

Szegedi Tudományegyetem
Informatikai Intézet
SZTE Informatika Doktori Iskola

Vizuális objektumok affin és nem-lineáris regisztrációja

Ph.D. értekezés tézisei

Sánta Zsolt

Témavezető:
Dr. Kató Zoltán
egyetemi tanár

Szeged
2018

Bevezetés

A számítógépes látás területén vizuális objektumok számítógépes értelmezésével és vizsgálatával foglalkozunk. A cél olyan automatizált rendszerek kidolgozása, amelyek képesek hatékonyan megvalósítani az élőlények vizuális érzékelését. Történetileg az első alkalmazások 2D képek feldolgozásával foglalkoztak. Napjainkban a modern 3D képalkotó berendezések, mint a Lidar, a Microsoft Kinect™ vagy az Intel® RealSense™ elterjedésével a 3D adatok feldolgozása egyre nagyobb fontosságot kap.

A legtöbb számítógépes látásbeli problémában egy fontos előkészítő lépés a *képregisztráció*, amely önmagában is érdekes terület. Itt a fő célunk, hogy leképezéseket határozzunk meg az aktuálisan vizsgált színtér különböző megfigyelései között. A leggyakoribb problémában a bemeneti adatok koordináta-rendszereit szinkronizáljuk, meghatározva a közös origót és a fő tengelyeket. Bár sok probléma esetén elegendő ez az egyszerűbb illesztés, komplexebb deformációk kezelése is gyakori probléma. Bár a szakirodalomban bizonyos problémákra találhatunk hatékony megoldásokat, több nyitott kérdéssel is találkozhatunk.

A disszertáció összefoglalja a Szerző kutatási eredményeit a képregisztráció több alterületéről. A javasolt eljárások transzformációk széles körét lefedik, mint például affin, perspektív és nem-lineáris modellek.

Háromdimenziós objektumok nem-lineáris regisztrációja

Az alábbi szekcióban javasolunk egy általános keretrendszert 3D objektumok közötti nem-lineáris regisztrációjára. A továbbiakban felszín illesztéssel foglalkozunk, ahol a fizikai deformáció kizárólag az objektumok felszínén mérhető. Mivel a javasolt algoritmus csak geometriai információkkal dolgozik, egy bináris regisztrációs problémáról beszélhetünk. A javasolt keretrendszer egy sikeres 2D algoritmus alapötletét használja [1], de a kiterjesztés nem egyértelmű: több olyan 3D specifikus sajátossággal kellett foglalkozni, amelyek 2D-ben nem jelentkeznek. A javasolt keretrendszert kidolgoztuk voxeles [Sánta and Kato, 2012a][Sánta and Kato, 2012b], valamint nyitott és zárt háromszögfelszíni hálókkal megadott objektumokra [Sánta and Kato, 2018][Sánta and Kato, 2016a][Sánta and Kato, 2013a]. Lehetséges valós alkalmazásokként tüdő CT felvételek, agyi felszínek és emberi arcok illesztését prezentáljuk.

Először formalizáljuk az illesztési problémát: adott egy *sablon* és egy *megfigyelés* objektum, melyeket rendre $\mathcal{F}_t \subset \mathbb{R}^3$ és $\mathcal{F}_o \subset \mathbb{R}^3$ jelöl. Keressük a két objektum között ható φ transzformációt, amely tetszőleges $\mathbf{x} \in \mathcal{F}_t$ esetén létezik olyan $\mathbf{y} \in \mathcal{F}_o$ amely teljesít az alábbi relációt:

$$\varphi(\mathbf{x}) = \mathbf{y} \tag{1}$$

A klasszikus pontmegfeleltetés alapú módszerek nagy számú, a fenti relációt kielégítő pontpárokat nyernek ki, majd ezeket felhasználva meghatározzák a keresett φ transzformáció paramétereit. A javasolt módszerben ehelyett kiintegráljuk az egyes megfeleltetéseket az objektumok felett, amely az alábbi formulához vezet [Sánta and Kato, 2018][Sánta and Kato, 2013a]:

$$\int_{\mathcal{F}_o} \mathbf{y} \, d\mathbf{y} = \int_{\varphi(\mathcal{F}_t)} \mathbf{z} \, d\mathbf{z}. \quad (2)$$

