

Hosszú ideje foglalkoztatja és még sokáig foglalkoztatja majd a kutatókat, megoldásfejlesztőket a Pattern Recognition – vagyis az alakfelismerés, mintaazonosítás – problematikája. Jelen esetben tekintsünk el a háromdimenziós elrendezésektől és maradjunk a kétdimenziós objektumok területén, azon belül is egy speciális, de napjainkban már igen széles körben elterjedt és használt technológia és megoldás, az optikai karakterfelismerésnél. A látó élőlények a szem által közvetített képet használják alakfelismerésre. A számítástechnika is ezt a sémát követi, amikor valamely képfellevő eszköz előállította kép tartalmát igyekszik felderíteni, azonosítani. Tudjuk, hogy a gépi mozgókép állóképek sorozata, így nyugodtan szorítkozhatunk az állóképek vizsgálatára, különös tekintettel az optikai karakterfelismerésre, amely betűk, számok, jelek – összefoglalóan karakterek –, illetve az ezek által alkotott szövegek felismerésére irányul.

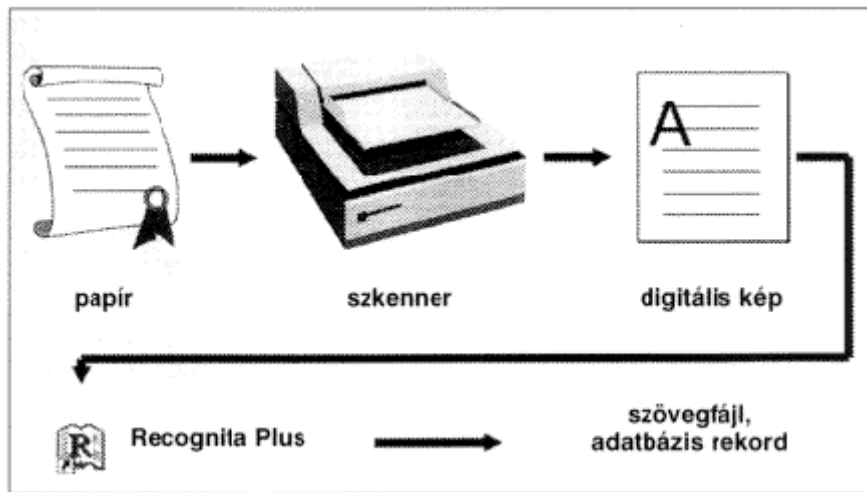
Mivel számítógépes megoldásról van szó, a számítógép pedig számokkal dolgozik, olyan képalkotó, illetve képbeviteli eszközre van szükségünk, amely digitális képet ad. Ilyen eszközök a digitális szkennerek vagy lapolvasók és a digitális kamerák. A 80-as évek elején-közepén, amikor az akkori idők egyik vezető intézménye, a Számítástechnikai Kutatóintézet és Innovációs Központ (SZKI) matematikai, valamint elméleti laboratóriumaiban megkezdődtek a mesterséges-intelligencia kutatások, különböző képfeldolgozási eljárásokat is kidolgoztak. Szükségessé vált ugyanis az egyre nagyobb számban érkező műholdképek képi elemzése. Itt először katonai célú feladatok voltak, később azonban már szerepet kaphatott olyan polgári alkalmazás is, mint egy nagy kiterjedésű gabonátábla műholdképéből megállapítani az esetleges betegségeket, rovarkárokat. Hamarosan megkezdődtek a kísérletek analóg kamerák digitalizált képének feldolgozására is. Tekintettel azonban a nyolcvanas évek közepén beszerezhető kamerák és képdigitalizálók műszaki színvonalára, a generált képek felbontása, részletgazdagsága – vagy mondhatjuk inkább, részletzegénysége – miatt csak viszonylag egyszerű feladatokat lehetett megoldani, például, hogy megkülönböztessünk egy kört egy függőleges vonaltól.

Igazán komoly lehetőséget a lapszkennerek megjelenése adott. A személyi számítógépek dinamikus fejlődésével és elterjedésével a nyolcvanas évek végére létrejött számítástechnikai környezet (300 dpi [pont/hüvelyk] felbontású szkennerek, 16 bites személyi számítógépek) megfelelő alapot adott a már jól használható karakterfelismerő technológia létrehozásához. A különböző alakfelismerő feladatok megoldásának az SZKI-ban logikus folytatása volt a karakterfelismerés, mint problémakör kutatása, piacképes technológia és arra épített termék kidolgozása, folyamatos továbbfejlesztése.

Ismerkedjünk meg néhány, az iparágban használatos fogalommal

- A karakterfelismerést általában az OCR betűszóval jelölik (az angol Optical Character Recognition rövidítése). Ezt használhatjuk általánosságban is, de leginkább a géppel nyomtatott, írógéppel írt szövegek felismerését értjük alatta.
- Az OMR – Optical Mark Recognition – speciális jelek, jelölőnégyzetek tartalmának felismerését szolgáló technológia neve.
- Az ICR – Intelligent Character Recognition – pedig a kézzel írt „nyomtatott” nagy- és kisbetűk felismerési technológiája.

Miután egy, a képfeldolgozásban jobban elmélyült egyetemi hallgató néhány havi munkával képes (mondjuk, diplomamunkaként) működő OCR-prototípust létrehozni és az így készült programok egész jó teljesítményt tudnak nyújtani valamely, előre definiált karakterkészleten, joggal hihetnénk, hogy a használható OCR létrehozása nem túl nagy feladat. A gyors kezdeti siker után azonban évek kemény munkája következik, ha a való élet dokumentumainak széles skáláját figyelembe véve akarunk jól használható megoldást készíteni.



