

Óbudai Egyetem

Doktori (PhD) értekezés
tézisfüzete



**Tartalom alapú keresési algoritmusok képi
adatbázisokban**

Sergyán Szabolcs

Témavezetők:
Szeidl László DSc
Rövid András PhD

Alkalmazott Informatikai Doktori Iskola

Budapest, 2011. július 15.

1. A kutatás előzményei

Az internet használatának gyors növekedésével, valamint a digitalizálás és tároló eszközök árának csökkenésével egyre népszerűbbé vált szövegek, képek, grafikák, hanganyagok digitális formában való készítése és tárolása. Ez növelte annak igényét is, hogy az eltárolt tartalmak között hatékony keresést lehessen végrehajtani. Ennek az általános problémának egy része a képi anyagok tárolásának és közöttük való keresésnek a megvalósítása.

Képi adatbázisokban történő keresésre alapvetően két különböző módszer létezik: a szöveg alapú, illetve a tartalom alapú megközelítés. A szöveg alapú rendszerek fejlesztése már az 1970-es években megkezdődött. Ezekben a rendszerekben a képekhez manuálisan rendelnek hozzá szöveges leírókat, amelyek ezt követően az adatbázisban történő keresés alapjául szolgálnak. Ennek a módszernek két hátránya van. Az első, hogy jelentős mértékű emberi munkát kíván a szöveges indexelés megvalósítása. A második pedig, hogy a szöveges indexek pontossága az emberi érzékelés szubjektivitásától függ [5, 15]. A képek közötti keresések egyik fontos feladata, hogy a képek tartalma alapján szemantikai jellemzőket nyerjünk ki, melyek alapján a felhasználó igényeihez illeszkedő találatokat kapunk egy keresés során. A szöveges alapú keresés hátrányainak kiküszöbölése érdekében indult el az 1980-as évek elején a tartalom alapú kereső rendszerek (CBIR - Content-based Image Retrieval) rohamos, máig is tartó fejlődése [2, 16, 10, 9].

A tartalom alapú kereső rendszerekben a képek saját vizuális tartalmuk alapján vannak indexelve. Leggyakrabban figyelembe vett jellemzők a szín, a textúra, illetve az alak.

A tartalom alapú képkereséssel kapcsolatban első meghatározó cikket Chang és Liu publikálta 1984-ben [1], amelyben a szerzők bemutattak egy kép indexelési és absztrakciós eljárást képeslap adatbázisban való keresésre.

Az alapvető különbség a tartalom alapú és a szöveges visszakereső rendszerek között az, hogy az utóbbinak elhagyhatatlan része az emberi beavatkozás szükségessége. Az emberek viszont hajlamosak arra, hogy magas szintű jellemzőket, például fogalmakat használjanak kulcsszóként, szöveges leíróként. Ezzel szemben a számítógép által automatikusan előállított jellemzők a gépi látás és képfeldolgozás területén alkalmazott eljárások használatával állítanak elő főként alacsony szintű jellemzőket. Ilyen alacsony szintű jellemzőnek tekinthetjük a színt, textúrát, alakot, térbeli elhelyezkedést, stb. Általában viszont nincs közvetlen kapcsolat az alacsony- és magas szintű jellemzők között [15]. Bár számos kifinomult algoritmust fejlesztettek már ki a szín, alak és textúra jellemzők leírására, ezek az eljárások nem képesek pontosan modellezni a képeken található szemantikai információkat, és így számos korlátja van annak, hogy olyan tartalom alapú kereső rendszer jöjjön

létre, amely a felhasználók igényeit teljes mértékben ki tudja elégíteni széles spektrumú képi adatbázisokban történő keresés esetén [12]. CBIR-rendszerekben végzett alapos kísérletek azt igazolták, hogy az alacsony szintű leírók sok esetben nem alkalmasak az emberi agyban kialakuló magas szintű szemantikai fogalmak leírására [17]. Emiatt a tartalom alapú kereső rendszerek teljesítménye még mindig messze elmarad a felhasználók elvárásaitól.

