



Digitális képek keresése a weben

Bevezetés

A digitális képek rohamos szaporodása miatt megnőtt az érdeklődés ezek hatékonyabb tárolása, indexelése és visszakeresése iránt. Ugyanakkor kevésbé vizsgálták még azt a kérdést, hogy az emberek hogyan is keresnek képeket a weben? Mivel a web annyira univerzális, a felhasználók képkeresési stratégiájának megismerése hasznos támpontot adhat más multimédia-gyűjtemények keresőrendszereinek megtervezéséhez is.

A jelenlegi képkereső módszerek vagy fogalmi vagy tartalmi elvűek, vagy pedig e két megközelítést kombinálják. Előbbire jó példák a tárgyszavazott képgyűjtemények, melyek a vizuális információkra kidolgozott teauruszok valamelyikét használják: az egyes képeket az ezekből vett fogalmakkal látják el a gyűjteményeket indexelő munkatársak. Sok webes keresőrendszer ezt a módszert automatizálja úgy, hogy a képet vagy egyéb multimédia-állományt környező szövegből, illetve a fájlok nevéből készít kereshető indexeket. Ez a megoldás természetesen azt feltételezi, hogy ezek a szöveges információk relevánsak a multimédia-objektum tartalmával kapcsolatban. A valóságban ez nem mindig van így, hiszen egyes szoftverek, vagy például a digitális kamerák véletlenszerű neveket adnak az állományoknak. A szöveggörnyezet felhasználása viszont általában elég jó eredménnyel jár.

A tartalomelvi technikáknál a képpontok szintjén indexelik a digitális képeket és a keresés a pixelek összehasonlításán alapul. Ezeknél a rendszereknél a képeket szín, mintázat, forma és elrendezés szerinti hasonlóság alapján lehet visszakeresni; s ezt esetleg kiegészíti még egyes szöveges információkra (pl. a fájlba ágyazott leíró adatokra) való keresési lehetőség is. Ilyen technológiákat építettek be például a *MediaSite.com* nevű üzleti alkalmazásba, vagy a *WebSeek* és a *SingingFish* multimédia-webkeresőkbe.

A hibrid, vagyis fogalmi+tartalmi elven működő rendszerek esetében általában van egy tanulási fázis, amelynek során megtanítják a rendszert arra, hogy a pixelek elemzésével hogyan írja le a képek tartalmát, és ezután ezeket a géppel hozzárendelt tárgyszavakat lehet felhasználni a visszakeresésnél. Az ilyen képfelismerő rendszereknek köszönhetően a felhasználók a szöveges dokumentumokhoz nagyon hasonló módon tudnak katalógizálatlan képhalmazokban is keresni, tehát nem kell valamilyen speciális keresési metódust vagy szakzsargont elsajátítaniuk. Néhány kereskedelmi szoftver már használja is ezt a technikát különböző dokumentumok automatikus metaadatolásához (pl. *Verity K-2 Architecture*).

Ugyanakkor nem tudni, hogy ezek a megoldások illeszkednek-e a felhasználók valódi igényeihez, keresési szokásaihoz és jellemzőihez? A weben keresgélők értik-e vajon a fogalomalapú indexelési sémák működését? És a képpontelemzést használó, tartalomalapú keresőszolgáltatások találkoznak-e vajon a felhasználók információs igényeivel? Úgy tűnik, nincs sok együttműködés a manuális módszerekkel dolgozó indexelők és az automatikus képfelismeréssel foglalkozó informatikusok között. És a korábbi vizsgálatok azt is kimutatták, hogy a képi dokumentumok osztályozását végző szakemberek nem igazán olyan kulcsszavakat használnak, amilyeneket a felhasználók beírnak a tényleges kereséseik közben. Az is kiderült, hogy az emberek ritkán keresnek a képek vizuális jellegzetességei alapján.