Míg a pontmegfeleltetés alapú módszereknél minden új megfeleltetés egy új egyenletet generál, a 2. formula csak meghatározott számú (\mathbb{R}^3 esetén pontosan három) egyenletet állít elő. Ahhoz, hogy több egyenletet generálhassunk, vegyük észre, hogy az 1. egyenlet (és így a 2. egyenlet is) továbbra is teljesül, ha egy megfelelően konstruált nem-lineáris $\omega : \mathbb{R}^3 \rightarrow \mathbb{R}$ függvényt alkalmazunk mindkét oldalon [1]. Így, alkalmazva független, nem-lineáris függvények $\{\omega_i\}_{i=1}^\ell$ halmazát egy ℓ egyenletből álló egyenletrendszert kaphatunk [Sánta and Kato, 2018][Sánta and Kato, 2013a]:

$$\int_{\mathcal{F}_o} \omega_i(\mathbf{y}) \, d\mathbf{y} = \int_{\varphi(\mathcal{F}_t)} \omega_i(\mathbf{z}) \, d\mathbf{z} \quad i = 1, \dots, \ell. \quad (3)$$

A keresett transzformáció, φ , paramétereit az egyenletrendszer ismeretlenjeként állnak elő. Ebből következően ℓ értéke legalább N , amely a transzformáció paramétereinek számát jelöli. Mivel a gyakorlatban számíthatunk arra, hogy a bemeneti adatot különböző zajok terhelhetik, érdemes egy túlhatározott rendszert generálni, így $\ell \geq N$. Ezután a kapott rendszert legkisebb négyzetes értelemben megoldjuk egy tetszőleges iteratív optimalizáló segítségével. A számítás megkönnyítése érdekében alacsony fokszámú hatványfüggvényeket választunk az $\{\omega_i\}_i^\ell$ halmaz elemeinek:

$$\omega_i(\mathbf{x}) = x_1^{n_i} x_2^{m_i} x_3^{o_i}, \quad (4)$$

ahol $\{(n_i, m_i, o_i)\}_{i=1}^\ell = \{(a, b, c) \mid a + b + c = O\}$ és $O \in \{0, \dots, O_{max}\}$.

Háromdimenziós objektumokkal dolgozva a gyakorlatban előforduló tartományok vagy felszínek vagy térfogatok lesznek. A leggyakoribb felszín reprezentáció a háromszögfelszíni háló, amellyel leírhatunk felszíni (nyílt felszín) vagy volumetrikus (zárt felszín) objektumokat. Tételezzük fel, hogy a 3. egyenletben háromszöghálók feletti felszíni integrálokat alkalmazunk, ekkor az egyenletek kifejezhetők a háló háromszögei feletti integrálok összegeként:

$$\sum_{o \in O_\Delta} \int_o \omega_i(\mathbf{y}) \, d\mathbf{y} \approx \sum_{\pi \in \varphi(T_\Delta)} \int_\pi \omega_i(\mathbf{z}) \, d\mathbf{z}, \quad (5)$$

ahol $T_\Delta \approx \mathcal{F}_t$, $O_\Delta \approx \mathcal{F}_o$ és $\varphi(T_\Delta) \approx \varphi(\mathcal{F}_t)$. A formula számítása lineáris a háromszögek

számában, a leírt rekurzív numerikus formulák alkalmazásával (egzakt és approximációs számítási módok).

Hasonló formalizmus vezethető le volumetrikus objektumokra zárt háromszögfelszíni hálók esetén. Ekkor az integrál a hálók háromszögeiből konstruált tetraéderek feletti integrálok összege lesz:

$$\sum_{o \in O_\Delta} \text{sgn}(\text{vol}(T_o)) \int_{T_o} \omega_i(\mathbf{y}) d\mathbf{y} \approx \sum_{\pi \in \varphi(T_\Delta)} \text{sgn}(\text{vol}(T_\pi)) \int_{T_\pi} \omega_i(\mathbf{z}) d\mathbf{z}, \quad (6)$$

ahol T_o az o háromszögből és az origóból generált tetraéder, míg T_π a T_o tetraéder által meghatározott folytonos tartomány lesz.

Volumetrikus objektumok esetén használhatjuk a voxel reprezentációt is. Itt a bemeneti adat identikus térfogati elemekből épül fel (amelyeket voxelnek hívunk) és minden elem térfogata 1 lesz. Vezessük be a F_t és F_o jelöléseket a *sablon* és *megfigyelés* objektumokra. Mivel itt volumetrikus objektumokkal dolgozunk, a felhasznált integrálok térfogati integrálok lesznek [Sánta and Kato, 2012a][Sánta and Kato, 2012b]. Ebből következően, a 3. egyenletből az alábbi egyenleteket kapjuk

$$\sum_{\mathbf{Y} \in F_o} \omega_i(\mathbf{Y}) \approx \int_{F_o} \omega_i(\mathbf{y}) d\mathbf{y}, \quad (7)$$

$$\sum_{\mathbf{Z} \in \varphi(F_t)} \omega_i(\mathbf{Z}) \approx \int_{\varphi(F_t)} \omega_i(\mathbf{z}) d\mathbf{z}, \quad (8)$$