1. ábra. Az OCR (Optical Character Recognition) technológia

Miről is van itt szó? Szkenneljünk be (olvassunk be), mondjuk, egy A4-es szövegoldalt, melyen a szövegen kívül természetesen lehet kép, grafika stb. (1. ábra). Számítógépünkön megjelenik az oldal digitális képe (image). Sokan nem ismerve a lényegét, úgy gondolják, hogy a dolog ezzel rendben is van, és csodálkoznak, amikor a képen szereplő szöveget nem tudják azonnal szövegszerkesztőjükkel feldolgozni. Nem is, hiszen ami e pillanatban ott van, az képpontok halmaza (egy karakternek kb. 25–30 x 40–50 képpont, pixel felel meg), mely halmaz alkotta kép lehet fekete-fehér, szürke skálás vagy színes. Ha a kép fekete-fehér, egy pont jellemzésére elegendő 1 bit, ha szürke skálás, kell hozzá 1 byte (8 bit), ha színes, akkor pedig 3 byte a vörös-zöld-kék (RGB) színek meghatározásához. Mielőtt a karaktert felismerne, egy sor lépést kell megtennünk, amit összefoglalóan kép-előfeldolgozásnak nevezünk. Ahhoz, hogy az OCR-felismerő algoritmus (vagy még inkább algoritmusok gyűjteménye) jó munkát végezzen, megfelelően pozícionált, csak a felismerendő szöveget tartalmazó tiszta képrészeket kell kapnia. Lehet, hogy a lapot véletlenül fejjel lefelé tettük be a lapolvasóba. Ezt egy speciális, a szövegsorokat alkotó képfoltokat elemző program felismeri, és a képet 180 fokkal elforgatja (rotáció). Amíg ez az automatizmus nem volt, a portré-tájkép-választást is „kézzel” kellett megoldani. Ugyancsak előfordulhat, hogy a lap nem pontosan pozícionálva kerül az olvasóba, ilyenkor úgynevezett ferdeségkorrekcióra (deskew) van szükség. Mindkét eset különösen gyakori lehet a vak felhasználóknál, akik számára az OCR az egyetlen lehetőség ahhoz, hogy mindennapi nyomtatott (nem Braille) dokumentumokat külső segítség nélkül olvashassanak.

A most már megfelelően pozícionált képről ezután el kell távolítani az oda nem illő, például szennyeződések okozta foltokat, pöttyöket. E feladat különösen kritikus, ha arra gondolunk, hogy az ékezeteket, jeleket kifejezetten káros lenne eltávolítandó objektumoknak tekinteni.

A következő lépés az egyes képrészek, vagyis a lapszerkezet azonosítása. Meg kell határozni, mi szöveg és mi nem szöveg (kép, fotó, grafika). Következik annak meghatározása, hogy a felismerő motor (OCR engine) milyen sorrendben fogja az egyes szövegrészeket, mondjuk, egy szövegszerkesztőnek átadni (gondoljunk itt például a többhasábos lapszerkezetre).

A lapszerkezet elemzése során meg kell találni az OCR-szempontról érdekes részeket. A szöveges dokumentumok egy- vagy többhasábosak, egy hasábon belül a szöveg bekezdésekre (paragrafusokra) tagolódik. A paragrafus jellemzője lehet, hogy első sora esetleg beljebb kezdődik, az utolsó sora pedig nem ér ki a sor végéig. A hasábok meghatározására viszont az őket elválasztó fehér terület szolgál. Két hasáb között általában néhány karakter szélességű fehér függőleges csík van, a program ezt a területet viszonyítja a mellette található sokkal szélesebb, fekete-fehér képpontokat tartalmazó területhez. A hasábos lapszerkezet felismerése nagyon fontos, mert egy felismert szöveget nem lehet úgy továbbadni, hogy például az első hasáb első sorát néhány szóköz után követi a második hasáb első sora. A hasábokat a felismerés után az olvasási sorrendnek megfelelően kell egymás után illeszteni.

A hasábokhoz hasonlóan oszlopokat tartalmaznak a táblázatok is, ezek olvasási sorrendje azonban más. A táblázatokat a függőleges oszlopok ellenére vízszintesen kell olvasni, szigorúan megtartva az egymás alatt elhelyezkedő karakterek pontos pozícióját. Itt az egyes szövegrészek „igazítása” segíthet: a hasábok általában egyidejűleg jobbra-balra, a táblázat oszlopai pedig csak balra, középre vagy csak jobbra igazítottak.

A szigorú tördelésű dokumentumoknál, például a verseknél – minden sor külön paragrafusként kezelendő, így az automatikus szerkezet-felismerésre elég kicsi az esély; tehát a programnak lehetőséget kell adnia a kézi beavatkozásra.

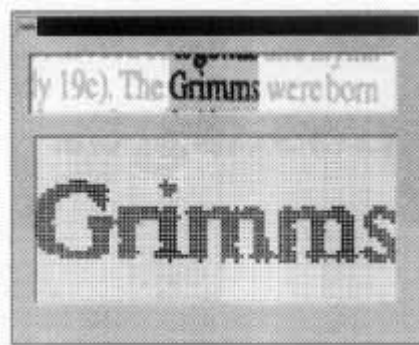
A felismerés során nem csak a karakter kódjára vonatkozó információkat kapjuk meg, sokat megtudunk a karakter attribútumáról (dőlt, aláhúzott, kövér, betűméret stb.) is, sőt a pontos elhelyezkedéséről a lapon. Ezek az adatok a kódokkal együtt átadhatók egy intelligens szövegszerkesztőnek, így a lap teljes formátuma tovább él, azaz minden a helyén van, csak éppen a szöveges képrészek már szöveggé szereznek.