A jelen tanulmányban bemutatott kutatás az egyik nagy webes keresőgépre beírt kérdések elemzésén alapul. A keresőkérdéseket összetettük három, régebben publikált, képekre vonatkozó osztályozási sémával, és megnéztük, hogy a webes keresések megoszlása mennyire tér el a korábbi, másfajta környezetekben végzett kutatások során kapott eredményektől, illetve hogy mennyire alkalmasak ezek a sémák a webes képkeresések osztályozására?

Szakirodalmi áttekintés

A képkeresőkkel foglalkozó vizsgálatok gyakran figyelmen kívül hagyják a felhasználói oldalt: például a szövegek és a képek utáni kutatás különbségeit, a képjellemzők fontosságát a keresésnél, és a képkereső kérdések sajátosságait; pedig ezek nagyon fontos kérdések, különösen a webes keresőrendszerek szempontjából. Például *Greisdorf* és *O'Connor* 2002-es tanulmányukban megállapították, hogy a felhasználók számára a találatként megjelenő képek relevanciája olyan fogalmi és tartalmi jellemzőktől is függ, amelyek nincsenek is rajta a képeken; továbbá hogy az érzelmeket kifejező szavak fontos keresőkifejezések a képek esetében. *Hertzum* 2003-ban publikált kutatása során egy nagy filmarchívumhoz érkezett e-mail kérések egy évnyi halmazát elemezte és azt találta, hogy a kérések beküldői nagyon sokféle jellemzővel írták le az igényeiket, de a levelek 43 százaléka nem tartalmazott kontextusra vonatkozó információt, vagyis hogy mivel kapcsolatban, milyen környezetben merült fel az adott igény. *Choi* és *Rasmussen* szintén 2003-ban a *Kongresszusi Könyvtár „American Memory”* nevű fotóarchívumában 48 kutató kéréseit gyűjtötte össze és négy csoportba sorolta ezeket: *speciális*, *általános*, *absztrakt* és *szubjektív*. A kérések 26.3%-a esett az első és 60.5%-a a második kategóriába.

A weben hatalmas mennyiségű kép- és egyéb multimédia-állomány van. Nem könnyű hatékony keresési stratégiákat kialakítani ezekhez, ugyanakkor például *Fukumoto* 2006-os cikkében arról számolt be, hogy a képeket keresők stratégiái rendszerint elég egyszerűek. *Jansen* és munkatársai az Excite webkeresőhöz érkezett kéréseket elemezték, amelyek hang-, video-, vagy képfájlokra vonatkoztak. Utóbbiak esetében egy felhasználó átlagosan 3.36 kérdést küldött a keresőgépnek és ezek 3.74 keresőszót tartalmaztak, melyek között nagy számban voltak egyedi, csak egy alkalommal előforduló szavak. A keresésre fordított átlagos idő és a keresőkifejezés is hosszabb a képek esetében, mint általában a webes kereséseknél, ez is jelzi, hogy a multimédia-információ megtalálása nagyobb szellemi feladat. Ugyanezt igazolta *Jansen* is 2003-as kutatása során, amikor nem az általános webkeresők, hanem az AltaVista-val társult multimédia-gyűjtemények használatát elemezte. A hangokra és a videókra vonatkozó, kevesebb mint három szóból álló keresőkifejezésekkel szemben a képek esetében átlagosan négy szót írtak be a felhasználók, sőt 28 százalékban még Boole-operátorokat is használtak, valamint hosszabb

ideig kerestek, mint más dokumentumtípusoknál.

Bár ezek a vizsgálatok világosabbá tették, hogy az emberek hogyan is keresnek vizuális információkat, de egyik esetben sem próbálkoztak a kutatók azzal, hogy a webes képkereséseket besorolják az ismert osztályozási sémák valamelyikébe, így az sem derült még ki, hogy ezek az osztályozások alkalmasak-e egyáltalán a webes keresésekhez. És arról sincs igazán információ, hogy alkalmazák-e őket a webes keresőgépek, segítve ezzel is a felhasználóikat abban, hogy releváns képeket találjanak.