A lekérdezéseknek három szintjét különböztetjük meg a CBIR rendszerekben [5].

- 1. szint:** A kép primitív jellemzői (szín, textúra, alak, térbeli elhelyezkedés, stb.) alapján történő keresés. Tipikus esete a példa kép alapján történő keresés: „keress ehhez hasonlót”.
- 2. szint:** A képen található objektumok jellemzőiből logikai következtetések alapján kinyert adott típusú azonosítók alapján történő keresés. Például „találd meg egy virág képét”.
- 3. szint:** Absztrakt attribútumok alapján történő keresés, beleértve magas szintű leíróit a képnek, amelyek a kép készítésének körülményeiről árulnak el információt. Például „találj egy képet, amin örömteli tömeg látható”.

A 2. és 3. szintet együttesen szemantikai kép visszakeresésnek nevezzük, az első és második szint közötti hézag pedig az ún. szemantikai hézag [3, 16].

A jellemzők (tartalom) kinyerése az alapja a CBIR rendszereknek [14]. A jellemzők két nagy csoportra oszthatók. Az elsőbe tartoznak az általános jellemzők, mint a szín, a textúra, illetve az alak, az utóbbiba pedig alkalmazás függő leírók, például az emberi arcok [31, 30], ujjlenyomatok.

Az elmúlt évtizedekben számos tartalom alapú képkereső rendszert fejlesztettek ki, melyek részletes felsorolását az értekezésemben tettem meg. Ezen rendszerekben a kép jellemzőinek kinyerése során jelentős szerepe van a kép szín- vagy textúra alapján homogén összefüggő régióinak, amelyek detektálása során szegmentáló algoritmusokat alkalmaznak. Ezen algoritmusok paraméter értékei általában a kép tulajdonságaitól függő értékek, ezért automatikus meghatározásuk fontos részfeladata a jellemzők előállításának.

A megvalósított rendszerek számos távolság- és hasonlósági mértéket használnak a jellemzővektorok összehasonlítására, ezek kidolgozása még nem tekinthető a kutatási terület lezárt részének.

A tartalom alapú kereső rendszerek egy fontos típusa a skicc alapú kereső rendszer, amelynél a felhasználó által felvázolt szabadkézi rajz alapján keresünk hasonló alakú objektumot tartalmazó képeket egy adatbázisból.

Értekezésemben a képhez hozzárendelt jellemző vektorok előállítását támogató előfeldolgozó eljárások javításával, hasonlósági mértékek vizsgálatával és skicc alapú kereső rendszerrel foglalkozom részletesen.

2. Célkitűzések

A képi adatbázisokkal szemben elvárás, hogy a benne tárolt képeket hatékony módon tudjuk visszakeresni. Ehhez az adatbázisoknak két lényeges funkcióval kell rendelkezniük. Egyrészt az egyes képekhez a képen lévő tartalmat jól jellemző indexeket kell előállítani és tárolni, másrészt ezen indexek összehasonlításával egy keresésnek a felhasználó igényeinek legmegfelelőbb eredményeket kell szolgáltatnia.

A tartalom alapú képkereső rendszereknél a képet leíró indexeket, vagy tulajdonság vektorokat a számítógép állítja el automatizáltan a képen tárolt információk (pixelek intenzitása, szomszédsági viszonyok, stb.) felhasználásával. Ezen rendszerek esetén a kulcsszavak előállítása sokkal gyorsabb, mint az ember által előállított szöveges leírók használata esetén, valamint új igényekhez illeszkedő indexelés is egyszerűbben megvalósítható.

Az indexek előállítása előtt a képet úgy kell átalakítani, hogy abból hatékonyan lehessen az információkat kinyerni. Ezt a fázist a képi adatbázisokban való keresés során előfeldolgozásnak tekintjük, mely számos képfeldolgozási eljárás használatát igényli.