A karakterfelismerés néhány buktatója

Az OCR-folyamat első lépése a szegmentálás, amely nem más, mint az egyes karaktereket alkotó képpontok csoportosítása, összerendelése. Első ránézésre a feladat egyszerűnek tűnik, vagyis az egymással közvetlen kapcsolatban lévő képpontok alkotnak egy karaktert. De gondoljunk a több különálló rész alkotta elemekre, mint a kettőspont vagy a felkiáltójel, nem beszélve az ékezetes karakterekről. A gondot csak fokozza, ha a beszkenelt kép túl világos. A karakterek vonala ilyenkor elvékonyodik és a karakterek széttöredeznek (2. ábra). Ellenkező esetben, ha a kép sötét vagy „kövér”, fonttal állunk szemben, az egyes karakterek összeérhetnek (3. ábra). A legrosszabb a kettő kombinációja. A 4. ábrán látható képet olvashatjuk 3000-nek vagy akár 3(XX)-nek is.

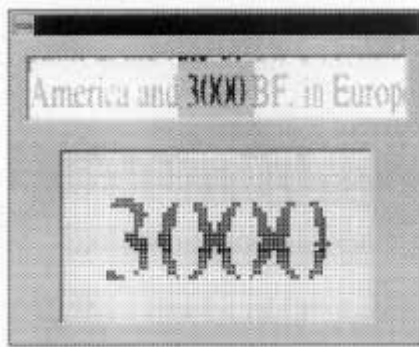


2. ábra. Több részre tört karakterek

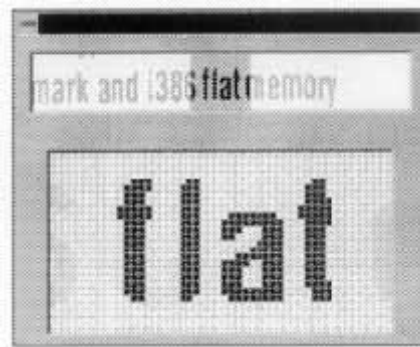


3. ábra. Egymáshoz érő karakterek

A következő lépés a *feature extraction* (jellemzők kigyűjtése): numerikus értékek sorozatát (tulajdonságvektor) rendeljük a karakter alakjához. A legegyszerűbb eset, amikor egyesek és nullák kétdimenziós elrendezésével képezzük le a karakter képpontok geometriai elhelyezkedését. Ez, amit kezdetben az OCR pionírok is használtak, a matrix matching (mintaillesztés) algoritmus. Noha egyes esetekben ma is nagyon hasznos lehet ez a megközelítés, a ma használt algoritmusok méret- és fontfüggetlen tulajdonságokat kezelnek: például a görbületek, hurkok száma, jellemző pontok helyzete, valamint más topológiai és statisztikai jellemzők. Az ideális tulajdonságvektornak két, egymással konfliktushelyzetben álló követelményt kell teljesítenie. Tudnia kell megkülönböztetni egymástól a hasonló alakú karaktereket (például 5 és S, C és G), miközben rugalmasan kell tudnia kezelni az azonos karakterek különböző variációit (például különböző fontoknál). Végül is ezeket a követelményeket nem lehet mindig maradéktalanul teljesíteni. Vannak karakterek, amelyek csak néhány képpontban különböznek egymástól (mint a t, az f, lásd az 5. ábrán). Az ilyenfajta eseteknél karakter-specifikus szabályokat kell alkalmazni. A szegmentálási problémák, valamint a hasonlóságokból adódó gondok a tipikus forrásai az OCR-hibáknak.



4. ábra. Tört és egymáshoz érő karakterek kombinációja, ez a legrosszabb



5. ábra. Karakterek néhány pixel különbséggel

Varázsszerek

Annak érdekében, hogy csökkentsük a széttöredezett, illetve egymással összeérő karakterek okozta gondot, beavatkozhatunk – már az elején – a képalkotási folyamatba, meghatározva, hogy egy szürke skálás képen hol legyen az a vágási szint, amely végül is eldönti, hogy egy szürke képpontot a továbbiakban fehérnek vagy feketének fogunk-e tekinteni, vagyis alkotó eleme lesz-e egy karakternek

vagy sem. A vágási szintet állíthatjuk manuálisan (vakok számára azonban ez nem lehet megoldás), de sokkal jobb eredményt érhetünk el azzal a kifinomult technikával, amely figyelembe véve az egyes képrészletek különböző megvilágítását, egyetlen oldalon belül is képes más-más vágási szinteket meghatározni a legjobb eredmény elérése érdekében. Ez a kifinomult technika sem segít azonban olyan esetekben, ahol már az eredeti dokumentumon például festék van egy helyen, amikor keskeny hézagnak kellene lennie. A többszörösen másolt dokumentumokon vagy faxokon található leggyakrabban ilyen folt.

