2D és 3D bináris objektumok lineáris deformáció-becslésének numerikus megoldási lehetőségei

Tanács Attila[1], Joakim Lindblad[2], Nataša Sladoje[3], Kató Zoltán[1]

[1] Képfeldolgozás és Számítógépes Grafika Tanszék
Szegedi Tudományegyetem
{tanacs,kato}@inf.u-szeged.hu
[2] Centre for Image Analysis
Swedish University of Agricultural Sciences, Uppsala, Sweden
joakim@cb.uu.se
[3] Faculty of Technical Sciences
University of Novi Sad, Serbia
sladoje@uns.ac.rs

Absztrakt. A képregisztráció a képfeldolgozás egy alapfeladata, amelyet különböző időpontokban és/vagy nézőpontokból készített felvételek közötti geometriai viszony meghatározására használnak. Jelen cikkünkben egy általános keretrendszert adunk meg, amely tetszőleges n-dimenziós bináris képek/objektumok közötti lineáris regisztrációs problémát képes megoldani megfeleltetések létesítése nélkül. A megközelítésünk polinomiális egyenletrendszerek generálásán és megoldásán alapul. Szintetikus teszteken keresztül a polinom egyenletrendszer különféle numerikus megoldási lehetőségeit vizsgáljuk, analitikus, iteratív legkisebb négyzetes és kombinált módszereket veszünk figyelembe. A tesztjeink alapján az iteratív és a kombinált megközelítés biztosítja a legpontosabb eredményt. A tesztek megmutatják azt is, hogy képpont lefedettségi információ felhasználásával pontosabb objektum körvonal leírás kapható, ami a regisztráció pontosságát növeli. Mivel az egyenletrendszer megalkotásához csak egyszer kell a képeket bejárnunk, ezért módszereink alkalmasak nagy méretű regisztrációs problémák gyors megoldására, függetlenül a deformáció mértékétől.

1. Bevezetés

A képregisztráció a képfeldolgozás egy fontos területe, célja képek közötti geometriai viszonyok meghatározása. Az utóbbi évtizedekben számos megközelítés született problémák széles skálájára [1].

2D-ben a szegmentált alakzatok illesztése egy fontos regisztrációs probléma. Ilyen feladatoknál jellemzően két lépésből áll az illesztés: Először egy szegmentáló algoritmus előállítja az alakzatokat, majd ezután történik meg ezek illesztése. Ez a megközelítés különösen akkor hasznos, amikor a kép intenzitásértékei erős nemlineáris torzításnak vannak kitéve, amely modellezése nehéz, pl. a röntgen

képalkotás bizonyos eseteiben. Amennyiben könnyen detektálható objektumok találhatók a képen (pl. csontok vagy implantátumok röntgen vagy CT képen), akár egy egyszerű szegmentáló módszer is alkalmazható.

Parametrikus transzformációbecslő módszert javasoltak Haege és munkatársai szürkeárnyalatos képekre [2, 3], ahol az intenzitásértékek felhasználásával lineáris egyenletrendszert kaphatunk. A helyes működéshez elvárás, hogy ne legyen jelen a képen nemlineáris intenzitás-torzítás, valamint a háttértől eltérő intenzitásértékű objektumok egymásnak minél pontosabban megfeleltethetők legyenek. Ha képek között az átfedés nem teljes, akkor meg kell határozni egy átfedő területet, ami nem triviális feladat. Ha feltételezzük, hogy egymásnak megfeleltethető alakzatok/objektumok egyszerűen szegmentálhatók a képeken, és nem feltételezzük az intenzitástartomány állandóságát, bináris objektumok illesztésére vezethetjük vissza a problémát. Ilyen esetekre polinomiális egyenletrendszerre vezető megoldást Domokos és munkatársa javasolt [4]. Korábbi publikációkban ezen módszer különböző variánsai összehasonlításra kerültek más szakirodalmi módszerekkel [5, 6].

A 3D képalkotás orvosi és ipari alkalmazásokban manapság általánosan elterjedt. Bináris objektumok jöhetnek létre például a képalkotás eredményeként (pl. diszkrét tomográfia), generálódhatnak geometriai leírásokból (pl. CAD modell), vagy egymásnak megfeleltethető területek szegmentálásával. A 3D regisztrációs problémát a klasszikus módszerek általában vagy geometriai jellemzők kinyerésével, vagy a képpontintenzitások közvetlen felhasználásával oldják meg. Az adatok közötti megfeleltetéseket rendszerint iteratív keresés segítségével próbálják meghatározni. A leggyakrabban alkalmazott geometriai jellemzők közé a pontok, felszínek [7], vázak [8], vagy ponthalmazok [9] tartoznak. Képpontok hasonlóságán alapuló módszerek jellemzően nem-bináris képek esetén kerülnek alkalmazásra, de bináris esetben is felhasználhatók [1]. Mivel az iteratív keresés ezen módszerek esetében minden lépésben felhasználja a képjellemző/intenzitás adatokat, emiatt nagy méretű képi adatok esetében a megoldás időigénye jelentősen megnő.

Fontos döntés, hogy milyen geometriai transzformáció tételezhetünk fel az objektumok között? Nemlineáris transzformációk alkalmazásával lehetőség van pl. szövetek lokális alakváltozásainak detektálására. Ilyen megközelítés esetén viszont figyelembe kell venni a különböző szövetek eltérő deformációs jellemzőit, ami nagyon nehézzé és alkalmazás-specifikussá teheti a megoldást. Sok esetben elegendő globális lineáris transzformációt alkalmazni, különösen ha a gyors számítási idő elvárás.

Korábbi publikációinkra építve jelen cikkünkben egy egységes keretrendszert vezetünk be, amely tetszőleges n-dimenzióban használható. Ez a megközelítés polinom egyenletrendszerek generálásán és megoldásán alapul és a szegmentáláson túl nem igényli megfeleltetések keresését. Az egyenletrendszer a képek egyszeri bejárásával előállítható, annak megoldása már független a képek méretétől. Ez különösen alkalmassá teszi a módszerünket nagy méretű, globális lineáris regisztrációs problémák hatékony megoldására. A digitális képtér csak korlátozott pontosságot képes biztosítani, így az (5)-(9) egyenletekben szereplő integrálok

csak közelíthetők diszkrét összegként. A digitális adatábrázolás lehetőségeinek jobb kihasználása érdekében korábbi publikációinkban az objektumok képpont lefedettségi reprezentációjának felhasználását javasoltuk [13, 11] ahol a képpontok számértéke arányos a képpont térrészének a valós objektum által történő lefedettségével. Bizonyításra került, hogy egyező térbeli felbontás esetén a bináris reprezentációhoz képest egy ilyen fuzzy reprezentáció nagyobb pontosságot biztosít egyes képjellemzők becslésében [12].