A képen található tárgyak meghatározása érdekében általában szegmentáló, valamint élkereső eljárások használata indokolt. Ezek az eljárások viszont általában előre meghatározandó paraméter értékeket használnak és eredményük nagy mértékben függ az alkalmazott paramétereiktől. A hatékony szegmentálás, illetve éldetektálást megvalósító paraméterezés viszont általában függ a képtől, így általános paraméterek meghatározása nem célravezető. Célkitűzésem, hogy olyan algoritmusokat dolgozzak ki, melyek lehetővé teszik a szegmentáló és élkereső algoritmusok képi tartalomtól függő automatikus paraméterezését, így mellőzhető az emberi beavatkozás szükségessége képi adatbázisok indexelése során.

A képi adatbázisok indexelésének előkészítéséhez kapcsolódóan meg kívánom vizsgálni, hogy a gyakran használt előfeldolgozó algoritmusok milyen más jellegű, hatékonyabb működést biztosító megközelítésekkel helyettesíthetők.

Az indexelt adatbázisokban hasonló képek keresése hasonlósági és távolsági mértékek használatával történik. Az egyik leggyakrabban használt képleíró a kép színezettségéről információt szolgáltató színhisztogram. Célkitűzésem, hogy olyan új hasonlósági mértéket vezessenek be, mely javítja a hasonló képek keresésének eredményességét színhisztogramok összehasonlítása esetén.

A tartalom alapú keresés egy speciális területe a skicc alapú keresés. A skicc alapú kereső rendszerek egy szabadkézi vonalrajzot hasonlítanak össze képekkel oly módon, hogy a képeken található tárgyakat mérik össze a megadott skicc-cel. Egy ilyen rendszer megvalósításához szükséges, hogy a képeken található tárgyak meghatározó éleit jól tudjuk detektálni, valamint úgy tudjuk átalakítani, hogy azok összemérhetőek legyenek a szabadkézi vonalrajzokkal. Céлом, hogy olyan algoritmust dolgozzak ki, mely a képeken fellelhető tárgyakat úgy tudja transzformálni, hogy azok skiccekkal hatékonyan összehasonlíthatók legyenek.

A keresési eredmények sok esetben nem illeszkednek megfelelő mértékben a felhasználó elvárásaihoz, ezért a tartalom alapú kereső rendszerekbe releváns visszacsatolást lehetővé tevő felhasználói beavatkozást szoktak beépíteni. Célkitűzésem, hogy a megvalósított skicc alapú kereső rendszerbe olyan visszacsatolási lehetőséget építsek be, mely által a keresések eredményessége jelentős mértékben javítható.

3. Vizsgálati módszerek

3.1. Szegmentáló algoritmusok automatikus paraméterezése

Szegmentáló algoritmusok automatikus paraméterezésénél abból a hipotézisből indultam ki, hogy ha rendelkezem egy kép több régiófelbontásával is, akkor a legkedvezőbb régiófelbontás a leginkább hasonló a többi régiófelbontáshoz. Ennek megállapításához szükséges a régiófelbontások hasonlóságának értelmezése, amelyet az értekezésemben részletesen ismertetett algoritmussal valósítottam meg.

Ezen algoritmus első lépésében két régiófelbontásban található régiókat párosítom össze, majd második lépésként értelmezem az előállt régiópárok hasonlóságát. Az így megkapott hasonlóságok alapján végül értelmezem a két régiófelbontás hasonlóságát is.

3.2. Élkereső algoritmusok automatikus paraméterezése

Kifejlesztett algoritmusom abból a hipotézisből indul ki, hogy ha rendelkezésre áll egy kép több különböző paraméterértékkel előállított élmátrixa, akkor az élmátrixok közül az felel meg leginkább a felhasználó elvárásainak, amely a legtöbb más paraméterezéssel előállított élmátrixhoz hasonlít. Ennek meghatározásához elő kell állítanunk egy kép több paraméter értékhez tartozó élmátrixát, valamint értelmeznünk kell ezen élmátrixok között is egy hasonlósági mértéket. Ezen hasonlóság értelmezésének részletes ismertetése az értekezésemben található meg.