ahol $i = 1, \dots, \ell$. Ugyanakkor a $\varphi(F_t)$ tartomány előállítás nem hatékony nagy méretű képek esetén, ezért alkalmazzuk az integráltranszformációt az egyenletekben. Ezzel az egyenlet formája a következő lesz:

$$\sum_{\mathbf{Y} \in F_o} \omega_i(\mathbf{Y}) = \sum_{\mathbf{X} \in F_t} \omega_i(\varphi(\mathbf{X})) |J_\varphi(\mathbf{X})|, \quad (9)$$

ahol $i = 1, \dots, \ell$ és J_φ a φ transzformáció Jacobi mátrixa. Míg a javasolt keretrendszer bármilyen parametrikus transzformációs modell esetén alkalmazható [Sánta and Kato, 2012a][Sánta and Kato, 2012b], polinomiális transzformációk esetén egy hatékony hatékony numerikus számítási módot kaphatunk, mivel az egyenletek jobb oldalán polinomokkal dolgozunk [Sánta and Kato, 2012a].

A javasolt algoritmusokat nagy elemszámú szintetikus generált adathalmazokon teszteltük. A háromszögfelszíni hálós reprezentációnál vékony fémlemez spline (TPS) modellt, míg a voxeles esetben TPS és polinomiális transzformációkat használtunk. Az eljárások robusztusságát igazoltuk szegmentálási hibákkal és modell túlillesztéssel szemben. Megvizsgáltuk a TPS modellnél alkalmazott kontrollpontok pozícióinak jelentőségét az illesztés pontosságára nézve. A javasolt eljárások hatékonyabbnak bizonyultak két

korábban publikált algoritmusnál (CPD [2] és GMMREG [3]). Végül az algoritmusok meggyőző eredményeket értek el tüdő CT, agyi felszín és 3D arcok illesztésében.

Kétdimenziós képek robusztus regisztrációja

Az alábbiakban ismertetjük a 2D képek regisztrációjában elért eredményeinket. A munka során két különböző problémával foglalkoztunk, mindkettő a korábban vázolt algebrai keretrendszerrel kapcsolatos. Az első problémában a módszer nagyobb mértékű, nem egyenletes szegmentálási hibákkal szembeni toleranciáját vizsgáltuk [Sánta and Kato, 2014].

A második problémában az alakzatok belső struktúrájának megőrzésével foglalkoztunk affin és nem-lineáris transzformációk esetén. Ezt a problémát korábban regularizációval kezeltük, itt a képek intenzitásértékeit használtuk a kidolgozott algebrai rendszerben [Sánta and Kato, 2016b].

Mindkét kidolgozott megoldást nagy méretű szintetikus és valós adathalmazon teszteltük. Emellett összehasonlítottuk eredményeiket az elmúlt években publikált módszerek eredményeivel.

Regisztrációs keretrendszer

A szekcióban bemutatásra kerülő módszereket a [1, 4] cikkekben prezentált 2D keretrendszerre építve dolgoztuk ki. A feladat itt a *sablon* és a *megfigyelés* 2D képrégiók közötti transzformáció meghatározása. Vezessük be a régiók pontjainak jelölésére a $\mathcal{F}_t \subset \mathbb{R}^2$ és $\mathcal{F}_o \subset \mathbb{R}^2$ halmazokat. A cél egy az alkalmazásnak megfelelő $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ transzformáció meghatározása, amely a *sablon* képet a *megfigyeléshez* illeszti.

A fenti jelöléseket használva, felírhatjuk az alábbi relációt egy tetszőleges $\mathbf{x} = [x_1, x_2]^T \in \mathcal{F}_t$ és $\mathbf{y} = [y_1, y_2]^T \in \mathcal{F}_o$ pontpár között:

$$\varphi(\mathbf{x}) = \mathbf{y} \tag{10}$$

Míg a reláció koordinátáinként egy-egy korlátot ad a keresett transzformációra, a gyakorlatban ilyen megfeleltetések előállítása nehéz feladat, főleg nagy szabadságfokú deformációk esetén.

Követve a [1, 4] cikkekben kidolgozott keretrendszert, a pontmegfeleltetések kinyerése helyett integráljuk a 10. reláció mindkét oldalát az alakzatok felett:

$$\int_{\varphi(\mathcal{F}_t)} \mathbf{z} \, d\mathbf{z} = \int_{\mathcal{F}_o} \mathbf{y} \, d\mathbf{y}. \tag{11}$$

Mivel a transzformált tartomány legenerálása időigényes lehet, alkalmazzuk az integrál-

transzformációt a 11. egyenletben, hasonlóan a 3D keretrendszerhez:

$$\int_{\mathcal{F}_t} \varphi(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x} = \int_{\mathcal{F}_o} \mathbf{y} dy. \quad (12)$$

Ekkor az integráltranszformáció behozza a φ transzformáció Jacobi determinánsát az egyenletek bal oldalán, ami a transzformáció hatására bekövetkező relatív területváltást fejezi ki. A $J_\varphi(\mathbf{x})$ mátrixot a transzformáció parciális deriváltjaiból képezzük.