Jelen cikkünk új eredménye a regisztrációs eljárás során előálló polinomiális egyenletrendszerek gyors és hatékony megoldási lehetőségeinek részletes tárgyalása. Megvizsgáljuk a túlhatározott egyenletrendszer esetén alkalmazható legkisebb négyzetes megoldást, a nem-szinguláris esetben használható közvetlen analitikus megoldást, valamint ezek együttes alkalmazásával előálló kombinált módszer tulajdonságait.

2. Regisztrációs keretrendszer

Röviden áttekintjük a cikkünkben felhasznált, korábban publikált 2D affin regisztrációs módszert [10] és kiterjesztjük tetszőleges n dimenzióra. Jelöljük a *sablon* és a *megfigyelés* képek objektumpont-koordinátáit $\mathbf{x}, \mathbf{y} \in \mathbb{P}^n$-el a projektív térben. A projektív tér használata lehetővé teszi az affin transzformációk egyszerű jelölésmódú megadását az ún. homogén koordináták alkalmazásával. Mivel az affin transzformáció nem változtatja meg a pontok utolsó (homogén) koordináta-értékét, vagyis az mindig 1 értékű lesz, így a további jelölésekben esetenként az Euklideszi teret is felhasználjuk, mindig az egyszerűbb jelölésmódot alkalmazva.

Jelölje $\mathbf{A}$ az ismeretlen, meghatározásra váró affin transzformációt. Az egyenlőség relációt az alábbi módon definiáljuk:

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y}.$$

A fenti egyenletek akkor is fennállnak, ha megfelelően választott $\omega : \mathbb{P}^n \to \mathbb{P}^n$ függvényeket alkalmazunk mindkét oldalra [4]:

$$\omega(\mathbf{A}\mathbf{x}) = \omega(\mathbf{y}) \quad \Leftrightarrow \quad \omega(\mathbf{x}) = \omega(\mathbf{A}^{-1}\mathbf{y}). \tag{1}$$

Bináris képek nem tartalmaznak radiometrikus információt, ezért reprezentálhatók a $\mathbb{1} : \mathbb{P}^n \to \{0, 1\}$ karakterisztikus függvényükkel, ahol 0 és 1 értékek kerülnek hozzárendelésre a háttér és az objektum pontjaihoz. Jelölje $\mathbb{1}_t$ a sablon, $\mathbb{1}_o$ a megfigyelés karakterisztikus függvényét. A pontmegfeleltetések elkerülése végett integráljunk az $\mathcal{F}_t = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{1}_t(\mathbf{x}) = 1\}$ és $\mathcal{F}_o = \{\mathbf{y} \in \mathbb{P}^n | \mathbb{1}_o(\mathbf{y}) = 1\}$ objektumpont-tartományok felett:

$$|\mathbf{A}| \int_{\mathcal{F}_t} \omega(\mathbf{x})\, d\mathbf{x} = \int_{\mathcal{F}_o} \omega(\mathbf{A}^{-1}\mathbf{y})\, d\mathbf{y}\,, \quad \text{és} \tag{2}$$

$$\int_{\mathcal{F}_t} \omega(\mathbf{A}\mathbf{x})\, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathcal{F}_o} \omega(\mathbf{y})\, d\mathbf{y}\,. \tag{3}$$

A transzformáció Jacobi determinánsa ($|\mathbf{A}|$) meghatározható:

$$|\mathbf{A}| = \frac{\int_{\mathcal{F}_o} d\mathbf{y}}{\int_{\mathcal{F}_t} d\mathbf{x}} \ . \tag{4}$$

A javasolt módszer alapötlete az, hogy elegendő számú lineárisan független egyenletet generáljunk az (1)-(3) képletek alkalmazásával. Legalább annyi egyenletre van szükségünk, amennyi az n-dimenziós affin transzformáció szabadsági foka. Ez 2D-ben 6, 3D-ben pedig 12. Mivel az ω-függvények az ismeretlenekre is hatnak, nem kaphatunk lineáris egyenletrendszert. Ha közvetlen analitikus megoldást szeretnénk, az egyik legjobb választás a polinomok használata.

Ennek elérésére a $\omega(\mathbf{x})^{(k)}_{p_1,\ldots,p_n} = x_1^{p_1} \ldots x_n^{p_n}$ formulát használjuk, amely a transzformált pont k-adik koordinátáját adja, ahol $p_1, \ldots, p_n \in \mathbb{N}_0$. Például a Eq. (2) egyenletből ezek a függvények az alábbi polinomiális egyenleteket generálják (harmadfokig bezárólag):

$$|\mathbf{A}| \int_{\mathcal{F}_t} 1 \ d\mathbf{x} = \int_{\mathcal{F}_o} 1 \ d\mathbf{y} \ , \tag{5}$$

$$|\mathbf{A}| \int_{\mathcal{F}_t} x_a \ d\mathbf{x} = \sum_{i=1}^{n+1} q_{ai} \int_{\mathcal{F}_o} y_i \ d\mathbf{y} \ , \tag{6}$$

$$|\mathbf{A}| \int_{\mathcal{F}_t} x_a x_b \ d\mathbf{x} = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} q_{ai} q_{bj} \int_{\mathcal{F}_o} y_i y_j \ d\mathbf{y} \ , \tag{7}$$

$$|\mathbf{A}| \int_{\mathcal{F}_t} x_a x_b x_c \ d\mathbf{x} = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} q_{ai} q_{bj} q_{ck} \int_{\mathcal{F}_o} y_i y_j y_k \ d\mathbf{y} \ , \tag{8}$$

ahol $1 \leq a, b, c \leq n$, $a \leq b \leq c$, és q_{ij} jelöli az $\mathbf{A}^{-1}$ inverz transzformáció elemeit. Magasabb fokszámú polinomok formalizmusa analóg módon megkapható.

Általában ezek a függvények vegyes momentumokat vezetnek be az egyenletek mindkét oldalán, így egy koordinátánként nem szétválasztható egyenletrendszert kapunk. A gyakorlatban ez nem okoz problémát az iteratív megoldóknál, de analitikus módszerrel ezek nem oldhatók meg hatékonyan (a megoldási módszereket a következő fejezetben tárgyaljuk). Szétválasztható egyenletrendszert kaphatunk az ω függvények további korlátozásával: a p_i paraméterek közül csak az egyik lehet nem nulla. Ebben az esetben a k-adik szeparált egyenletrendszer kompaktabb formában is felírható:

$$\begin{aligned} |\mathbf{A}| \int_{\mathcal{F}_t} x_k^p \ d\mathbf{x} = \sum_{i_1=0}^{p} \binom{p}{i_1} \sum_{i_2=0}^{i_1} \binom{i_1}{i_2} \ldots \sum_{i_n=0}^{i_{n-1}} \binom{i_{n-1}}{i_n} \\ q_{k1}^{p-i_1} \ldots q_{kn}^{i_{n-1}-i_n} q_{k,n+1}^{i_n} \int_{\mathcal{F}_o} y_1^{p-i_1} y_2^{i_1-i_2} \ldots y_n^{i_{n-1}-i_n} \ d\mathbf{y} \ . \end{aligned} \tag{9}$$

3. Numerikus megoldás

Síkbeli és térbeli esetre is kétféle megoldási stratégiát vizsgálunk: az egyenletrendszer közvetlen, analitikus megoldását, valamint az iteratív, legkisebb négyzetes értelemben optimális keresőt. Mindkét megközelítés rendelkezik előnyös és hátrányos tulajdonságokkal.

