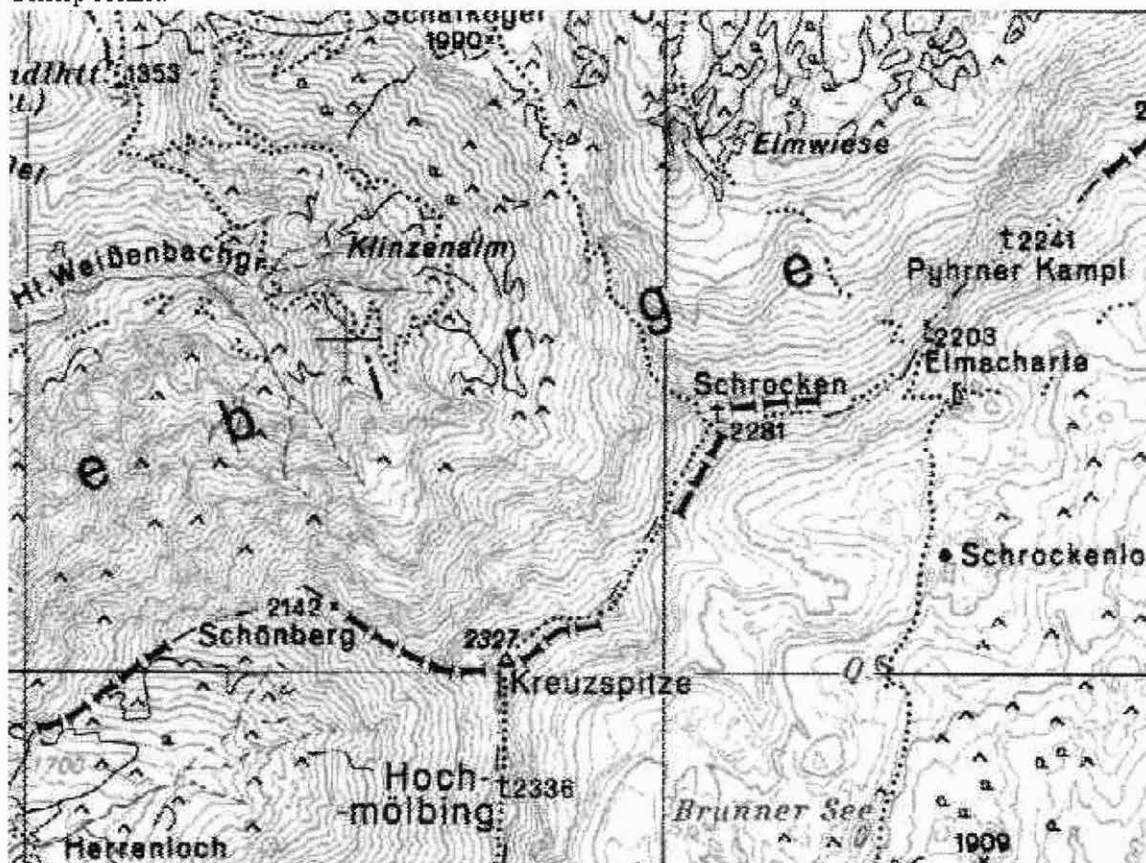
Elérési útvonal: map.tif

Bal-felső koordináták: 0 , 0

Szélesség és magasság: 1000 , 700

Minőség: 50%

Térkép részlet:



Teljes kép letölthető itt: [map.tif](#)

## 2. kép

*A Map-server megjelenítési funkciója*

## Tapasztalatok

A Sintagma projekt karakterfelismerő szoftverének használatával kapcsolatban szerzett tapasztalataink kedvezőek voltak. A projekt tervezői egy nagyon összetett és bonyolult – eddig sikeresen senki által még meg nem oldott – feladatra kívántak számítógéppel támogatott megoldást kínálni. Valamely ország, térség földrajzi névadatbázisának elkészítése manuálisan nagyon időigényes munka, s ezt jelentősen megkönnyítené, ha a térképek névanyagát számítógépes módszerrel, olcsón és kevés hibával lehetne kinyerni.

Hogy valamelyest érzékelní lehessen a program tervezői előtt tornyosuló „hatalmas hegyet”, gondoljuk végig, mi is a térkép és mi a szerepe a térképen található alfa-numerikus jeleknek.

Pragmatikus értelmezésben a térkép olyan modell, amelynek bizonyos értelmű helyettesítő szerepe van, illetve lehet. A modellképzés – írja *Klinghammer István* akadémikus – a következő kérdéssémának alávetett; miről, kinek, mikor és mihez készült, készül a térkép.