Mivel ez az egyenlet sem szolgáltat elegendő korlátot a paraméterek meghatározására, alkalmazzunk az egyenlet mindkét oldalán egy független, nem-lineáris $\{\omega_i \mid \omega_i : \mathbb{R}^2 \rightarrow \mathbb{R}\}_{i=1}^\ell$ halmazt a 11. és a 12. egyenletekben:

$$\int_{\varphi(\mathcal{F}_t)} \omega_i(\mathbf{z}) dz = \int_{\mathcal{F}_o} \omega_i(\mathbf{y}) dy, \quad (13)$$

$$\int_{\mathcal{F}_t} \omega_i(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})| d\mathbf{x} = \int_{\mathcal{F}_o} \omega_i(\mathbf{y}) dy, \quad (14)$$

ahol $i = 1, \dots, \ell$. Polinom ω függvények polinom egyenletrendszereket eredményeznek, és a [1, 4] cikkekben leírtak szerint a kapott keretrendszer robusztus a legtöbb geometriai és szegmentálási zajjal szemben.

Viszont a korábbi munkákban a nagyobb mértékű szegmentálási hibák (kitakart vagy hozzávett alakzatrészek) esetei egy triviális hibaforrásként voltak kezelve. Sok ipari és orvosi alkalmazásban kizárólag a deformációval kell dolgoznunk, mivel a képkészítés körülményei jól kontrolláltak. Ezzel szemben nem kontrollált környezetben (pl. megfigyelőrendszerek képeinek feldolgozásakor) mindig számíthatunk a takarás problémájának előfordulására [Sánta and Kato, 2014].

Kitakart alakzatok affin illesztése

Az előzőekben definiált keretrendszer alkalmazásához, tegyük fel, hogy a keresett fizikai deformáció közelíthető egy $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ homogén affin mátrix segítségével:

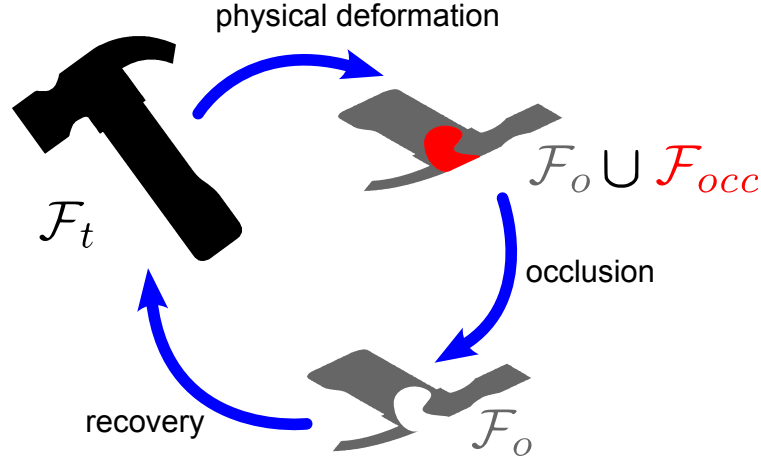
$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}. \quad (15)$$

Az affin transzformációk a legtöbb valós alkalmazásban jól használhatók, ahol lineáris vagy perspektív torzulásokkal dolgozunk. Ez utóbbi esetben az affin modell a deformáció elsőrendű approximációjaként jelenik meg, ami egy kényelmes lehetőség lyukkamerákkal készített képek regisztrációjakor, ha a kamera és az objektum távolsága nagy az objektum méretéhez képest [5].

Alkalmazva a fenti jelöléseket, a 10. egyenletben megadott reláció a következő

lesz [Sánta and Kato, 2014]:

$$\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{y}} \Leftrightarrow \bar{\mathbf{x}} = \mathbf{A}^{-1}\bar{\mathbf{y}}. \quad (16)$$



1. ábra. A kapcsolat a sablon és a kitakart megfigyelés között.