- Az analitikus megoldás esetén az n-dimenziós affin transzformáció szabad paramétereinek megfelelő számú egyenletet tudunk használni, míg az iteratív esetben dolgozhatunk túlhatározott egyenletrendszerrel is, amely nagyobb stabilitást adhat.[1]
- Gyors és hatékony analitikus megoldhatósághoz szükséges, hogy az egyenletrendszer koordinátánként szétválasztható legyen. Az iteratív technikánál nincs ilyen elvárás.
- Az analitikus módszerek több száz, vagy akár több ezer lehetséges megoldást is adhatnak, ezek közül sok, de akár az összes is komplex lehet. Ezek között található a globális megoldás is, függetlenül a transzformáció magnitúdójától. Szükség van egy megoldás kiválasztó sémára, amely ezek alapján egy valós megoldást biztosít. Az iteratív megközelítés egy valós eredményt ad, viszont a keresés elakadhat lokális minimumokban. Ezen lokális minimumok elkerüléséhez összetetett keresési stratégiára van szükség. A megoldás számítási ideje jobban megbecsülhető az analitikus megoldás esetén, az iteratív technikánál nagyobb különbségek is kialakulhatnak.
- Az analitikus megoldást hatékonyan affin esetre tudjuk megadni, az iteratívnál megszorításokat vezethetünk be a transzformáció fokszámára vonatkozóan. Valós képregisztrációs problémáknál a tükrözés kerülendő, mindkét megközelítés alkalmas az ilyen megoldások kiszűrésére.

A numerikus stabilitás érdekében az objektumpont-koordinátákat a $[-0.5, 0.5]$ tartományba normalizáltuk a $\mathbf{N}_t$ és $\mathbf{N}_o$ transzformációkkal, amelyek eltolást és skálázást tartalmaznak. Vegyük észre, hogy az (5)-(9) egyenletrendszerek minden ismeretlenje az integrálokon kívül található, így azokat elegendő egyszer kiértékelni. Időbonyolultsága $\mathcal{O}(N)$, ahol N az objektumpontok számát jelöli, mivel minden összegzés a képek egyszeri bejárásával előállítható.

A következő részben a két megoldási megközelítést tárgyaljuk alaposabban.

3.1. Iteratív legkisebb négyzetes megoldás

A legkisebb négyzetes megoldás túlhatározott egyenletrendszerek esetén használatos. A koordinátánkénti szétválaszthatóság nem feltétel, használhatjuk a

[1] Megjegyezzük, hogy természetesen egy túlhatározott egyenletrendszer esetén is lehetséges $n \cdot (n+1)$ egyenletből álló rendszert generálni a transzformáció paraméterei szerint vett parciális deriváltak alkalmazásával. Ekkor viszont már koordinátánként nem lesz szétválasztható az egyenletrendszer n darab független részre, ami olyan magas polinom fokszámhoz vezethet, amelyet hatékonyan már nem tudunk analitikus módszerrel megoldani.

(6)-(9) egyenletrendszert. Ez 2D-ben $2 + 3 + 4 = 9$, míg 3D-ben $3 + 6 + 10 = 19$ egyenletet generál. A numerikus stabilitás érdekében újabb 9, illetve 19 egyenletet veszünk még a Eq. (3)-nek megfelelően, ami a két ponthalmaz szerepének felcserélését jelenti. Megjegyezzük, hogy ez a lépés nem vezet be új ismeretleneket, mivel a $\mathbf{A}$ transzformáció paramétereit nem-szinguláris esetben egyértelműen meghatározzák a q_{ij} elemek. Így egy maximálisan harmadfokú, túlhatározott egyenletrendszerhez jutunk. Ez az egyenletrendszer legkisebb négyzetes értelemben megoldható pl. a Levenberg-Marquardt (L-M) módszerrel.

Azt találtuk, hogy a kezdeti forgatási paraméter értékek erősen befolyásolják a regisztrációs eredményt. Olyan gyakorlati alkalmazásokban, amelyekben ismert, hogy a forgatásbeli különbségek kicsik az objektumok között (pl. a képalkotás protokollja garantálja), vagy jól becsülhetők, az iterációt elegendő egy kezdő paraméterállásból indítani. Ha a fenti feltétel nem garantálható, akkor az optimalizálást több kezdőpontból is elindíthatjuk, és a megoldások közül a legkisebb hibát adót választjuk. 2D-ben 4, 3D-ben 27 különböző kezdő paraméterállást használtunk fel, szisztematikusan az origó körüli 90°-os, valamint a tengelyek körüli 120°-os elforgatásoknak megfelelően. A skálázó, nyíró és eltolási paraméterek az identikus értékeiket kapták.

Az összes keresési irány teljes kiértékelése időigényes, és az esetek egy részében nem is szükséges. Amennyiben a keresést az optimum közeléből indítjuk, az algebrai hiba gyors csökkenést mutat, vagyis kevés iterációszám alatt is az optimum közelébe érünk. Ezért először kevés iterációszámmal indítjuk a keresést és a legkisebb hibát adót választjuk ki, és azzal folytatjuk tovább a teljes keresést. 2D-ben 50, 3D-ben 20 iterációt engedélyezünk az első körben. Egy másik alkalmazott heurisztikánk szerint, ha a kezdeti alacsony iterációszámú szisztematikus keresés esetén az algebrai hiba egy megadott küszöbérték alá esik, akkor nincs szükség további jelöltek keresésre, ezzel folytathatjuk a teljes keresést. Ez a 3D esetben fontos, ahol a küszöbértéket 100-ra állítottuk. 2D-ben csak négy keresési irány van, ezek mindegyikét teszteljük.