Képkeresőgépek a weben

A nagy webkeresőket ugyanúgy használhatjuk a multimédia-dokumentumokhoz, mint ahogy a szöveges információk kereséséhez szoktuk. Ez egyrészt előnyös, mert nem kell új módszereket megtanulni, másrészt viszont nagyobb szellemi munka szükséges ahhoz, hogy megfogalmazzuk a megfelelő szövegekörnyezetet, amelyben a keresett multimédia-fájl valószínűleg előfordul: vagyis egy nem szöveges információt szöveges keresőkérdéssel kell átalakítanunk. Egyes keresőgépeknél kiegészítő nyomógombok vagy fülek jelennek meg, ha multimédia-tartalomra akarunk keresni. A ClipArt Searcher vagy a WebSeek esetében pedig a keresőmezőn kívül nagyobb tematikus kategóriák közül is választhatunk. A találati lista a legtöbb esetben kis bélyegképeket és fájlneveket tartalmaz, de például a WebSeek színeloszlást mutató hisztogramokat is megjelenít a képek és a videók mellett. A FaganFinder (www.faganfinder.com/img) oldalán egy helyen megtaláljuk valamennyi fontos képkereső rendszert.

Ami a képek indexelését illeti, a legtöbb rendszer viszonylag egyszerű megoldásokat használ. A legegyszerűbb módszer természetesen a fájlnevek leindexelése és kereshetővé tétele, de mivel egyre több program és berendezés automatikusan ad neveket a képeknek és a videóknak, ez már kevésbé járható út. Ennél jobb eredménnyel jár, ha a környező szöveges információt gyűjtik egy adatbázisba, feltéve hogy a multimédia-állomány szöveges weblapba van beágyazva, és hogy a szöveg tartalmi kapcsolatban van vele. Volt olyan kutatás is, amelyben azt vizsgálták, hogy mennyire hatékony az a módszer, ha a HTML oldalak fejlécébe és a multimédia-fájlokba ágyazott metaadatok alapján történik az indexelés. De az 1999-es jelentés

szerint akkoriban még nem nagyon terjedt el a metaadatok használata.

Osztályozási sémák

A web világán kívül vannak már osztályozási szisztémák képekre és képekre vonatkozó keresőkifejezésekre. Három ilyet mutatunk be a következőkben:

Enser és McGregor 1992-ben 2722 keresőkérdést elemeztek, melyeket egy kb. 10 millió tételes képadatbázis használói tettek fel, és ezeket négy csoportba sorolta: 1. *egyedi*; 2. *egyedi, minősítővel*; 3. *nem egyedi*; 4. *nem egyedi, minősítővel*. Az „egyedi” olyan kérdést jelent, amelynél a keresett objektum képe megkülönböztethető minden más hasonló típusú objektumtól. A „minősítő” pedig a keresés szűkítésére, finomítására szolgál. (A „Bill Clinton” kérdés pl. az első, a „Bill Clinton 1980” a második, a „középkorú férfi” a harmadik, az „asztalnál ülő középkorú férfi” pedig a negyedik csoportba tartozik.) A kutatás eredménye az lett, hogy a kérdések csaknem 70 százalékban egyedi személyre, tárgyra vagy eseményre vonatkoztak, és 34 százalékukban volt valamilyen – többnyire időbeli – minősítő.

Jörgensen 1998-ban elsőéves könyvtáros-hallgatókkal és középiskolásokkal folytatott le egy vizsgálatot, melyben a *Society of Illustrators* egyik albumából véletlenszerűen kiválasztott képeket kellett leírniuk a résztvevőknek. Minden kísérleti alany hat ilyen kivetített képet nézett meg, és azután a leírásaikban szereplő jellemzőket Jörgensen a következő 12 csoportba sorolta: 1. *absztrakt fogalmak*, 2. *színjellelmezők*, 3. *tartalmi elemek*, 4. *leíró jelzők*, 5. *(művészet)történeti információk*, 6. *helyre vonatkozó jellemzők*, 7. *tárgy*, 8. *ember*, 9. *emberrel kapcsolatos dolgok*, 10. *viszony kifejezése*, 11. *személyes benyomás*, 12. *vizuális elemek*.