3.3. HOSVD alapú eljárások használata az előfeldolgozáshoz

Legyen

$$f(\mathbf{x}), \mathbf{x} = (x_1, \dots, x_N)^T, x_n \in [a_n, b_n], 1 \leq n \leq N, \quad (1)$$

egy n -változós sima függvény. Az $f(\mathbf{x})$ függvény az alábbi módon approximálható egyváltozós ortonormált rendszert alkotó sima függvények segítségével:

$$f(\mathbf{x}) = \sum_{k_1=1}^{I_1} \dots \sum_{k_N=1}^{I_N} \alpha_{k_1, \dots, k_N} p_{1, k_1}(x_1) \cdot \dots \cdot p_{N, k_N}(x_N). \quad (2)$$

Az approximáció pontossága nagy mértékben függ a p_{n, k_n} egyváltozós függvények számától. A p_{n, k_n} függvények lehetnek

- ortonormált polinomok,
- trigonometrikus függvények,
- speciálisan meghatározott ortonormált rendszert alkotó függvények.

Színes képek tenzor reprezentációja a következő módon valósítható meg. Legyen $f(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, x_3)^T$ a digitális képet leíró függvény, ahol x_1 és x_2 a képpont koordinátáját, x_3 pedig a színt komponens sorszámát fejezi ki, azaz RGB színtérben ez a piros esetben 1, zöld esetben 2, kék esetben pedig 3 értéket vesz fel. Az $f(\mathbf{x})$ függvény kifejezhető az alábbi alakban:

$$f(\mathbf{x}) = \sum_{k_1=1}^{I_1} \sum_{k_2=1}^{I_2} \sum_{k_3=1}^{I_3} \alpha_{k_1, k_2, k_3} \tilde{w}_{1, k_1}(x_1) \cdot \tilde{w}_{2, k_2}(x_2) \cdot \tilde{w}_{3, k_3}(x_3). \quad (3)$$

A képpontok piros, zöld és kék komponensei egy $m \times n \times 3$ méretű tenzorban \mathcal{B} tárolhatók. \mathcal{B} felírható az alábbi módon:

$$\mathcal{B} = \mathcal{D} \boxtimes_{n=1}^3 \mathbf{U}^{(n)}, \quad (4)$$

ahol \mathcal{D} az ún. magtenzort jelöli, az $\mathbf{U}^{(n)}$, $1 \leq n \leq 3$ mátrix oszlopvektorai pedig az n -edik, $1 \leq n \leq 3$ dimenzióhoz tartozó egyváltozós ortonormált rendszert alkotó $\tilde{w}_{n,k_n}(x_n)$ függvények diszkretizált változatainak felelnek meg.

3.4. Távolsági- és hasonlósági mértékek

A színhisztogramok - mint vektorok - összehasonlítására alkalmas mérték a kvadratikus távolság, amely lehetővé teszi, hogy a nem összetartozó hisztogram vödrök távolságát is figyelembe vegyük valamilyen súllyal. Ez azért fontos az összehasonlítás során, mert a különböző színek hisztogramvödrökbe sorolásakor előállhat, hogy egymáshoz közeli színek más-más vödörbe, míg távoli színek azonos vödörbe kerülnek. Vizsgálataim során olyan súlyozást kerestem kvadratikus távolság használatakor, amellyel a színhisztogramok összehasonlítása jobb eredményt szolgáltat más gyakran használt módszerekhez képest.