Ezeket a paramétereket tapasztalati úton határoztuk meg. További heurisztikák is bevezethetők lennének. Például ne csak egy potenciális jelöltet, hanem többet is megtartsunk, mindegyikből indítsunk teljes keresést, és a legjobbat válasszuk. Várhatóan ez további, igaz kis mértékű, javulást hozhatna. Ez igazából rávilágít az iteratív technika gyenge pontjára: egyensúlyoznunk kell a számítási idő és stabilitás között egy összetett keresési stratégia paramétereinek finomhangolásával.

A keresés közben megszorításokat alkalmazhatunk az $\mathbf{A}$ transzformációra. Egy általános $\mathbf{A}$ transzformáció tükrözéseket is tartalmazhat, amely nem kívánatos a legtöbb gyakorlati alkalmazásban. Ezt elkerülendő, magas algebrai hibát rendelünk azokhoz az esetekhez, ahol $|\mathbf{A}| < 0$, vagyis tükrözés is jelen van. Mivel az $\mathbf{A}$ mátrix segítségével az affinnál alacsonyabb szabadsági fokú transzformációkat is le tudunk írni, könnyen át tudjuk alakítani a keresést merev-test vagy hasonlósági transzformációt igénylő problémák megoldására. Ehhez az algebrai hiba számításakor a kisebb szabadsági fokú transzformáció paramétereiből meg tudjuk határozni az $\mathbf{A}$ affin mátrix elemeit, amit felhasználhatunk változat-

lan formában a kereséshez. Az 1. algoritmus összegzi a javasolt iterációs technika lépéseit.

Algorithm 1: Iteratív megoldási stratégia

Bemenet: Sablon és megfigyelés képek.
Kimenet: Transzformációs paraméterek ($\mathbf{A}$).

Step 1 Objektumpont koordináták normalizálása az $\mathbf{N}_t$, $\mathbf{N}_o$ transzformációkkal
Step 2 Konstruáljuk meg az egyenletrendszert (Eq. (5)–(8) képletek)
Step 3 Legkisebb négyzetes megoldás
- Az LM kereső inicializálása (orientációk, kezdeti iterációszám)
- Minden iterációban számítsuk ki az E algebrai hibát
  Ha szükséges alkalmazzuk a megszorításokat a paraméterekre
  Opcionális: Ha E értéke elegendően alacsony, akkor a további orientációk kihagyása
- $\mathbf{A}^* =$ A legjobb orientációból indított teljes kereséssel eredménye

Step 4 Normalizálás inverze: $\mathbf{A} = \mathbf{N}_o^{-1} \cdot \mathbf{A}^* \cdot \mathbf{N}_t$

3.2. Analitikus megoldás

A közvetlen analitikus megoldáshoz pontosan annyi egyenletre van szükségünk, amekkora a szabadsági foka az n-dimenziós affin transzformációnak, vagyis 2D-ben 6, 3D-ben 12. Mivel a hatékony megoldhatósághoz fontos, hogy koordinátánként szétválasztható egyenletrendszert kapjunk, ezért (9) alapján választunk ω-függvényeket. Így 2D-ben harmadfokú, 3D-ben negyedfokú polinomiális egyenletrendszerhez jutunk.

Az előálló több száz, vagy akár több ezer valós és/vagy komplex megoldás közül ki kell választanunk a legjobbnak tűnő megoldást. Erre különféle kiválasztási sémákat használhatunk.

(a) Csak a valós megoldások figyelembe vétele, a komplexek elhagyásával. A potenciális jelöltek közül azt válasszuk, amelyikre a transzformáció Jacobi determinánsa és a képpontok száma alapján számított közelítő érték a legközelebb van egymáshoz. A korábbi publikációinkban ez a módszer került alkalmazásra [4, 13, 10], és mint kiderült, a csupa komplex megoldást szolgáltató esetek miatt állt elő az ott tapasztalt 5%-os megoldhatatlansági arány. Erre a sémára RealDet néven hivatkozunk a szintetikus tesztekben.

(b) A valós megoldások mellett a komplexeket is vegyük figyelembe, azok valós részeinek használatával. Mivel a komplex megoldások a komplex konjugált párjaikkal jelennek meg, a komplex megoldások felét eleve elvethetjük. Ezután a kiválasztás ugyanúgy megy, mint az (a) esetben. Ez a megoldás mindig garantálja a megoldhatóságot (amennyiben van legalább 1 olyan, amely nem tartalmaz tükrözést). Erre a sémára CplxDet néven hivatkozunk a tesztek

során. Ezt a megközelítést az motiválja, hogy az egyébként valós megoldások komplexszé válhatnak zaj vagy diszkretizálási hiba hatására.

(c) Hasonló (b)-hez, viszont, a megoldás kiválasztásakor a Jacobi determináns helyett azt vizsgáljuk, hogy melyik megoldás adja a legkisebb hibát egy másik, túlhatározott egyenletrendszerbe behelyettesítve. Az egyszerűség kedvéért erre a célra azt a túlhatározott egyenletrendszert használjuk, amelyet a legkisebb négyzetes megoldásnál is. A séma címkéje EqsSel a tesztekben.

Megjegyezzük, hogy mindhárom eljárás képes kiszűrni a nemkívánatos tükröző transzformációkat, amelyek esetében a forgató almátrix determinánsa negatív. Ebből következik, hogy a (b) és (c) sémák csak akkor nem szolgáltatnak eredményt, ha minden megoldás tartalmaz tükrözést. Ilyen eset viszont nem fordult elő a szintetikus tesztjeinkben. A 4. részben megmutatjuk, hogy általában a (c) séma adja a legjobb megoldást, használata 3D-ben elengedhetetlen. A 2. algoritmus összegzi az analitikus módszer lépéseit.

Algorithm 2: Analitikus megoldási stratégia

Bemenet: Sablon és megfigyelés képek.
Output: Transzformációs paraméterek ($\mathbf{A}$).

Step 1 Objektumpont koordináták normalizálása az $\mathbf{N}_t$, $\mathbf{N}_o$ transzformációkkal
Step 2 Konstruáljuk meg az egyenletrendszert (Eq. (9) képlet)
Step 3 Oldjuk meg az egyenletrendszereket
Step 4 Válasszuk ki a legjobbnak tűnő megoldást ($\mathbf{A}^*$)
(a), (b) vagy (c) séma alkalmazásával
Step 5 Normalizálás inverze: $\mathbf{A} = \mathbf{N}_o^{-1} \cdot \mathbf{A}^* \cdot \mathbf{N}_t$

3.3. Kombinált megközelítés

Ahogyan a szintetikus teszteknél látni fogjuk (4. fejezet), az analitikus megoldás általában a globális optimumhoz közeli megoldást szolgáltat a transzformáció magnitúdójától függetlenül, de az iteratív technika túlszárnyalja regisztrációs pontosságban. Mivel ez utóbbit a globális optimum közeléből célszerű indítani, ezért bonyolult, és zajjal terhelt esetben időigényesebb keresési stratégiát igényel. (Ne feledjük, hogy a két esetben más egyenletrendszerek kerülnek megoldásra!) Egy kombinált megközelítéssel kihasználhatjuk mindkét módszer előnyös tulajdonságait: az analitikus megoldásával inicializálhatjuk az iteratív keresőt, kiváltva ezzel annak összetett heurisztikus keresési stratégiáját. Erre a módszerre Combined néven hivatkozunk.