Míg tökéletes alakzatok esetén a fenti reláció teljesül, takarások előfordulásával az állítás már nem lesz igaz. Jelöljük a megfigyelés kitakart részeit \mathcal{F}_{occ} módon (lásd példaként az 1. ábrán) [Sánta and Kato, 2014]. Ezután írjuk fel az alakzat szintű relációt kitakart objektumokra:

$$\mathbf{A}(\mathcal{F}_t) = \mathcal{F}_o \cup^* \mathcal{F}_{occ} \Leftrightarrow \mathcal{F}_t = \mathbf{A}^{-1}(\mathcal{F}_t \cup^* \mathcal{F}_{occ}), \quad (17)$$

ahol \cup^* jelöli két halmaz diszjunkt unióját (ahol biztosan tudjuk, hogy a két halmaz metszete üres). Alkalmazva a fenti tartományokat a 14. egyenletben az alábbi rendszerhez vezet:

$$\int_{\mathbf{A}(\mathcal{F}_t)} \omega_i(\bar{\mathbf{z}}) d\bar{\mathbf{z}} = \int_{\mathcal{F}_o \cup^* \mathcal{F}_{occ}} \omega_i(\bar{\mathbf{y}}) d\bar{\mathbf{y}} \quad (18)$$

$$\int_{\mathcal{F}_t} \omega_i(\bar{\mathbf{x}}) d\bar{\mathbf{x}} = \int_{\mathbf{A}^{-1}(\mathcal{F}_o \cup^* \mathcal{F}_{occ})} \omega_i(\bar{\mathbf{p}}) d\bar{\mathbf{p}}, \quad (19)$$

ahol $i = 1, \dots, \ell$. Az egyszerűség kedvéért tegyük fel, hogy az ω_i függvények az alakzatok pontjainak Descartes-féle koordinátáin hatnak csak. Vegyük észre, hogy tetszőleges nem nulla \mathbf{A} affin transzformáció esetén teljesül a következő:

$$\mathbf{A}^{-1}(\mathcal{F}_o \cup^* \mathcal{F}_{occ}) = \mathbf{A}^{-1}(\mathcal{F}_o) \cup^* \mathbf{A}^{-1}(\mathcal{F}_{occ}) \quad (20)$$

Behelyettesítve a fenti megfigyelést a 18. és a 19. egyenletbe, majd alkalmazva a

Lebesgue integrálok alaptulajdonságait a következőket kapjuk:

$$\int_{\mathbf{A}(\mathcal{F}_t)} \omega_i(\bar{\mathbf{z}}) d\bar{\mathbf{z}} = \int_{\mathcal{F}_o} \omega_i(\bar{\mathbf{y}}) d\bar{\mathbf{y}} + \int_{\mathcal{F}_{occ}} \omega_i(\bar{\mathbf{y}}') d\bar{\mathbf{y}}' \quad (21)$$

$$\int_{\mathcal{F}_t} \omega_i(\bar{\mathbf{x}}) d\bar{\mathbf{x}} = \int_{\mathbf{A}^{-1}(\mathcal{F}_o)} \omega_i(\bar{\mathbf{q}}) d\bar{\mathbf{q}} + \int_{\mathbf{A}^{-1}(\mathcal{F}_{occ})} \omega_i(\bar{\mathbf{q}}') d\bar{\mathbf{q}}' \quad (22)$$

ahol $i = 1, \dots, \ell$. A [4] cikkben ismertetett affin keretrendszerhez hasonlóan, egyszerű hatványfüggvényeket használunk az $\{\omega_i\}_{i=1}^{\ell}$ halmazban és $\ell \geq 6$ választása esetén túlhatározott egyenletrendszert kapva, megoldjuk legkisebb négyzetes értelemben [Sánta and Kato, 2014].

A kapott keretrendszerben több algebrailag ekvivalens megoldást kaphatunk. Egy triviális megoldás például, ha a transzformációnak a nulla mátrixot választjuk és a teljes *megfigyelést* takarásként kezeljük. A probléma kezeléséhez regularizációt alkalmazunk. Itt az egyenletrendszer megoldása mellett minimalizáljuk a takarás méretét is, ami az alábbi probléma megoldásához vezet:

$$\int_{\mathcal{F}_t} \omega_i(\bar{\mathbf{x}}) d\bar{\mathbf{x}} = \int_{\mathbf{A}^{-1}(\mathcal{F}_o)} \omega_i(\bar{\mathbf{q}}) d\bar{\mathbf{q}} + \int_{\mathbf{A}^{-1}(\mathcal{F}_{occ})} \omega_i(\bar{\mathbf{q}}') d\bar{\mathbf{q}}' \quad (23)$$

such that $|\mathbf{A}^{-1}(\mathcal{F}_{occ})| \rightarrow \min \quad i = 1, \dots, \ell,$

ahol $|\mathbf{A}^{-1}(\mathcal{F}_{occ})|$ jelöli a $\mathbf{A}^{-1}(\mathcal{F}_{occ})$ tartomány területét. Megjegyezzük, hogy a 22. egyenlet segítségével egy hasonló probléma határozható meg, amely ekvivalens lesz a 23. egyenletben ismertetett problémával.