3.5. Skicc alapú kereső rendszer

Megvalósított skicc alapú kereső rendszeremben a képek előfeldolgozását követően a képekből az élhisztogram leíró (EHD - edge histogram descriptor) [6] és az irányított gradiensek hisztogramja (HOG - histogram of oriented gradients) [4] alkalmazásával nyertem ki jellemzőket. A releváns visszacsatolást megvalósító eljárás kifejlesztésekor pedig a skála invariáns jellemző transzformációt (SIFT - scale invariant feature transform) [11] alkalmaztam. Ezen eljárások részletes ismertetése megtalálható az értekezésemben.

3.6. Tesztelési eljárások

A kifejlesztett algoritmusok tesztelése gyakran szubjektív döntéseken alapul, hiszen a képi adatbázisokban történő kereséstől azt várjuk el, hogy olyan eredményt szolgáltatson, mely a felhasználó szubjektív igényeinek megfelel. A szubjektivitás kiküszöbölése érdekében többféle megoldás lehetséges.

A tesztelés során lehetséges mérőszámok definiálása, melyekre leggyakrabban a precíziót és a felidézést használják [7]. Tételizzük fel, hogy van N darab képet tartalmazó teszt

adatbázisunk, melyből Q darab számít relevánsnak találatnak egy keresés során, Z jelöli az elvárt releváns találatok számát, P pedig az eredmény lista hosszát. Ezek ismeretében meghatározható a rendszert jellemző két mérőszám:

$$precizio = \frac{Q}{P}, \quad (5)$$

$$felidezes = \frac{Q}{Z}. \quad (6)$$

A tesztelési eredmények összehasonlíthatósága érdekében több tesztadatbázist is készítettek, melyek lehetővé teszik, hogy különböző algoritmusokat ugyanolyan körülmények között lehessen tesztelni. A leggyakrabban használt két adatbázis az Amsterdam Library of Object Images [8] és a Columbia Object Image Library [13]. Mindkét adatbázis homogén háttérbe helyezett tárgyakról készített fotókat tartalmaz azonos megvilágítás mellett, ahol a tárgyakról 5° -os függőleges tengely körüli elforgatással több fénykép is készült.

Az eredmények „jósága” olyan módon is mérhető, hogy ha vizsgáljuk, hogy egy adott képhez mely N darab kép van a legközelebb, akkor ezt követően megnézzük, hogy az eredményül kapott képekhez legközelebbi N kép között megtalálható-e az eredeti kép. Így módon azokat a képeket tekintjük releváns találatoknak, melyekhez legközelebbi N kép között szerepel az eredeti kép is.

4. Új tudományos eredmények

1. téziscsoport

A képfeldolgozó algoritmusok sok esetben előre megadott paraméterek alapján szolgáltatnak eredményt. Ezen paraméterek értéke viszont gyakran függ a képen tárolt információktól. Képi adatbázisok indexelésénél szükséges, hogy ezeket a paramétereket a képen tárolt információk függvényében automatizált módon tudjuk előállítani.

1.1. tézis: Eljárást dolgoztam ki, mellyel szegmentáló algoritmusok automatikus paraméterezése valósítható meg. Az eljárás alkalmazhatóságát tesztek igazolják. [25, 24, 23, 22]

1.2. tézis: Eljárást dolgoztam ki élkereső algoritmusok automatikus paraméterezésének hatékony megvalósítására. Az eljárást más hasonló módszerrel összehasonlítottam és ezen tesztek igazolják, hogy pontosabb paraméterezést sikerült ilyen módon megvalósítani. [25, 18]

2. téziscsoport

Az előfeldolgozás számos esetben megkívánja valamilyen simító, illetve zajcsökkentő eljárás használatát.

Simítást többféle módon is meg lehet valósítani. Leggyakoribb erre a kép szűrése például átlagoló, vagy Gauss maszkkal. Gyakran használt eljárás a kép Fourier transzformálása oly módon, hogy az előálló trigonometrikus tagokból csak az első pár tagot tartjuk meg. Hasonló módszer lehet a magasabb rendű szinguláris érték dekompozíció (HOSVD - High Order Singular Value Decomposition) használata. Ebben az esetben a képet három dimenziós tenzornak tekintjük a vizsgálati módszereknél ismertetett módon. Ezt a három dimenziós tenzort ortonormált függvények kompozíciójaként állítjuk elő. Ha ebből az előállításból is csak pár tagot tartunk meg, akkor a Fourier transzformációhoz hasonlóan szűrést tudunk megvalósítani, amelynek eredménye részletgazdagabb, így a képi indexek legyártására alkalmasabb eredményt szolgáltat.