Chen 2001-es kutatása során 29 művészettörténet-sz-hallgató 534 keresését elemezte, amelyeket egy nagy művészeti adatbázisban folytattak le. Az adatbázis képeken kívül folyóiratcikkeket és középkori kéziratokat is tartalmazott. Chen nem választotta szét a különböző dokumentumtípusokra vonatkozó kereséseket, de többek között a képi keresések osztályozására korábban publikált sémák szerint is csoportosította őket, majd az Enser és McGregor-, illetve a Jörgensen-féle osztályozás kombinálását javasolta a következő módon: 1. *hely*, 2. *konkrét tárgy (egyedi név)*, 3. *művészettör-*

téneti információ, 4. *ember*, 5. *emberrel kapcsolatos*, 6. *konkrét tárgy (nem egyedi név)*, 7. *szín*, 8. *vizuális elemek*, 9. *leíró jelzők*, 10. *absztrakt fogalmak*, 11. *tartalom/esemény*, 12. *külső viszony/kapcsolat*, 13. *nézői benyomás*.

Nem nagyon vizsgálták még meg, hogy ezek az osztályozási sémák mennyire lennének használhatók a weben található képek indexelése és visszakeresése során, érdemes-e őket figyelembe venni a webes képgyűjtemények tervezésekor? Ennek a kérdésnek a megválaszolását tűzte ki célul a következőkben bemutatott kutatás.

A kutatás ismertetése

A kutatás keretében az Excite kereső 2001. április 30-i naplóját elemeztük, amely csaknem egymillió keresést tartalmazott. (Akkoriban az Excite az ötödik legnépszerűbb keresőgép volt a weben.) A naplófájl minden bejegyzése három adatot tartalmazott: időpont, felhasználó azonosító kódja, keresőkifejezés. Utóbbiakból kiválogattuk a képekkel kapcsolatos szavakat tartalmazó kereséseket, amelyek száma kb. 4500 volt, és ezekből véletlenszerűen kiválasztottunk 587 keresést a részletes elemzéshez. Ezek mindegyikét három ember egymástól függetlenül besorolta az Enser-McGregor-féle négy kategória egyikébe, és ha a keresőkérdés minősítőt is tartalmazott, akkor azt a Chen által finomított osztályozás szerint értékelték. Ezután a Jörgensen-féle osztályozás szerint is elvégezték a kérdések besorolását. Egy kérdést több csoportba is besorolhattak, sőt szükség esetén az osztályozási sémákat további kategóriákkal is bővíthették; és egy 1-től 7-ig terjedő skálán kellett jelezniük, hogy mennyire biztosak a besorolás helyességében. A továbbiakban azokat a besorolásokat fogadtuk el, amelyeknél legalább két értékelő véleménye egyezett.

Az Enser és McGregor által használt kategóriáknál a besorolások 12%-át kellett érvénytelennek tekinteni, mert a három értékelő mindegyike más rovatba helyezte el az adott keresőkérdést. A többi kérdés 71.9%-a a *nem egyedi, minősítővel* kategóriába került (Enser és McGregor eredeti kutatásánál ez csak 25% volt). Ha az *egyedi* kategóriát is számításba vesszük, akkor összesen a kérdések 87.1%-a tartalmazott legalább egy minősítőt, ami meglepően magas szám, tekintve a webes keresőkérdések rövidségéről elterjedt legendákat. A kérdések elhelyezése a négy fő kategória valamelyikébe nem okozott nagy gondot, tekintve hogy

ezek nagyon általánosak, de a minősítők besorolásánál már megoszlottak a vélemények. Az értékelők 2524 minősítőt osztályoztak. A leggyakrabban előfordult típusok csökkenő sorrendben: *Gyűjtemény*, *Pornográfia*, *Prezentálás módja*, *Költség*.