Azzal, hogy egy szűrkeskálás képből bizonyos megfontolásokkal fekete-fehér képet csinálunk, egy csomó, a képben lévő információt eldobunk. A kínai Tsinghua Egyetemen kísérleteket végeznek a vágás nélküli feldolgozásra, amikor is közvetlenül a szűrkeskálás kép alapján lehet meghatározni a karakterek legvalószínűbb alakját.

Noha egyfolytában karakterfelismerésről beszélünk, a speciális alkalmazási területeket leszámítva (például kevés adatot tartalmazó formanyomtatványok) valójában szöveget akarunk felismerni, így komoly segítségünkre lehetnek a nyelv-specifikus információk. Esetünkben azonban a lingvisztikai segítség kétélű fegyver, nem megfelelő alkalmazása félreviheti a felismerést. Az egyes OCR-megoldások, nem tudván a hagyományos képfeldolgozó eszközökkel tovább növelni a felismerés pontosságát, kivétel nélkül szótárak és spell-checking (helyesírás-ellenőrző) modulok támogatását veszik igénybe a felismerési folyamatban. A legkevesebb, hogy a helyesírás-ellenőrző rámutat a nem megfelelő szavakra, azonban az OCR-szoftverek ennél tovább mennek, a fel nem ismert karaktereket a legvalószínűbb megoldást jelentő szóval helyettesítik. Lehetnek viszont szép számmal szavak, amelyeket a szótárak, illetve helyesírás-ellenőrzők nem tartalmaznak, és persze elég nehéz azt automatikusan eldönteni, hogy vajon ezek korrekcióra szorulnak-e. Az angolban például sok rövid, számos betűvariációban létező szó van, így itt egy-egy rosszul felismert karakter könnyen túlélheti a nyelvi korrekciót. A másik probléma, hogy az OCR-hibák tipikusan csoportokban jelentkeznek, gyakorta a szó betűszáma is más, mint az eredetiben, így különösen nehéz automatikus becsléseket tenni. Tekintettel az ilyen típusú gondokra, csak az vezethet megfelelő eredményre, amikor szoros kapcsolat van a szó képe és a nyelvi információ között.

Noha az OCR-programok egyre pontosabbak, el kell fogadnunk, hogy akadnak hibák. Attól függően, hogy az egyes algoritmusokat hogyan írták meg, más-más OCR-ek más-más típusú hibákat vétének, és azokat következetesen. Erre alapozható az a technika, amely segít a felhasználónak a hibák gyors kijavításában. Amint a felhasználó kijavított egy hibát a dokumentum elején, a program végigmegy az egész dokumentumon, és automatikusan kijavítja az azonos típusú hibákat. Itt a tanuló- és a javítófunkció dolgozik. Ezzel végül is eljutottunk oda, hogy bár egy OCR-program legfontosabb jellemzője a felismerési pontosság, hasonlóan lényeges lehet a használat hatékonysága, amely a szkenneléstől a kész dokumentumig terjedő időt jelenti, magában foglalva értelemszerűen a szkennelésen, felismerésen kívül a javításra fordított időt.

Napjainkra az OCR meglehetősen kiforrott technológia lett, s könnyen hozzáférhető mindenki számára. A szkennerek mindennapi eszközzé váltak, és ma már a legolcsóbb szkennerekhez is adnak valamilyen OCR-t, amely a legegyszerűbb igényeket képes is kielégíteni.

A Recognita OCR-technológiája – egyike a világ legjobbjaival – egyedülálló abban a tekintetben, hogy a világon a legtöbb nyelv karakterkészletét képes felismerni. Ez a 114-féle nyelv az összes latin, görög és cirill betűs nyelvet jelenti. Ha figyelembe vesszük, hogy mindinkább elektronizálódó világunkban milyen fontos, hogy a korábban csak papíron meglévő szövegek elektronikus formában tárolódjanak, beláthatjuk az ilyen technológia kulturális jelentőségét. Különösen a kis nyelvek esetében – ahol a korlátozott piaci méret üzletileg nem teszi megalapozottá saját technológia kifejlesztését – tölthet be a

nemzetközi, soknyelvű OCR kulturális missziót is. Mivel a nyelv a nemzeti kultúra alapja, igen fontos, hogy az irodalom, publicisztika, de nemkülönben a tankönyvek, az oktatási segédanyagok elektronikusan terjeszthető, visszakereshető formába kerüljenek, és akkor nem beszélünk még a vak, látássérült, diszlexiás emberekről, akik e technika nélkül el vannak zárva a nyomtatott információ elől.

A Recognita helyzeti előnyét azzal szerezte, hogy már induláskor a soknyelvű Európában gondolkodott és az OCR-programban a belső karakterábrázolást két byte-ban valósította meg (egy byte a karakter „testnek”, egy másik az ékezetnek) garantálva a karakterek elegendő mennyiségét, míg az elsősorban amerikai versenytársak lényegében csak az amerikai piacot, így az angolt, valamint a nagy nyugat-európai nyelveket tekintették irányadónak, és megelégedtek 1 byte-tal, s e korlátozás miatt később sem lehetett elmozdulni a sok nyelv irányába.

A Recognita mint üzlet

A Recognita, mint üzleti vállalkozás létrehozásában hárman játszottak főszerepet: *Kovács Emőke* és *Marosi István* fejlesztői munkájának eredményeként jött létre az a technológia és a ráépülő termék, amelyre alapozva az üzleti vállalkozás beindulhatott, *Tállai Benedek* marketing-szakértelme pedig életre hívta magát a vállalatot, és útjára indította a szakmai eredmények mellett a Recognitát, mint sikeres üzleti vállalkozást.