2.1. tézis: Igazoltam, hogy HOSVD-alapú függvény approximációval az azonos számú ortonormált függvény megtartása esetén a transzformált képen végrehajtott simítás nagyobb részletgazdagságot eredményez mint ugyanolyan számú trigonometrikus függvény megtartása esetén a Fourier-alapú approximáció. [19]

Ennek az eljárásnak másik előnye, hogy nem jelennek meg a Fourier-alapú közelítésnél előálló ciklikus hullámok, melyek a Fourier sor alaptulajdonságai miatt jelennek meg a képeken.

3. téziscsoport

Képi adatbázisokban történő keresésnél az egyes képekhez hozzárendelt indexeket hasonlítjuk össze valamilyen hasonlósági mérték alkalmazásával. Szín alapú összehasonlítás esetén a figyelembe vett index gyakran a kép színhisztogramja. A színhisztogramok előállítása során a vödörkre osztás miatt információvesztés következik be. Az információvesztést a gyakran használt távolsági- és hasonlósági mértékek közül a kvadratikus távolság használatával lehet kompenzálni.

3.1. tézis: Hisztogramok kvadratikus távolságánál olyan új súlyokat vezettem be, melyek a keresés által szolgáltatott releváns találatok számát növelték a nem kvadratikus távolságot használó, illetve más súlyozású kvadratikus távolságot alkalmazó eljárásokhoz képest. [21, 20]

4. téziscsoport

A tartalom alapú képkeresés egy speciális területe a szabadkézi vonalrajzok (skiccek) alapján történő keresés. Ezen rendszereknél a felhasználó által készített szabadkézi rajzhoz hasonló alakú objektumokat keresünk képeken. A kereséshez szükséges, hogy a képeken található tárgyakat úgy alakítsuk át, hogy azok összemérhetőek legyenek a skiccekkel.

4.1. tézis: Kifejlesztettem egy olyan előfeldolgozó eljárást, mely skicc-alapú kereső rendszer esetén valós körülmények között készített képeket úgy alakít át, hogy azok összemérhetőek szabadkézi vonalrajzokkal. A kifejlesztett módszert más hasonló rendszerekkel azonos tesztkörülmények között összemértem. A mérési eredmények igazolták, hogy az előfeldolgozó eljárásom eredményesebben alkalmazható, mint más eljárások. [26]

A tartalom alapú képkereső rendszerekben a megtalált képeknek a felhasználó igényeihez minél jobban igazodniuk kell. Ennek érdekében visszacsatolási lehetősége is van sok esetben a felhasználónak, ahol megadhatja, hogy a megtalált képek közül melyeket tekint relevánsnak. A rendszer ezen visszacsatolás alapján a felhasználó igényeihez jobban illeszkedő találatokat állíthat elő.

4.2. tézis: Skicc alapú kereső rendszerbe beépítettem egy olyan felhasználói releváns visszacsatolási lehetőséget, mely az ún. SIFT-leíró használatával jelentős mértékben növeli a keresések eredményességét. [26]

5. Az eredmények hasznosítási lehetősége

A szegmentáló és élkereső algoritmusok automatikus paraméterezését megvalósító algoritmus nem csak képi adatbázisok indexelése esetén használható, hanem más esetekben is, amikor a felhasználónak nincs lehetősége megállapítani az adott képhez legjobban illeszkedő paraméter értéket. Az algoritmus hátránya viszont, hogy jelentős futási időt igényel, ami a képi adatbázisok indexelésénél nem probléma, más esetekben viszont problémát jelenthet.