A számítási komplexitás csökkentése érdekében az objektumokat poligonokkal határolt tartományokként kezeljük [Sánta and Kato, 2014]. Jelöljük az \mathcal{F}_t , \mathcal{F}_o és \mathcal{F}_{occ} alakzatok kontúrjait közelítő poligonokat rendre, T_{\diamond} , O_{\diamond} és M_{\diamond} módon. Ezeket az approximációkat használva, a kitakart területek meghatározhatók egyszerű Boole műveletek segítségével:

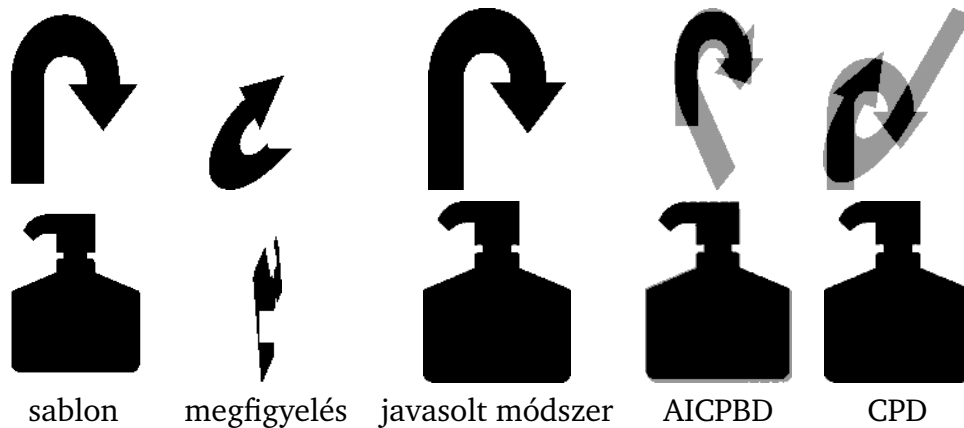
$$\mathbf{A}^{-1}(M_{\diamond}) = T_{\diamond} \setminus \mathbf{A}^{-1}(O_{\diamond}), \quad (24)$$

$$M_{\diamond} = \mathbf{A}^{-1}(O_{\diamond}) \setminus \mathbf{A}(T_{\diamond}), \quad (25)$$

ahol \setminus jelöli a két poligon különbségét. A hatványfüggvények integrálása poligonok felett hasonlóan hatékonyan végezhető, mint az integrálás 3D-ben háromszögfelszíni hálók felett. A komplexitás itt is minimális lesz a megfelelő rekurzív formulák alkalmazásával. Ezért a $\{\omega_i\}$ halmazt az alábbiak szerint definiáljuk:

$$\omega_i(\bar{\mathbf{x}}) = x_1^{n_i} x_2^{m_i}, \quad (26)$$

ahol $\{(n_i, m_i)\}_{i=1}^{\ell} = \{(a, b) \mid a + b = O\}$ és $O \in \{0, \dots, O_{max}\}$.



2. ábra. Példák a szintetikus adathalmazból. Az utolsó három oszlopban a *sablont* feketével, míg az illesztési hibákat szürkével jelöljük.

A tesztleinkben egy 9 egyenletből álló rendszert generáltunk, így $O_{max} = 3$ volt. Az algoritmust Matlab-ban implementáltuk (a poligonokon végzett Boole műveleteket MEX interfészekon keresztül valósítottuk meg). A kvantitatív kiértékeléshez egy 3000 szintetikus képből álló adathalmazon teszteltünk, amely elérhető a [6] módszer honlapján (<http://www.inf.u-szeged.hu/~kato/software/affbinregdemo.html>). A kitakarások szimulálására véletlenszerűen töröltük az objektumok 10%, 20% és 30% egybefüggő területeit. Az eredményeket összehasonlítottuk két korábban publikált módszer eredményeivel (AICPBD [7] és CPD [2]).

Összefoglalva, a javasolt algoritmus hatékonyan meghatározza a keresett transzformációt, ha az jól közelíthető affin modellel, és a kitakarás mértéke 20% alatt van.