4. Szintetikus tesztek

Egy bináris képekből álló adatbázis segítségével szintetikus teszteken keresztül vizsgáljuk az egyes megoldási stratégiákat és a fuzzy határvonal reprezentáció hatását. A képpárok ismert affin transzformációk alkalmazásával álltak elő.

Algorithm 3: Kombinált megoldási stratégia

Bemenet: Sablon és megfigyelés képek.
Output: Transzformációs paraméterek ($\mathbf{A}$).

Step 1 Objektumpont koordináták normalizálása az $\mathbf{N}_t$, $\mathbf{N}_o$ transzformációkkal
Step 2 Konstruáljuk meg az egyenletrendszert (Eq. (9) képlet)
Step 3 Oldjuk meg az egyenletrendszereket
Step 4 Válasszuk ki a legjobbnak tűnő megoldást ($\mathbf{A}^*$)
(a), (b) vagy (c) séma alkalmazásával
Step 5 Konstruáljuk meg az újabb egyenletrendszert (Eq. (5)–(8) képletek)
Step 6 Legkisebb négyzetes megoldás
$\mathbf{A}^{**}$ = Az optimális transzformáció paraméterei az $\mathbf{A}^*$-ból indított teljes kereséssel adódnak
Step 7 Normalizálás inverze: $\mathbf{A} = \mathbf{N}_o^{-1} \cdot \mathbf{A}^{**} \cdot \mathbf{N}_t$

Az eredmények mérése kétféle hibamértéket vezetünk be. Az ϵ-hiba az átlagos távolságot jelenti a sablon kép képpontjainak az ismert transzformációval vett ($\mathbf{Ap}$), valamint a regisztráció eredményeként kapott transzformációval előálló helye ($\widehat{\mathbf{A}}\mathbf{p}$) között. Ezt a mértéket a szintetikus teszteknél használhatjuk, mivel ott ismert az alkalmazott transzformáció. Gyakorlati problémáknál nem ez a helyzet. A δ-hiba a két alakzat közötti átfedés mértékét jellemzi.

$$\epsilon = \frac{1}{|T|} \sum_{\mathbf{p} \in T} \|\mathbf{Ap} - \widehat{\mathbf{A}}\mathbf{p}\|, \qquad \text{és} \qquad \delta = \frac{|R \triangle O|}{|R| + |O|} \cdot 100\%,$$

ahol $\triangle$ a szimmetrikus különbség, míg T, R és O a sablon, a regisztrált és a megfigyelés alakzatot jelöli. Megjegyezzük, hogy a hibaszámítások előtt a fuzzy reprezentációt binárissá alakítjuk egy alfa-vágással $\alpha = 0.5$ értékkel (a fuzzy tagsági függvény küszöbölésével).

4.1. Kvantitatív kiértékelés 2D-ben

A 2D teszthez 56 különböző alakzatot és transzformált változataikat, összesen 2240 képpárt használtunk fel. Mivel módszerünk jellegéből adódóan nagyméretű képek esetén működik legjobban, a korábbi cikkünkben ilyen képekkel dolgoztunk, a képek szélessége és magassága 500 és 100 pixel közötti volt. Szintén feltételeztük, hogy a fuzzy határvonal reprezentáció alacsonyabb képfelbontás mellett is használhatóvá teszi a módszerünket [13]. Emiatt a jelenlegi tesztünkben a képméreteket nyolcadára csökkentettük az egyes irányok mentén.

A transzformációs paramétereket egyenletes eloszlással, véletlenszerűen generáltuk. A forgatási érték tetszőleges lehetett. A skálázási tényezőt az $[1, 2]$ tartományból választottuk (maximálisan kétszeres nagyítás), de az esetek felében ennek reciprokát véve kicsinyítést hajtottunk végre. A nyírási paraméter tartománya $[-1, 1]$ volt, míg az eltolás maximuma 150 képpont. A sablonok bináris képek voltak, vagyis 0 vagy 1 fuzzy-tagsági értékkel rendelkeztek. A fuzzy határvonal reprezentációt 16×16 méretű túlmintavételezéssel közelítettük, összegezve

milyen arányban található háttér és objektum pont ezen a rácson. A fuzzy-tagság értékeket kvantáltuk k-bites egész értékekre ($(k = 1, \ldots, 8)$), így előállítva a különböző fuzzy reprezentációs szinteket. Néhány sablon és megfigyelés képpár látható az 1. ábrán.

1. ábra: Példák a 2D adatbázisból: Sablon (felső sor) és megfigyelés képek (alsó sor)

A különféle fuzzy reprezentációs szintekhez és kiválasztási sémákhoz tartozó hibaeloszlást és számítási időigényt mutat a 2. ábra, valamint ϵ és δ-hibák mediánjait a 1. táblázat.

Megfigyelhetjük, hogy a tapasztalat alátámasztja az elméleti eredményeket, vagyis a fuzzy reprezentáció jelentősen javítja a regisztrációt.

A korábbi cikkekben egyedüliként alkalmazott RealDet sémát jól láthatóan túlteljesíti az összes javasolt új módszer. Az L-M megoldás az egyértelműen legjobb a bináris esetben. Ez a módszer 8-bites fuzzy szinten is a kisebb medián hibát okoz, mint az analitikus módszerek, de figyeljük meg, hogy az EqsSel módszer 4 és 8-bites esetben jobb elfogadható hibaarányt szolgáltat. 8-bites reprezentációnál a kombinált módszer a legjobb, de érdemes megfigyelni, hogy a bináris esetben elmarad az iteratívtól.

Ez azt sugallja, hogy amennyiben az objektum reprezentáció szinte tökéletes, az analitikus módszer jobb kezdőpozíciót szolgáltat az iteratív kereséshez, mint az ahhoz javasolt keresési stratégia. Amikor viszont a reprezentációs hiba magasabb, az analitikus módszer rosszabb eredményt ad. Így a gyakorlati, hibára hajlamos alkalmazásoknál az iteratív technika választása jobbnak tűnik. A RealDet séma teljesítménye gyenge, az eredmények ≈20–30%-a nem elfogadható, amiben benne van a 6–7%-nyi megoldhatatlan eset is. Ez az arány jóval kisebb a legjobb analitikus (0.4–7.4%), iteratív (1.3–3.9%) és kombinált (0.4–7.0%) módszerek esetében, amelyek ráadásul mindig szolgáltattak eredményt.