A HOSVD módszer a képfeldolgozás területén széles körben alkalmazható lehet, így pl. simításra, képek nagyítása esetén interpolációs technikaként, valamint élek detektálására is. A képi adatbázisok esetében megvizsgálandó, hogy a HOSVD eljárást követően előálló képre jellemző ortonormált függvények a kép indexelésére alkalmazhatók e.

Hasonlóan nyitott kérdés még, hogy a HOSVD alapú eljárás lényegkiemelésre, az információ tömörítésére, valamint a képek osztályozására hatékonyan alkalmazható módszer-e.

A kidolgozott súlyozott kvadratikus távolság használatának továbbfejlesztési lehetősége lehet, ha az alkalmazott színtérben egyenletes eloszlástól eltérő színeloszlást feltételezve az egyes irányokban más-más súlyokat alkalmazunk.

A skicc alapú rendszerben megvalósított releváns visszacsatolás SIFT leíró használatával nem csak skicc alapú, hanem általános tartalom alapú kereső rendszerekben is alkalmazható lehet, de ezzel kapcsolatban kísérleteket még nem végeztem.

6. Irodalmi hivatkozások listája

- [1] S. Chang and S. Liu. Picture indexing and abstraction techniques for pictorial databases. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(4):475–483, 1984.
- [2] Y. Chen, J. Li, and J. Z. Wang. *Machine Learning and Statistical Modeling Approaches to Image Retrieval*. The Kluwer International Series on Information Retrieval. Kluwer Academic Publisher, Boston / Dordrecht / London, 2004.
- [3] Y. Chen, J. Wang, and R. Krovetz. An unsupervised learning approach to content-based image retrieval. In *IEEE International Symposium on Signal Processing and its Applications*, pages 197–200, July 2003.
- [4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 886–893, July 2005.
- [5] J. Eakins and M. Graham. Content-based image retrieval. Technical report, University of Northumbria, Newcastle, 1999.
- [6] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa. An evaluation of descriptors for large-scale image retrieval from sketched feature lines. *Computers and Graphics*, 34:482–498, October 2010.
- [7] R. Fabbri, L. D. F. Costa, J. C. Torelli, and O. M. Bruno. 2D euclidean distance transform algorithms: A comparative survey. *ACM Computing Surveys*, 40(1):1–44, February 2008.
- [8] J.-M. Geusebroek, G. J. Burghouts, and A. W. Smeulders. The amsterdam library of object images. *International Journal of Computer Vision*, 61(1):103–112, 2005.
- [9] C. Jörgensen. *Image Retrieval: Theory and Research*. Scacecrow Press, 2003.
- [10] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1):262–282, Januar 2007.
- [11] D. Lowe. Object recognition from local scale-invariant features. In *IEEE International Conference on Computer Vision*, volume 2, page 1150, 1999.

- [12] A. Mojsilovic and B. Rogowitz. Capturing image semantics with low-level descriptors. In *International Conference on Image Processing*, pages 18–21, Thessaloniki, Greece, October 2001. ISBN 0-7803-6725-1.
- [13] S. Nene, S. Nayar, and H. Murase. Columbia object image library (COIL-100). Technical report, Columbia University, February 1996. CUUCS-006-96.
- [14] Y. Rui, T. S. Huang, and S.-F. Chang. Image retrieval: Current techniques, promising directions, and open issues. *Journal of Visual Communication and Image Representation*, 10(1):39–62, March 1999.
- [15] I. Sethi and I. Coman. Mining association rules between low-level image features and high-level concepts. In *SPIE Data Mining and Knowledge Discovery*, volume 3, pages 279–290, 2001.
- [16] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, December 2000. ISSN 0162-8828.
- [17] X. Zhou and T. Huang. CBIR: From low-level features to high-level semantics. In *SPIE, Image and Video Communications and Processing*, volume 3974, pages 426–431, San Jose, CA, January 2000.