A nagy politikai, gazdasági változások hajnalán, 1989-ben alapította a Recognita részvénytársaságot az SZKI tulajdonosként, úgy, hogy egy angol üzletembert társtulajdonosként is bevont. Utóbbira azért volt szükség, mert az akkori kormányzat ötévi társaságiadó-mentességgel és további öt évre nagymértékű adókedvezménnyel jutalmazta azokat a frissen alapított vállalatokat, amelyekben a meghatározott mértékű tőkenagyság mellett a külföldi tulajdonosi részvétel adott mértéket elért. A kedvezményhez kellett még, hogy a vállalat stratégiai fontosságú területen működjön, a szoftverfejlesztés, -gyártás pedig stratégiainak minősült. A Recognita filozófiájához tartozott, hogy a két főrészvényes mellett a vezetés és az alkalmazottak is birtokoljanak részvényeket. A Recognitát tehát a megfelelő feltételekkel alapították, így mivel működésének az első öt évében folyamatosan számottevő nyereségre tett szert, a kormányzat által adott adómentességet kitűnően ki tudta használni.

Az üzletet viszonylag akadálytalanul beindították, ugyanis abban az időben a piac meglehetősen üres volt, különösen Európában nem volt erős konkurencia. Így a szerény tőkeellátottságú vállalat kevés marketingdollárja is elegendő volt ahhoz, hogy két-három éven belül kiterjedt „láthatóságot”, vagyis piaci ismertséget érjen el a Recognita márkanév számára. A cég- és a márkanév azonossága pedig segítette abban, hogy egyidejűleg az előbbi is ismertté váljon. Mivel a technológia a maga soknyelvűségével eleve feltételezte a nemzetközi piaci jelenlétet, a cég fő terméke – a RecognitaPlus – a kezdetektől főleg exportra készült, ami többek között abban nyilvánult meg, hogy először a termék angol nyelvű változata készült el, majd a piaci kapacitás függvényében a német változat, és csak azután jöhetett a magyar. A leghatékonyabb marketingeszköznek a külföldi szakfolyóiratokban megjelenő méltató cikkek, a konkurensokkal való összehasonlító elemzések bizonyultak. Ezen kívül a vállalat marketing erőforrásait elsősorban a szakkiállítási részvételre koncentrált: a két legfontosabb az egyre nagyobb szerepet játszó hannoveri CeBIT és a Las Vegas-i Comdex volt.

A cégalapítást követő két évben 1990-ben és 1991-ben a Recognita két leányvállalatot alapított fő piacain, Németországban és az Egyesült Államokban. A német vállalatot éppen az egyesítés előtt jegyezték be Lipcsében, így még nagyon gazdaságosan (NDK-s költségszinten) lehetett létrehozni. Az amerikai vállalat székhelyét a Recognita a Szilícium-völgyet választotta, azon belül Sunnyside városát. A

német vállalat célja a piaci jelenlét erősítése volt Európa legnagyobb felvevőpiacán, szervezve a marketingakciókat, a sajtókapcsolatokat, a termékelosztást, és megadva az eladás előtti és utáni szolgáltatásokat a vevőknek. Az amerikai vállalat feladata viszont elsősorban az átfogó partnerkapcsolatok kiépítése volt, továbbá az, hogy a szilícium-völgybeli személyes kapcsolatok ápolásával „első kézből” gyűjtsék össze az információt, amit ha jól csinálnak, komoly gazdasági előnyöket eredményezhet. Mindemellett az volt a nem titkolt szándék, hogy a leányvállalat értékelhető (3% fölötti) piaci részesedést szerezzen Amerikában.

A kilencvenes évek elején a Recognita volt az egyetlen „önálló” magyar kiállító a CeBIT-en, és sokáig az egyetlen magyar kiállító a legrangosabb amerikai informatikai vásáron, a Comdexen. Ezzel a cég egyfajta küldetést is teljesített, mert a szakmai közvélemény a Recognita jelenlétét magyar jelenlétként is értékelte.

1994 fordulópont volt a Recognita életében, ugyanis mind a külső, mind a belső feltételek jócskán megváltoztak. Ekkorra az amerikai leányvállalat globális aktivitásának eredményeként a vállalat számos stratégiai fontosságú partnerszerződést kötött amerikai és távol-keleti szkennergyártókkal, akik lényegében az európai piacot is lefedték (Hewlett Packard, Canon, Microtek, Mustek, Kye Genius stb.). Ezek a szerződések az úgynevezett OEM + upgrade üzleti modell alapjai. E modell szerint a Recognita rendkívül alacsony darabár mellett licencjogot ad az „eredeti” szkennergyártónak, az OEM-nek (Original Equipment Manufacturer). A szkennergyártó e licencjog alapján minden készülék mellé letesz egy Recognita OEM-verziót (Recognita Select, Recognita Standard), mely a csúcsmo­dell Recognita Plus csökkentett funkciókészletű változata. Amikor tehát a végfelhasználó megveszi a szkennert, talál benne végeredményben egy használható megoldást adó Recognita-programot, felismerési pontossága megközelíti a csúcsmo­dell ebbéli képességét, de egy sor kényelmi funkció, a használatot segítő tulajdonság hiányzik belőle. A termék egyszerű használat közben rendszeresen felszólítja a felhasználót, hogy regisztráljon, azaz juttassa el adatait a Recognita vállalatnak, másrészt felhívja a figyelmet arra, hogy ezen OEM-változattal a termék birtokosa jogosult a csúcstermékekre fellépni, upgrade-elni (frissíteni). Ez azt jelenti, hogy a teljes árnál lényegesen kedvezőbb áron juthat a legtöbbet tudó változathoz. Ez az OEM + upgrade modell robbanásszerűen kezdett elterjedni 1994-ben. A nagy amerikai versenytársak (Caere, Xerox) mind erőteljesebben aktivizálódtak Amerikán kívül is. Hatalmas harc kezdődött a szkennergyártókért, a harcban a Recognita egyre-másra alulmaradt. Már csak elvétve sikerült az egész világra szóló szerződést kötni, de abból is kimaradt Amerika mint terjesztési terület. A szerződések hatálya földrajzilag azokra a területekre szűkölt (kelet-európai, görög, török, valamint kisebb nyugati nyelvek), ahol a Recognita monopolhelyzetben volt. Azt, hogy a nagy versenytársak a Recognitával szemben terjeszkedni tudtak az OEM-partnereknél, nem technológiai, hanem marketingfő­lényüknek köszönhetőek.