Szemiotikai kifejezéssel élve a térképmo­dell pragmatikus entitás, és így egy többjegyű állít­mány megvalósulása: a térkép az eredeti (a va­lóság) egy modellje, egy térképkészítőtől, egy térképhasználónak, egy bizonyos időben, egy bizonyos intencióra vonatkoztatva.<sup>1</sup>

A térbeli információk megjelenítésére különböző modelltípusok alkalmasak:

- rajzi modellek (térképek),
- technikai modellek (dombortérképek),

- szemantikai modellek.

Míg a rajzi és technikai modelleket lényegében a szemléletesség, az előállíthatóság és a manipulálhatóság jellemzi, addig a szemantikai modellek mindig jelmodellek, a fogalmak, az érzelési és elképzelési rendszerek, egyáltalában a gondolati rendszerek számára. A külvilágról belsőleg alkotott modellek képzésével férünk tulajdonképpen hozzá a külvilág jelzés-konstellációjához, jelzésállományához. E belső modelleket kiegészítik a kognitív modellek, ezek a tulajdonképpeni gondolati operációkat, a percepció modellek átszerkesztését és feldolgozását segítik elő.<sup>2</sup>

E modellképzés során a közvetítő szerepet vállaló térképalkotó az ismereteket a befogadó, felhasználó felé az évszázadok alatt kialakult kifejezésformák alkalmazásával tolmácsolja. A térképi nyelv és a hétköznapi emberi nyelv szembeállításával világosan megmutatja, hogy a kartográfiai jelrendszer csak térhez kötött információkat alkalmaz. Ha a közvetítő és a befogadó nem „egy nyelven beszélnek”, az ismeretek nem juthatnak célba, ugyanis e művek „fordítások” útján nem közvetíthetők! Ezt a gondolatot fogalmazta meg *Jaques Bertin* „A térkép csak attól a pillanattól kezdve hasznos, amikor az olvasó az ábrázolt területe saját földrajzi tudatképében el tudja helyezni.”<sup>3</sup>

A természetes nyelvek és térképi kifejezésformák felépítésének lényeges különbségeit a következő ábra is jól szemlélteti.<sup>4</sup>

Ismérv	Nyelv		Kartográfiai kifejezési forma
	beszélt	írott	
A legkisebb egység (jel, alapelem)	hang	betű	Pont, vonal, felület
Összetett jel (szuperjel, konstrukció eszköz)	Szó, mint egy tárgy predikálója, mint tulajdonnév vagy, mint indiká- tor (kapcsolatok megadása) (rövidítés, formális jel is)		Szignatúra (jel) Diagram Árnyalat (tónus) Írás (szó)
információközlés	mondat		A megformálási (ábrázolási) eszközök hozzárrendelése és rendezése
Jelvariációk (hangtani és rajzi)	Tempó Hangerő Hangszín intonáció	Forma, nagyság, szín, árnyalat, kitöltés	
	A mondat szórendje		A rajzi súlypontozás

A térképjelek mindig fogalmi információkat tartalmaznak, melyek az ábrázolandó valóság és tényállás gondolati, racionális konstrukciói. A térképi jel mint grafikai jelzés az alapvető rajzi elemeken, a ponton, a vonalon és a felelületen alapul.

Az azonos jelenségek ábrázolására már a 16. századi, mérések után készült térképek alkotói is hasonló jeleket kezdtek alkalmazni. Az idők során e jelek elveszítették eredeti formájukat és átalakultak a napjainkban is használt piktogrammokká. A térképjeleket – mint fogalmi ismeretet – egyetlen térkép sem tudná eredményesen közvetíteni, ha nem használna alfanumerikus adatokat.

A betűk és számok nem esetlegesen, hanem az esetek többségében igen szigorú szempontok szerint kerültek a rajzban elhelyezésre. A betűmetszés – most a különböző nyomtatási eljárások egyedi sajátosságait e helyütt figyelmen kívül hagyva – korán önálló szakmává vált. Az írásmetsző művész ügyessége, rátermettsége meghatározta egy-egy térkép szépségét, használhatóságát. Írás (=névanyag) hiányában a térkép szép „kép” lehetne, míg képi elemek hiányában csak „értelmetlen szavak, kifejezések összefüggéstelen halmazává” válna. A fen-

tiek alapján megállapíthatjuk, hogy a tér-kép és a névírás együtt eredményezi azt az ismeretközlést, amely elvezetheti a felhasználót a készítő által remélt térbeli valóság megismeréséhez is. A térkép megírása összetett vagy egyedülálló betűkből, szavakból, mondatokból és számokból állhat, ezeket kell oly logikusan, az ábrázolt mű méretarányával összhangban álló módon elhelyezni, hogy segítse a felhasználót az értelmezésben és ne zavarja, vagy akadályozza.