Az analitikus módszerek esetén azok kiválasztási hatékonyságát is vizsgáltuk. Mivel pontosan ismerjük az alakzatok közötti geometriai transzformációt, minden lehetséges megoldásra ki tudjuk értékelni az ϵ-hibát és meghatározhatjuk ezek minimumát – ez az elméleti optimum, amit a kiválasztási sémáink el tudnak érni. Azon esetek aránya, amikor a séma által adott megoldáshoz és az optimálishoz tartozó ϵ-hibák távolsága nagyobb, mint 1, 6.52% volt a bináris, és 0.36% a 8-bites reprezentáció esetén. Megállapíthatjuk, hogy a bináris esetben egyik séma sem közelíti meg az optimálisan elérhetőt, de 8-bites fuzzy reprezentációval az EqsSel már összemérhető vele.

1. táblázat: Medián hibák különböző fuzzy reprezentációs szintekhez. Az esetek százaléka, ahol az ϵ-hiba 1 alatti ($\approx$ kitűnő) és 5 alatti ($\approx$ vizuálisan elfogadható) szintén feltüntetésre került. A vastagított számok a kategórián belüli legjobb eredményeket mutatják.

RealDet kiválasztási séma

k	ϵ (pxls.)	δ (%)	$\epsilon > 1$ (%)	$\epsilon > 5$ (%)	Nem reg. (%)
1	0.58	3.65	43.4	29.2	*6.5*
2	0.25	1.44	31.3	23.7	*6.3*
4	0.09	0.61	26.4	21.0	*6.3*
8	0.05	0.34	24.5	20.5	*7.1*

CplxDet kiválasztási séma

k	ϵ (pxls.)	δ (%)	$\epsilon > 1$ (%)	$\epsilon > 5$ (%)	Nem reg. (%)
1	0.47	2.49	34.6	17.3	0
2	0.19	0.96	19.4	9.3	0
4	0.06	0.41	11.8	4.2	0
8	0.04	0.23	7.0	2.5	0

EqsSel kiválasztási séma

k	ϵ (pxls.)	δ (%)	$\epsilon > 1$ (%)	$\epsilon > 5$ (%)	Nem reg. (%)
1	0.39	1.89	25.8	7.3	0
2	0.17	0.88	13.3	3.2	0
4	0.06	0.36	7.5	1.1	0
8	0.03	0.20	3.9	**0.4**	0

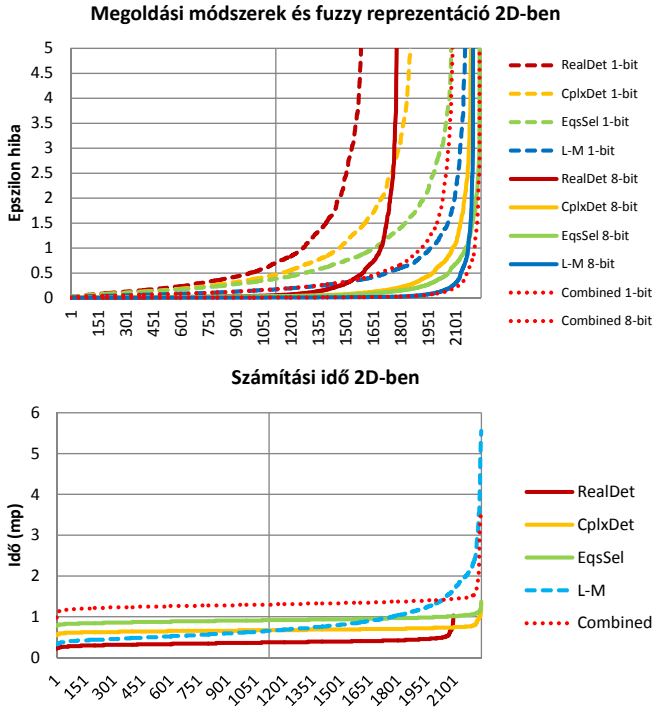
L-M megoldás

k	ϵ (pxls.)	δ (%)	$\epsilon > 1$ (%)	$\epsilon > 5$ (%)	Nem reg. (%)
1	**0.17**	**0.76**	**12.2**	**3.9**	0
2	**0.07**	**0.32**	**6.4**	**2.2**	0
4	**0.02**	**0.13**	3.3	1.4	0
8	**0.01**	**0.08**	2.5	1.3	0

Combined kiválasztási séma

n	ϵ (pxls.)	δ (%)	$\epsilon > 1$ (%)	$\epsilon > 5$ (%)	Nem reg. (%)
1	0.18	0.79	14.6	7.0	0
2	**0.07**	**0.32**	7.0	3.0	0
4	**0.02**	**0.13**	**3.0**	**1.0**	0
8	**0.01**	**0.08**	**1.6**	**0.4**	0

Az algoritmusokat Matlab-ban implementáltuk. Az analitikus megoldásokhoz a *solve*, az iteratív technikához az *fsolve* függvényt használtuk fel. A teszteket egy 2.4 GHz-es Intel Core2 Duo processzorral rendelkező laptopon futtatuk. Az átlagos számítási idő 1 másodperc alatti, kivéve a kombinált módszert, ahol ez 1.3 másodperc körül alakult. Az analitikus megoldók időigénye konstansnak tekinthető. Az iteratív technikánál nagyobb különbségek is kialakulhatnak. A fuzzy reprezentációs szint nem befolyásolja egyik módszer időigényét sem.



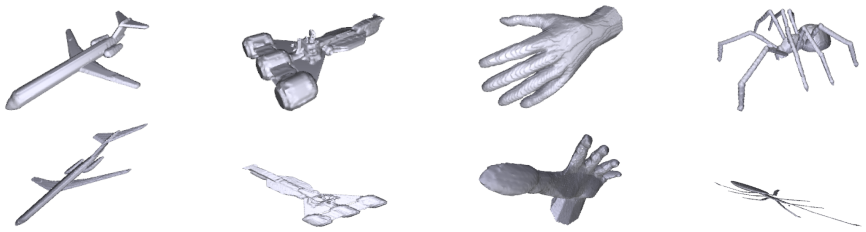
2. ábra: Hiba és számítási idő eloszlások a 2240 esetre. Az értékek a legjobbtól a legrosszabbig lexikografikusan rendezésre kerültek.