Ezzel párhuzamosan megkezdődött a piaci szereplők átrendeződése. Az egyik legjobb technológiájú amerikai Expervision például rossz marketingdöntésekkel egyszerűen padlóra küldte magát. A Caere egy jól irányított fúzióval eltüntette az OEM-üzletben számára veszélyessé váló, szintén amerikai Calerát. Oroszországban a semmiből jött elő és gyors növekedésnek indult a jól átgondolt, sikeres technológiai megoldásokat gyártó Bit (jelenleg Abbyy). A Recognita legfontosabb piaci területére, Nyugat-Európába egyre nagyobb tömegben lovagoltak be a szkennerek hátán a versenytársak: a Caere terméke, az OmniPage, és a Xeroxé, a TextBridge. Az éleződő konkurenciaharc erősítette a Recognita vállalaton belüli problémáit, és kikényszerített egy sor, a jövőt szem előtt tartó döntést.

Időközben a tulajdonosi összetétel is megváltozott. 1993-ban az angol tulajdonos halálát követően üzletrészt egy amerikai kockázatitőke-társaság, a Magyar–Amerikai Vállalkozási Alap (MAVA) mint *pénzügyi befektető* vette át. Ekkor a tulajdonosi megoszlás a következő volt: SZKI 50 százalék, MAVÁ 30

százalék, alkalmazottak 20 százalék. A kockázati tőke természetrajzáról annyit érdemes tudni, hogy általában érzelemmentes, még kisebbségi jelenléte is erőteljesen hat az adott vállalatban, katalizálja a vállalat fejlődését, és arra készül, hogy előbb-utóbb nagy haszonnal túladjon befektetésén. 1994-ben, amikor az SZKI vagyonát az ÁVÜ privatizálta, a MAVÁ kedvező áron meg tudta szerezni az SZKI 50 százalékos részesedését, és ezzel tulajdonosi hányadát 80 százalékra tornázta fel. Ugyancsak 1994-ben, az év elején vette át a vállalat irányítását Reszler Ákos, e sorok írója. Első szempontom az volt, hogy ne ártsak, ne rontsam el ezt a híres magyar szoftver fenomént, a Recognitát. Az évben volt egy Siker című rendezvény a Budapesti Kongresszusi Központban, ahol mindenki arról beszélt hogyan építsük fel, éadjuk el a sikert, én sikert viszont egy olyan előadással arattam, amelynek címe ez volt: "Hogyan NE rontsunk el egy sikeres vállalkozást!"

1995 és 1996 meglehetősen ellentmondásos évek voltak a vállalat életében. Egyrészt a pénzügyi mérleget tekintve a vállalat katasztrofális helyzetbe került, ugyanis e két év alatt kemény veszteségek mellett teljes vagyonát elvesztette, másrészt a vezetés minden stratégiai elképzelése bejött. Rendkívül intenzív technológiafejlesztés eredményeként 1995 végére a Nevada Las Vegas Egyetem felmérésének eredményeként a Recognita a Xerox-szal együtt az első helyen végzett a világ OCR-technológiáinak versenyében. Egy, az amerikai védelmi minisztérium által pénzelt kutatócsoport a Las Vegas-i egyetemen évről évre nagyon alapos összehasonlító elemzésnek vetette alá a világ OCR-termését. Ezzel a teszttel a Recognitának szerencséje volt, mert így egy független fórum nagyon alapos, több hónapos, sokféle és több millió karakteren alapuló teszt eredményeként meglehetősen megbízható rangsort tudott felállítani a termékek között. Persze keserves volt látni, mondjuk az ötödik helyezést, és innen meggyőzni egy OEM-partnert, hogy velünk kössenek szerződést és ne a jobbakkal, de ezek a részletes elemzések a hibákat is pontosan megmutatták, így már „csak” ki kellett azokat javítani. 1996-ra a Recognitának tehát világszínre számító technológiája volt.

A nagyon céltudatos intenzív technológiafejlesztéssel egyidejűleg egy másfél éves folyamat eredményeként, 1996 végére a cég kiépítette és bevezette, azaz a mindennapi munkában alkalmazta azt a nyugat-európai minőségbiztosítási rendszert, amelyre ugyanez év végén megkapta az ISO 9001 minőségbiztosítási tanúsítványt; megszerzése azonban csak része volt a vállalati minőségpolitikának. A legfontosabb feladat olyan belső rendszer megteremtése volt, amelynek segítségével a fejlesztési és más folyamatok professzionális módon kézben tarthatók, azaz tervezhetők, követhetők, szükség esetén korrigálhatók. Itt igen fontos szerephez jutott a hazai szoftvervállalatoknál még egyáltalán nem általános folyamat- és termékdokumentálás.