A betűk kiválasztásában az alábbi szempontokra kellett és kell figyelemmel lenni:

1. Az olvashatóság, illetve érthetőség. A betűk könnyen és hiba nélkül akkor olvashatók, ha mentesek minden felesleges vonaltól, ha az olvasó szeme a betűk vonalait mind nagyság, mind vastagság tekintetében egyenlően fogja fel. Néhány térképnyomtatási típust nem számítva, a térképek névírásának szépsége a térképraező vagy a betűművész tehetségétől függött. Az 1920-as évektől már szinte az összes nyomtatott térkép névanyagát nyomdailag szedett betűkkel, készítették, készítik.

2. A betűnagyság, illetve kifejezőképesség. A térképeken használt betűk nagysága általában nem függött a méretaránytól. Nem túl szerencsés, ha a betűk túlságosan vékonyak, vagy ha

a méretük a 0,8 mm-nél kisebb. A betűtípusok és -nagyságok változékonysága olvashatóbbá, míg az ennek ellenkezője, monotonná teheti a térképet.

3. Elrendezés, esztétika. A névírás elrendezése gondos körültekintést kívánt és kíván meg, a sűrű megírással ugyanis nem növeljük a térkép ismeretközvetítő képességét, hanem romboljuk. Természetesen a névírásnak szoros a kapcsolata a térképi jelekkel, ezért arra kellett és kell törekedni, hogy a felhasználó számára egyértelmű legyen a grafikai jel és a reá vonatkozó alfa-numerikus jel, vagy jelek kapcsolata. Az írás olvashatóságát térképek esetén még az is befolyásolja, hogy az lehetőleg párhuzamosan legyen valamely térképi vonallal (például térképkeret, szélességi vagy hosszúsági vonallal) elhelyezve. Olyan írássorok, amelyek valamely domborzati elemre, vagy vízhálózatra, stb. vonatkoznak, azal párhuzamosan kerüljenek elhelyezésre. Ha olyan írássort kell elhelyezni a térképen, amely valamely hosszúsági vonallal párhuzamosan haladna, akkor a képzeletbeli középvonaltól keletre eső neveket felülről lefelé, míg a fenti vonaltól nyugatra esőket alulról felfelé vezetően célszerű írni. Így a betűk majd nem állnak fejjel lefelé, vagy felfelé.

Ha a fent megfogalmazottak fényében vizsgáljuk meg a Sintagma program keretében készült szoftver használhatóságát, akkor tapasztalatainkat a következőkben foglalhatjuk össze:

1. A nem gépi úton és/vagy nem betűsablon használatával megírt térképeken végzett karakterfelismerés eredménye igen rossznak mondható.

a.) Oka részben a megkívánt nagy felbontásban (legkevesebb 600 dpi, de inkább 1200 dpi) ke-reshető. A térképek nagy mérete miatt így olyan „gigantikus” méretű állományok keletkeztek, amelyekkel a munka nem egyszerű. Számptalan olyan „zaj” található, elsősorban az 1920 előtt nyomtatott térképeken, amelyekre szinte minden térképhez időtől-időre a programot külön kellene betanítani, ami felette időigényes. Így a tisztán

és helyesen kinyerhető alfa-numerikus adatok mennyisége egy-egy térkép esetén nem érte el a 25%-ot, azaz a nevek 75%-át egyedileg kellett a helyes karakterekkel kiegészíteni.

b.) A térképnyomatás során hosszú időn át a térképek alfa-numerikus adatait is kézzel vitték fel arra a felületre, amelyről a sokszorosítás történt. Tehát a programot egy-egy metsző, rajzoló egyéni stílusára is meg kellene tanítani, ami szintén felette időigényes. Ugyanakkor azt is megfigyelhetjük, hogy egy-egy metsző vagy rajzoló attól függően, hogy milyen szerepe volt egy-egy alfa-numerikus adatnak az adott térkép egészét illetően, eltérő stílust használhatott, használt.

2. Ha a már említett program keretében kidolgozott szoftvert modern nyomtatott, majd digitálisan raszteres formátumban rögzített kataszteri térképeken teszteltük, az eredmény akár kiválónak is mondható. Véleményünk szerint a földnyilvántartás számítógépes adatbázisba áttemelésével az 1950 után nyomtatott kataszteri térképek alfa-numerikus adattartalmát ilyen módon lehetne automatizáltan gyorsan kinyerni. Az így keletkezett térképek száma nem csekély, e szoftver sikeresen épülhetne be akár a „Nemzeti kataszteri programba”, így segítve a kataszteri nyilvántartás modernizálását.

## Jegyzetek

1. KLINGHAMMER István: A kartográfia kialakulása napjainkig. (Akadémiai doktori disszertáció). Budapest, 1991. 45. p.
2. KLINGHAMMER i. m. 45. p
3. KLINGHAMMER i. m. 92. P.
4. KLINGHAMMER i.m. 49. p.