7. A tézispontokhoz kapcsolódó tudományos közlemények

- [18] L. Csink and S. Sergyán. Automatic parametrization of edge detection algorithms. In *5th International Symposium on Intelligent Systems and Informatics*, pages 119–121, Subotica, Serbia, August 2007. ISBN 987-86-7031-131-3.
- [19] A. Rövid, I. J. Rudas, S. Sergyán, and L. Szeidl. HOSVD based image processing techniques. In *10th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, pages 297–302, Cambridge, UK, February 2011. ISBN 978-960-474-273-8.
- [20] S. Sergyán. Special distances of image color histograms. In *5th Joint Conference on Mathematics and Computer Science*, page 92, Debrecen, Hungary, June 2004.
- [21] S. Sergyán. Color histogram features-based image classification in content-based image retrieval systems. In *6th International Symposium on Applied Machine Intelligence and Informatics*, pages 221–224, Herl’any, Slovakia, January 2008. ISBN 978-1-4244-2106-0, IEEE Catalog Number: CFP0808E-CDR.
- [22] S. Sergyán. Content-based image retrieval using automatically determined color regions of images. In *7th International Symposium on Applied Machine Intelligence and Informatics*, pages 41–45, Herl’any, Slovakia, January 2009. ISBN 978-1-4244-3802-9, IEEE Catalog Number: CFP0908E-CDR.
- [23] S. Sergyán and L. Csink. Kísérletek a szín-alapú tartomány felismerés terén. In *Informatika a Felsőoktatásban 2005 Konferencia*, Debrecen, Hungary, August 2005. ISBN 963-472-909-6.
- [24] S. Sergyán and L. Csink. Automatic parametrization of region finding algorithms in gray images. In *4th International Symposium on Applied Computational Intelligence and Informatics*, pages 199–202, Timisoara, Romania, May 2007. ISBN 1-4244-1234-X, IEEE Catalog Number: 07EX1788.
- [25] S. Sergyán and L. Csink. Automatic parametrization of image processing algorithms. *SCIENTIFIC BULLETIN of "Politechnica" University of Timisoara*, 54(1):53–58, 2009. ISSN 1224-600X.

- [26] B. Szántó, P. Pozsegovics, Z. Vámosy, and S. Sergyán. Sketch4Match - Content-based image retrieval system using sketches. In *9th IEEE International Symposium on Applied Machine Intelligence and Informatics*, pages 183–188, Smolenice, Slovakia, January 2011. ISBN 978-1-4244-7428-8, IEEE Catalog Number: CFP1108E-CDR.

8. További tudományos közlemények

- [27] A. Kiss, T. Németh, S. Sergyán, Z. Vámosy, and L. Csink. Recognition of a moving object in a stereo environment using a content based image database. In *3rd Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence*, pages 65–74, Herl’any, Slovakia, January 2005.
- [28] S. Sergyán. Content based image retrieval in database of segmented images. In *4th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence*, pages 380–388, Herl’any, Slovakia, January 2006. ISBN 963-7154-39-6.
- [29] S. Sergyán. Color content-based image classification. In *5th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence and Informatics*, pages 427–434, Poprad, Slovakia, January 2007. ISBN 978-963-7154-56-0.
- [30] S. Sergyán. Classification of image databases using face detection. In *6th International Symposium on Intelligent Systems and Informatics*, Subotica, Serbia, September 2008. ISBN 978-1-4244-2407-8, IEEE Catalog Number: CFP-0884C-CDR.
- [31] S. Sergyán. A new approach of face detection-based classification of image databases. *Acta Polytechnica Hungarica*, 6(1):175–184, 2009. ISSN 1785-8860.
- [32] S. Sergyán and L. Csink. Consistency check of image databases. In *2nd Romanian-Hungarian Joint Symposium on Applied Computational Intelligence*, pages 201–206, Timisoara, Romania, May 2005. ISBN 963-7154-39-6.