1996-ra kimerültek a vállalat pénzügyi lehetőségei, viszont volt világszínre technológiánk, és a legigényesebb minőségi követelményeknek megfelelően szervezett csapat. Megérték tehát a feltételek a stratégiai, azaz hosszú távú, úgynevezett *szakmai befektető* bevonására. Megkezdődtek a tárgyalások japán, német és amerikai vállalatokkal, és 1996 végére létrejött a megállapodás az amerikai Caere Corporationnel, a Recognita legnagyobb versenytársával a fúzióra: a Recognita 100 százalékban a Caere Corp. európai leányvállalata lett. Ez annak idején az év sikertörténete volt, nemcsak azért, mert egy amerikai vállalat először vásárolt meg 100 százalékban magyar szoftvervállalatot, hanem mert a vételár nagyon sokat elárult arról, ahogy a vevő a Recognitát értékelte. A korábbi ügyletekben a részvények névértékük 100, illetve 138 százalékában cseréltek gazdát, ennél azonban a vevő a névérték 750(!) százalékát fizette a részvényekért. Mindezt egy olyan pillanatban, amikor a vállalat könyv szerinti értéke, saját vagyona a nullánál is kevesebb (negatív) volt. Sokan aggódtak akkor a Recognitáért, mert „a nagyhal megeszi a kishalát” hatása érvényesült a Calera felvásárlásakor, azaz a Caere nyomtalanul magába olvasztotta a Recognitánál jóval nagyobb piaci versenytársát is. A fúziót/akvizíciót követő évek azonban bizonyították, hogy itt egészen másról volt szó. A Recognita megmaradt magyar vállalatnak, nem kerültek ide amerikai vezetők. A vállalat saját műszaki, minőségi és szervezeti kultúrája erősödött, az

anyavállalattal való kölcsönhatásban nemcsak a magyarok tanultak és gazdagodtak, hanem sok minden átkerült az anyavállalat életébe, valamint beépült termékeibe. A budapesti projektek eredményeként előálló termékeket a Caere világméretű elosztási csatornáin keresztül értékesítette, lényegesen nagyobb bevételt elérve, mint korábban az az önálló Recognitánál jellemző volt. Így a magyar szoftvermérnökök munkája által hozott bevétel ugyanolyan mértékű lett, mint amerikai kollégáiknál. Azaz elmondhatjuk, hogy itt a magyar munka termelékenysége elérte az amerikai színvonalat, sőt, tekintve, hogy a magyar munka sajnos máig lényegesen olcsóbb, a mi munkánk termelékenysége magasabb is lett az amerikaiénál. A Caere-rel kötött házasságot követő 3-4 évben, kihasználva az anyavállalaton keresztüli piaci lehetőségeket, a Recognita ugrásszerűen megnövelte bevételeit, és a hazai iparági átlagnál lényegesen nyereségesebb lett. Így nem csak elvesztett részvénytőkéjét tudta visszatölteni, 1999 végére az úgynevezett „saját vagyona” a korábbiak a többszöröse lett.

A vállalat teljesítményének elismerése rangos díjakban is megnyilvánult. A cég egymás után háromszor elnyerte az európai informatikai díjat, a „The European IT Prize”-t. Az elismerésért vívott versenyben az EU országain kívül a kelet-, és közép-európai országok, valamint Izrael vállalatai is indulhattak. 6–800 pályázó volt évente, és 25 pályamunkát díjaztak. A Recognita az egyetlen vállalat, amely képes volt ezt a díjat háromszor is elnyerni.

Díjazott termékek:

- Recognita Form (1996): kitöltött formanyomtatványokat feldolgozó program,
- Recognita Reader (1997): speciálisan vakok számára készült OCR-program,
- Recognita Plus (1998): általános célú OCR-program.

A ScanSoft – Nuance korszak

2000-ben a Microsoft elhatározta, hogy OCR-technológiát épít az Office-ba, ehhez azonban nem a piacvezető Caere/Recognitát, hanem a Xerox technológiáját választotta. Ez a momentum időlegesen drámai átrendeződést eredményezett a tőzsdén, a Caere-részvények leestek, a Xerox Imaging Systems leányvállalat utódjaként tőzsdére lépett ScanSoft részvényei pedig felmentek. Így történhetett, hogy a kishal megette a nagyhalat, a lényegesen kisebb ScanSoft felvásárolta a Caere-t.és itt következett a Recognita újabb sikere, a ScanSoft vezetés megismerve és felismerve a budapesti fejlesztői csapat szakmai kvalitásait, és az amerikaival összevetve előnyös költségeket, fokozatosan minden fejlesztői munkát megszüntetett Amerikában, és minden k+f munkát a Recognitának adott. Érdemes a nevet megemlíteni, Paul Ricci, aki a Caere-felvásárlást levezényelte nem sokkal később erőteljes stratégiai váltással látványos növekedési pályára állította a vállalatot. A képfeldolgozás (Imaging) jól beállt, nagyon profitábilis üzlet, de mint az ilyen érett korba jutott üzletek általában, nem képes produkálni azt a növekedési rátát, amelyet a tőzsde értékelhető árfolyamnyereséggel honorálna. Esetünkben a dinamikus növekedést produkáló új területet a nyelvi technológiák és termékek jelentik. Egy csődaukción a ScanSoft fölvásárolta a Lernout & Hauspie üzletét, benne beszéd-szintetizáló- és beszéd-felismerő megoldásokkal, köztük a már akkor is jól ismert Dragon programmal. A beszéd-üzlet belépésével a képfeldolgozásra utaló ScanSoft név zavaróvá vált, így egy semleges nevet kellett választani, ez lett a Nuance. Az elmúlt 8 évben a Nuance akvizíciók sorozatával erőteljes organikus fejlődés mellett 200 főről 5,000-re, és 60 millió \$-ról 1,300 millió \$-ra nőtt. Ebből az Imaging ma mindössze 10%-ot képvisel.