Kovariáns függvények affin és nem-lineáris regisztrációja

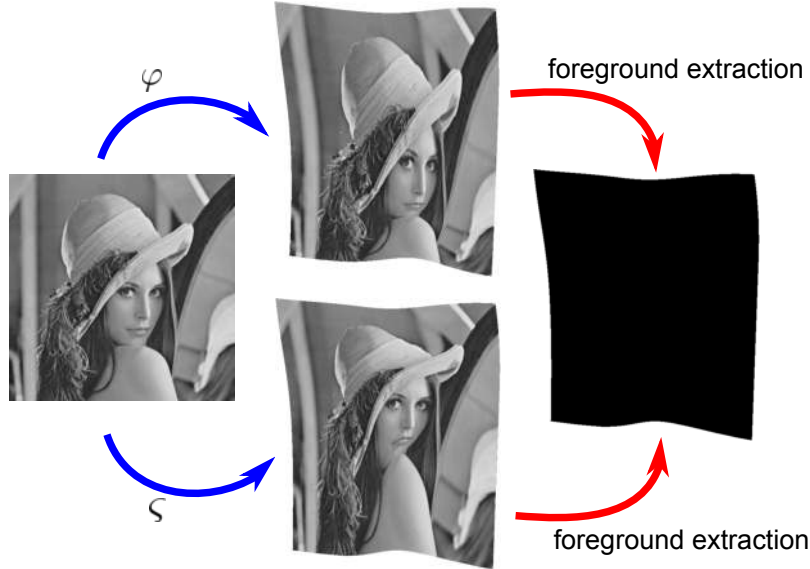
A szekció második témájában egy hasonló keretrendszert javasolunk képi alakzatok regisztrációjára affin és nem-lineáris modellek segítségével [Sánta and Kato, 2016b]. Az eddigiekben kizárólag geometriai információt használtunk a transzformáció meghatározására. Bár több gyakorlati esetben ez elegendőnek bizonyul, sokszor előfordul, hogy a fizikai deformáció nem mérhető a geometriai adatokból.

Egy gyakori megoldás az ilyen problémák esetére, ha több információt állítunk elő a megoldó számára. Amennyiben 2D képekkel dolgozunk, a legegyszerűbb megoldás az intenzitás információk használata.

Tegyük fel, hogy adott két intenzitás függvény a *sablon* és *megfigyelés* alakzatokhoz: $T : \mathcal{F}_t \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ és $O : \mathcal{F}_o \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Feltételezve, hogy a 10. reláció teljesül, ezek a függvények kovariánsak lesznek φ transzformációra nézve, így [Sánta and Kato, 2016b]:

$$T \circ \varphi = O \quad (27)$$

$$T(\mathbf{x}) = O(\varphi(\mathbf{x})) = O(\mathbf{y}). \quad (28)$$



3. ábra. Az alakzat alapú regisztrációk bizonytalansága. Ebben a példában két különböző nem-lineáris transzformációt alkalmaztunk ugyanarra a képre, amely két eltérő megfigyelést eredményezett. Ugyanakkor, szegmentálás után ugyanazt a maszkot határoztuk meg, így egy tisztán bináris algoritmus identikusan illeszti a két megfigyelést.

A javasolt módszerben a cél kombinálni a geometriai módszer robusztusságát az intenzitásértékek regularizáló hatásával [Sánta and Kato, 2016b]. Ehhez egyszerűen szorozzuk meg a 10. egyenletet a 28. egyenlettel:

$$\int_{\mathcal{F}_t} \varphi(\mathbf{x}) T(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x} = \int_{\mathcal{F}_o} \mathbf{y} O(\mathbf{y}) d\mathbf{y}. \quad (29)$$

Ezután, a 14. rendszerhez hasonlóan, alkalmazzuk az $\{\omega_i\}$ halmazt a koordinátákra:

$$\int_{\mathcal{F}_t} \omega_i(\varphi(\mathbf{x})) T(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x} = \int_{\mathcal{F}_o} \omega_i(\mathbf{y}) O(\mathbf{y}) d\mathbf{y}, \quad (30)$$

ahol $i = 1, \dots, \ell$. A fenti összevont egyenletrendszer legkisebb négyzetes értelemben megoldjuk, ahol az ismeretlenek adják a keresett transzformáció paramétereit [Sánta and Kato, 2016b].

A módszer hatékonyságát kísérleti úton igazoltuk szintetikus és valós tesztalmozakon. Az előbbiekhöz egy véletlenszerűen generált halmazt állítottunk elő affin és TPS modellek alkalmazásával.

A javasolt eljárást sikeresen alkalmaztuk a vázolt bizonytalansági probléma kezelésére [Sánta and Kato, 2016b]. A 3D esethez hasonlóan megvizsgáltuk a TPS modell alkalmazása esetén használt kontrollpontok pozíciójának hatását az eredményre. A tesztjeinkben arra jutottunk, hogy a módszer jelentősen gyorsítható a *coarse-to-fine* hierarchikus stratégia alkalmazásával [5]. Az implementációban egy Gauss képpiramisra épülő technikát használtunk [8]. Az eljárás robusztusságát intenzitás zajjal szemben

igazoltuk. A módszer eredményeit összehasonlítottuk két korábban publikált megoldás eredményeivel (DROP [9] és SPECDEM [10]). A javasolt keretrendszert valós képeken is teszteltük, két adathalmazt használva. Az első halmazban póló mintázatokról készített képeket regisztráltunk. A második halmazban egy publikusan elérhető textúrázott papírlapok hajlítását bemutató képeket illesztettünk [11].