4.2. Kvantitatív kiértékelés 3D-ben

A megoldási lehetőségeket 3D objektumokon is teszteltük. 15 objektum összesen 1500 transzformált változatát használtuk. A sablon objektumok 200 ezer és 2 millió közötti képpontszámúak voltak. A transzformáció paraméterek az alábbi tartományokból kerültek ki véletlenszerűen: elforgatás $[0, 2\pi)$; skálázás $[0.5, 1.5]$; nyírás $[-1, 1]$; eltolás $[0, 1]$ (nagyobb eltolásnak az alkalmazott normalizálás miatt nincs szerepe). Minden sablonra 100 véletlenszerű transzformációt hajtottunk végre, amelyek közül 25 merev-test, 25 nem-uniform skálázó, és 50 teljes affin volt. Az 3. ábrán láthatunk néhány példát az adatbázisból.

A fuzzy reprezentációhoz $n \times n \times n$ ($n \in \{1, 2, 4, 8\}$) méretű felülmintavételezést alkalmaztunk képpontonként, a hozzá tartozó fedettségi értékek a rácson vett objektum és háttérpontok számának arányával került közelítésre[2]. A δ-hiba számítása előtt az objektumokat binarizáltuk. A regisztrációs pontosságokat a 2. táblázat, az ϵ-hibák és a számítási idő eloszlását a 4. ábra mutatja.

[2] Megjegyezzük, hogy ez a fuzzy reprezentáció eltér a 2D-ben alkalmazottól, mivel a 3D eset – nagy méretének következtében – más kezelést kívánt.



3. ábra: Példák a 3D adatbázisból: sablon (felső sor) és deformált verziói (alsó sor)

2. táblázat: Medián hiba értékek különböző fuzzy reprezentációs szintekre. Az esetek százaléka, ahol a δ-hiba 1 alatti ($\approx$ kitűnő) és 10 alatti ($\approx$ vizuálisan elfogadható) szintén feltüntetésre került. A vastagított számok a kategórián belüli legjobb eredményeket mutatják.

EqsSel kiválasztási séma

n	ϵ	δ	$\delta > 1\%$	$\delta > 10\%$	Idő (mp)
1	0.2681	1.0613	52.3%	0.87%	2.15
2	0.1164	0.5569	28.1%	**0.07%**	2.14
4	0.1034	0.5106	19.1%	**0.00%**	2.15
8	0.1060	0.5351	18.96%	**0.00%**	2.14

L-M megoldás

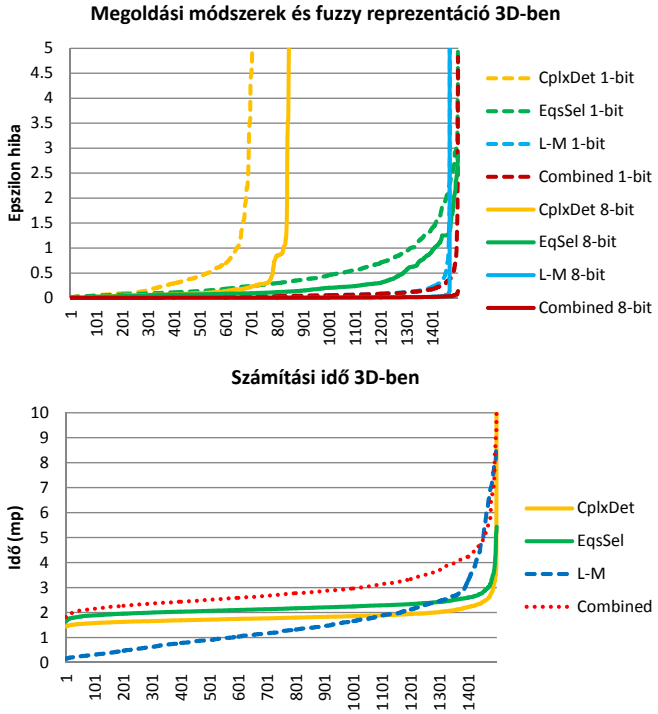
n	ϵ	δ	$\delta > 1\%$	$\delta > 10\%$	Idő (mp)
1	0.0361	0.1555	10.67%	2.13%	**1.54**
2	**0.0108**	0.0627	3.13%	2.27%	**1.56**
4	**0.0069**	0.0470	2.47%	2.20%	**1.54**
8	**0.0065**	**0.0402**	2.47%	2.20%	**1.52**

Combined kiválasztási séma

n	ϵ	δ	$\delta > 1\%$	$\delta > 10\%$	Idő (mp)
1	**0.0359**	**0.1546**	**8.95%**	**0.00%**	2.32
2	**0.0108**	**0.0624**	**1.07%**	**0.07%**	2.42
4	**0.0069**	**0.0469**	**0.33%**	**0.00%**	2.23
8	**0.0065**	**0.0402**	**0.27%**.	**0.00%**	2.31

Az eredmények mutatják, hogy már a bináris reprezentáció is képes kiváló eredményt adni. Viszont már az alacsonyabb fuzzy reprezentációs szintek is tisztán látható javulást hoznak, különösen a kiváló regisztrációs eredmények számában.

Az analitikus megoldás CplxDet sémával nagyon gyengén teljesít, 44-54% közötti a nem elfogadható regisztrációs eredmények aránya a különböző fuzzy szintekre. A RealDet sémát nem is teszteltük. Az iteratív L-M módszer alacsonyabb medián hibát ad, mint a legjobb analitikus módszer, de az elfogadható regisztrációs eredmények számában megelőzi az EqsSel séma. A 2D esettel ellentétben a kombinált módszer jól láthatóan a legjobb választás minden fuzzy reprezentációs szinten, beleértve a bináris esetet is. A feltüntetett számítási



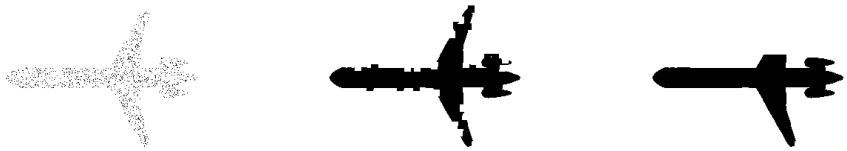
4. ábra: Hiba és számítási idő eloszlások a 3D adatbázis esetén

idők nem tartalmazzák az egyenletrendszer generálásnak idejét, ami átlagosan 1 tizedmásodperc. Az iteratív módszert a Matlab *fsolve* függvényével valósítottuk meg. Az analitikus megoldáshoz a PHCPack csomagot [14] használtuk, mivel a Matlab nem volt képes a negyedfokú polinomiális egyenletrendszer megoldására.