A Nuance minden imaging termékének fejlesztéséért a Budapesti Recognita-utód felel. A hajdanvolt 50 fős Recognita fejlesztményei évi 3 millió \$ forgalmat produkáltak, a mostani 150 fős csapat termékei 160 millió \$-t hoznak.

A zászlóshajó, a klasszikus OCR-termék a piacon legjobban bevezetett nevet, a Caere OCR termék nevét viseli, és ez az OmniPage. Benne megtalálható mindhárom felismerő motor, a Recognita, az OmniPage és a xeroxos TextBridge.

Irodai dokumentumkezelésre kifejlesztett program a PaperPort, amely szkennelési, szervezési, és kép- illetve dokumentum megosztási lehetőségeket kínál. A PDF Converter program igényes üzleti megoldás PDF fájlok létrehozására, konvertálására és a PDF-alapú információcserére.

Mit hozhat a jövő?

A jövő már itt van! A szkennerek robbanásszerű terjedése mellett belépett egy másik úgynevezett Still Image Capturing (álló-kép előállító) eszköz a digitális kamera. Nem beszélve a mobil eszközökben lévő kamerákról! Egy ilyen eszközzel akár az utcán lekapott falragasz szövegét is OCR-ezhetjük, amennyiben a program képes például háromdimenziós ferdeségkorrekcióra, valamint a digitális fényképezőgépek készítette alacsonyabb felbontású kép megfelelő feldolgozására. Itt az OCR-program nem az eszközben, hanem valamilyen távoli kiszolgálón "ül". A kép elmegy a szervernek, az feldolgozza, és a szövegfájlt visszaküldi. Kitűnő terep Cloud-os alkalmazásoknak.

Mintegy két évtizedig a Nuance-Recognita PC-s, azaz desk-top alkalmazásokkal volt jelen. 2010 óta működésünket kiterjesztettük szerveres – hálózatos irányba is. A holland X-Solutions és az amerikai eCopy cégek felvásárlásával kiértelt technológiákhoz, termékekhez, és jól felépített piacokhoz jutottunk. A fejlesztést mindkét helyről áthoztuk Budapestre. Multifunkciós (MFP) eszközök gyártóival (Xerox, Canon, HP, Ricoh, stb.) együttműködve vállalati elektronikus dokumentumkezelő rendszerekhez kapcsolódunk ShareScan nevű termékünkkel.

Epilógus

Az OCR (mint üzlet) legnagyobb gondja, hogy a piac mérete – legalábbis ami a bevételi lehetőségeket illeti – nem növekszik. Hiába nő a szkennerek száma robbanásszerűen, egyre nagyobb lehetőségeket teremtve az OCR-nek, a szoftver árcsökkenése „feléli” a darabszám növekedést, vagyis, a darabszám és az eladási ár szorzata lényegében stagnál. Ez a piaci szereplők körében egyre erősödő koncentrációt eredményez. A régebbiek sorozatos egyesülések árán igyekeznek növekedni, új szereplőknek pedig egyszerűen nem éri meg belépni egy ennyire beállt piacra.

Az OCR, mint technológia erősen kötődik a papírhoz, mivel alapvetően papír alapú dokumentumok feldolgozását célozza. A világ pedig rohamtempóban elektronizálódik. Az e-mailek világában ma már kinek jut eszébe levelet vagy faxot küldeni. Ha csak ezt néznénk, abba is hagyhatnánk minden további erőfeszítést. A helyzet azonban nem ennyire reménytelen. Az archívumokban, könyvtárakban heverő hatalmas mennyiségű feldolgozatlan anyag rosszabb minőségű anyagok feldolgozására is alkalmas OCR-technológiát igényel. Változatlanul azt mondjuk, hogy egy OCR legfőbb fokmérője a minél pontosabb karakter-(szöveg) felismerés. Ezen túl azonban egyre inkább igény a dokumentum szerkezethelyes, alakhelyes visszaadása is (azonos tördelés, azonos fontok stb.).

A PDF, mint "elektronikus papír" dokumentum, sok esetben már a szkennerek kimenetén is megjelenik, és ez a PDF Image is terepe az OCR-nek. A PDF-kép mögött ott lehet az OCR-ezett szövegfájl, mely lehetőséget ad a PDF-dokumentumban való keresésre (Searchable PDF).

Irodalom

Kovács E. – Marosi I. – Benesóczky M. – Lánckzy J.: Lessons in Character Recognition
Document World, 1996. July–August, p. 30–34.

Reszler Ákos: A Recognita, mint technológia és mint üzlet.
Természet Világa, 2000. II. különszám, Informatika