A Jörgensen-féle osztályozási rendszerénél csak 2%-át kellett elvetni a kérdéseknek a miatt, mert mindhárman másként értékelték őket. Itt a leggyakoribb típusok a *Képgyűjtemény*, a *Leíró jellemzők* és az *Ember* voltak. Jörgensen eredeti 12 kategóriája nem teljesen alkalmas a webes képkeresések osztályozására, így további három típust kellett bevezetni: *Költség*, *URL cím* és *Képgyűjtemény*. A költség – amit tipikusan a „free”, vagyis „ingyenes” szó jelez a keresőkérdésekben – fontos szempont a web használói számára; ez érthetően nem merült fel a korábbi „laboratóriumi” környezetben folytatott kutatásoknál. A web hipertext jellege magyarázza, hogy URL címek is nagy számban fordulnak elő a keresőkifejezésekben, és hogy az emberek gyakran szűkítik a keresést gyűjteményekre, vagyis fontos számukra a kép lelőhelye. Az emberek és az emberekkel kapcsolatos dolgok dominálnak a kérdések között, míg az olyan tulajdonságok, mint például a szín, alig számítanak, pedig az ilyen képjellemezőkhöz a keresőgépek általában külön szűkítési lehetőséget nyújtanak.

Az eredmények értékelése

Enser és McGregor 1992-es kísérleténél a keresőkérdések mintegy 70 százaléka konkrét személyre vagy eseményre vonatkozott, és 34 százalékban volt valamilyen minősítő a kérdésben. A jelen kutatásban viszont, a webes keresőgép adatait elemezve az derült ki, hogy a keresések többsége nem egyedi személyre vagy tárgyra irányult, és hogy csaknem mindig (87.1%) volt valamilyen szűkítő jellemző is a kérdésben, legtöbbször képgyűjteményre (49.4%), pornográf tartalomra (16.2%), a kép prezentálásának módjára (12.3%) és a költségre/ingyenességre (10.8%). Jörgensen

1998-as tanulmányában a *Konkrét tárgy* volt a lista élén (29.3%), ezt követte az *Ember* (10.0%), majd a *Szín* (9.3%). A mi vizsgálatunk szerint azonban a webes kereséseknél a leggyakoribb a *Képgyűjtemény* (31.2%), a *Leíró jellemző* (19.0%), az *Ember* (18.8%), a *Tárgy* (13.5%) és a *Költség* (5.8%). Chen 2001-es kutatása során a *Hely* (23%) és a *Tárgy* (18%) típusú minősítőt találta a leggyakoribbaknak – ezek az adatok sem hasonlítanak a webhasználók képkereséseinek megoszlásához.

Látható tehát, hogy ezek a korábban publikált osztályozási sémák nem igazán használhatók a webes környezetben folyó képkeresésekhez. Az egyik lehetséges magyarázat, hogy a különböző felhasználói csoportok és a rájuk jellemző tematikájú képek eltérő keresési stratégiákat eredményeznek. Például egy jól osztályozott, nagy képarcívumnál valószínűleg nagy számban fordulnak elő igen speciális (*egyedi*) keresőkérdések. Egy csoport művészettörténész hallgató szintén egészen másféle típusú kereséseket végez egy szakadatbázisban, mint amilyeneket a hatalmas és rendkívül változatos tartalmú weben folytatnak az emberek. Kutatásunk egyik eredménye az, hogy célszerű öt új jellemzőt bevezetni, amelyeket gyakran használnak a webes képkereséseknél: *Gyűjtemény* (pl. „stock photography”), *Pornográfia* (pl. „gay”), *Prezentálás* (pl. „clipart”), *URL* (pl. „www.bhphoto.com”) és *Költség* (pl. „free”).

Ha alaposabban megismerjük, hogy valójában hogyan is keresnek az emberek digitális képeket, a jelenleginél jobb keresőrendszereket és -felületeket lehet tervezni. Ehhez a folyamathoz nyújtanak segítséget a webes képkeresések jellemzőihez igazított osztályozási sémák, melyeket például a keresőkifejezések kézi vagy automatikus metaadatoláshoz, címkézéséhez lehet felhasználni.

/JANSEN, Bernard J.: Searching for digital images on the web. = Journal of Documentation, 64. köt. 1. sz. 2008. p. 81–101./

(Drótos László)