4.3. Robusztusság vizsgálat 3D-ben

A gyakorlatban a képpárokon, objektumokon többféle degradáció is megjelenik, ilyenek például a hiányzó adatok, szegmentációs hibák. A módszerünk robusztusságának tesztelésére három féle degradációtípust modelleztünk a bináris képeinken (5. ábra). A hiányzó képpontok tesztelésére adott százaléknyi képpont került az objektumból véletlenszerűen törlésre. A szegmentációs hiba modellezésére az objektum körvonala mentén véletlenszerűen kis méretű régiókat távolítottunk el az objektumból vagy adtunk hozzá az objektumhoz. A hiányzó térfogatrészek vizsgálatához egy véletlenszerűen kiválasztott objektumpont környezetéből távolítottunk el adott mennyiségű objektumpontot. Terjedelmi okokból itt csak az eredmények értékelését közöljük, részletes táblázatok és grafikonok nélkül.

Az egyenletesen eltávolított objektumpontok általában nem okoznak problémát egyik megoldási módszer esetén sem, akár 90% eltávolítása után is elfo-



5. ábra: Példák képdegradációkra egy 3D objektum 2D szeletén: Képpontok 90%-nak véletlenszerű eltávolítása (balról), képpontok 10%-nak módosítása a határ mentén kis méretű régiókban (középen), és a 5%-nak eltávolítása egy képpontból kiindulva (jobbról).

gadhatók az eredmények. A szegmentációs hiba esetében 25%-os degradációs szintnél az esetek 70%-ban az objektumok illeszkedése elfogadható. Mint általában a terület-alapú módszerek, a mi megközelítésünk is érzékeny a hiányzó térfogatrészekre. $1-2\%$-os degradációs szint felett az eredmények megbízhatatlanná válnak.

Azt tapasztaltuk, hogy az analitikus megoldás önmagában nem versenyképes az iteratív technikával. A kombinált megközelítés felülmúlja az iteratívat kis mértékű degradációk esetén. A degradáció mértékének növekedésével ez az előny viszont eltűnik. Észrevehetjük azt is, hogy mivel az algebrai hiba magasabb degradációs szint esetén jelentősen növekedhet, több kezdő orientáció vizsgálatára van szükség, ami megnöveli a számítási időt. Összefoglalásként elmondhatjuk, hogy a módszerünk rendszerint jól teljesít, amíg az objektumunk geometriai momentumai nem változnak meg nagymértékben a degradáció hatására.

5. Összegzés

Cikkünkben egy egyesített regisztrációs keretrendszert mutattunk be amely képes 2D és 3D bináris objektumok között többféle lineáris deformáció becslésére. A regisztrációs probléma megoldását egy polinom egyenletrendszer generálása és megoldása adja. Új eredményünk a regisztrációs eljárás során előálló polinomiális egyenletrendszerek gyors és hatékony megoldási lehetőségeinek részletes tárgyalása. Azt tapasztaltuk, hogy az iteratív legkisebb négyzetes megoldás szisztematikus kereső eljárással kombinálva a legjobb választás 2D problémákra, míg 3D-ben az analitikus és az iteratív technikák kombinációját célszerű választani. Tesztjeinkben vizsgáltuk a korábbi publikációinkban bevezetett fuzzy reprezentációt, amely a diszkrét objektumok képpont lefedettségi információját jelenti, és kimutattuk ennek előnyeit.

Köszönetnyilvánítás

A szegedi szerzők munkáját az OTKA K75637 számú pályázat és a futurICT.hu nevű, TÁMOP-4.2.2.C-11/1/KONV-2012-0013 azonosítószámú projekt támogatta az Európai Unió és az Európai Szociális Alap társfinanszírozása mellett.

N. Sladoje és J. Lindblad munkáját anyagilag támogatta a Szerb Köztársaság Tudományos Minisztériuma az ON144018 és ON144029 projekteken keresztül.[3] Köszönetet mondunk Niku Corneának, aki a szintetikus képek egy részét biztosította számunkra.

Irodalom

1. B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, October 2003.
2. R. Hagege and J.M. Francos, "Linear estimation of sequences of multi-dimensional affine transformations," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, Toulouse, France, May 2006, IEEE, vol. 2, pp. 785–788.
3. R. Hagege and J.M. Francos, "Parametric estimation of affine transformations: An exact linear solution," *Journal of Mathematical Imaging and Vision*, vol. 37, no. 1, pp. 1–16, 2010.
4. C. Domokos, Z. Kato, and J.M. Francos, "Parametric estimation of affine deformations of binary images," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, Nevada, USA, April 2008, IEEE, pp. 889–892.
5. J. Kannala, E. Rahtu, J. Heikkilä, and M. Salo, "A new method for affine registration of images and point sets," in *Proc. of the 14th Scandinavian Conference on Image Analysis (SCIA)*. 2005, LNCS, pp. 224–234, Springer.
6. T. Suk and J. Flusser, "Affine normalization of symmetric objects," in *Proc. of the 7th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS)*. 2005, LNCS, pp. 100–107, Springer.
7. J. Zhang, Y. Ge, S.H. Ong, C.K. Chui, S.H. Teoh, and C.H. Yan, "Rapid surface registration of 3D volumes using a neural network approach," *Image Vision Comput.*, vol. 26, no. 2, pp. 201–210, 2008.
8. N.D. Cornea, M.F. Demirci, D. Silver, A. Shokoufandeh, S.J. Dickinson, and P.B. Kantor, "3D object retrieval using many-to-many matching of curve skeletons," *Int. Conf. on Shape Modeling and Appl.*, pp. 368–373, 2005.
9. J. Lindblad, V. Ćurić, and N. Sladoje, "On set distances and their application to image registration," in *Proc. of the 6th International Symposium on Image and Signal Processing and Analysis (ISPA)*. 2009, pp. 449–454, IEEE.
10. C. Domokos and Z. Kato, "Parametric estimation of affine deformations of planar shapes," *Pattern Recognition*, vol. 43, no. 3, pp. 569–578, 2010.
11. A. Tanács, J. Lindblad, N. Sladoje, and Z. Kato, "Estimation of linear deformations of 3D objects," in *Proc. of International Conference on Image Processing (ICIP)*. 2010, pp. 153–156, IEEE.
12. N. Sladoje and J. Lindblad, "Pixel coverage segmentation for improved feature estimation," in *Proc. of ICIAP*. 2009, vol. 5716 of *LNCS*, pp. 929–938, Springer.
13. A. Tanács, C. Domokos, N. Sladoje, J. Lindblad, and Z. Kato, "Recovering affine deformations of fuzzy shapes," in *Proc. of SCIA*. 2009, vol. 5575 of *LNCS*, pp. 735–744, Springer.
14. J. Verschelde, "Algorithm 795: Phcpack: a general-purpose solver for polynomial systems by homotopy continuation," *ACM Trans. Math. Softw.*, vol. 25, no. 2, pp. 251–276, 1999.

[3] N. Sladoje and J. Lindblad are financially supported by the Ministry of Science of the Republic of Serbia through Projects ON144018 and ON144029 of MI-SANU.