Ad-hoc mobil kamerahálózatok kalibrációja

Az alábbi szekcióban egy kamerahálózat póz meghatározására alkalmas algoritmust prezentálunk, amelyet az alacsony rangú textúrák rektifikációjában elért eredmények inspiráltak. A cél a hálózat lokalizációja egy valós 3D színtérben, egy 3D struktúrához képest [Sánta and Kato, 2013b]. Feltételezzük, hogy ilyen 3D struktúra nem áll rendelkezésre, ezért az algoritmusban meg kell határoznunk azt.

A javasolt eljárás három fő lépésre épül:

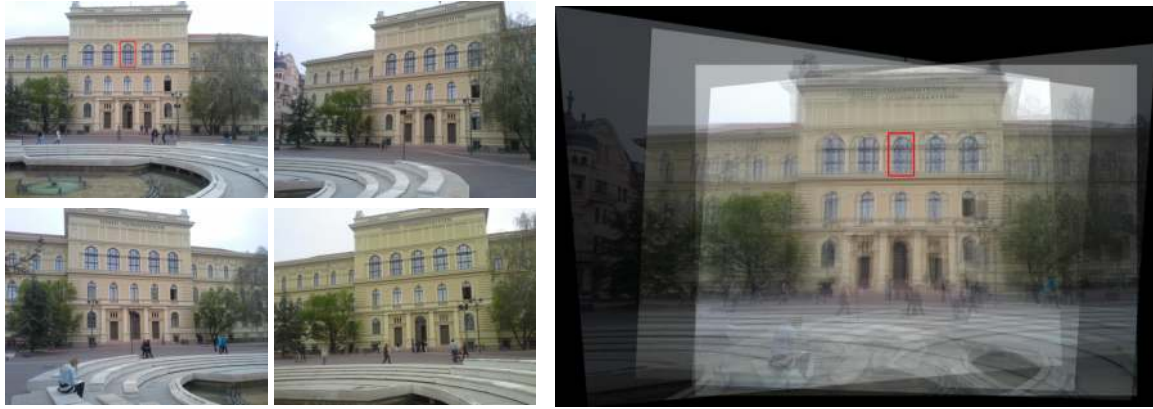
1. A relatív póz meghatározása a hálózaton belül egy tetszőleges *főkamerához* képest. Ehhez a *Normalizált 8 pontos algoritmust* használtuk, amellyel kinyertük az összes kamerapárra az esszenciális mátrixokat és így a relatív pózt.
2. Az abszolút póz meghatározása a *főkamera* és a 3D objektum között. Ezt a *Transform Invariant Low-rank Textures (TILT)* [12] algoritmus segítségével határoztuk meg. Az eljárásban egy olyan síkhomográfia kinyerése a cél, amelyet egy tetszőleges 2D foltra végrehajtva minimalizáljuk az eredmény sor- vagy oszloprangját.
3. A relatív skálafaktorok meghatározása a rendszeren belül, amely egy teljesen kalibrált hálózatot eredményez a 3D színtérben. Ehhez kifejlesztettünk egy egyszerű egy paraméterrel dolgozó regisztrációs eljárást, ami az első pontban meghatározott relatív pózra és a második pontban kinyert TILT homográfiára épül.

A kalibrációs algoritmust elosztott adatfeldolgozására terveztük: minden mobil a saját képét dolgozza föl és csak a legszükségesebb adatokat küldi tovább a többi készüléknek.

Bár az algoritmus több lépésében használunk egy kitüntetett *főkamerát*, ennek kiválasztása tetszőleges és nincs nagy befolyása az eredmény minőségére. Az eljárás robusztusságát és stabilitását egy nagy elemszámú szintetikus halmazon és valós, városi körülmények között készített képeken teszteltük.

A Szerző hozzájárulásának összefoglalása

A dolgozat eredményeit két fő téziscsoportban foglaltam össze, ahol az elsőben 3D objektumok, míg a másodikban 2D alakzatok regisztrációjával foglalkozom. A téziscsoportok és az elfogadott publikációim közötti kapcsolatot az 1. táblázatban prezentálom.



4. ábra. Eredmények a valós halmazon. Az utolsó képen a megjelenítjük a meghatározott homográfiák segítségével fuzionált kameraképeket.

I.) Háromdimenziós objektumok regisztrációja

A korábban publikált 2D keretrendszer által inspirálva [1], a módszer kiterjeszhető 3D objektumok regisztrációjára. A javasolt módszerben a regisztrációs problémát egy megfelelően konstruált nem-lineáris egyenletrendszer megoldására vezetjük vissza, ahol az ismeretlenek megadják a keresett transzformációs modell paramétereit.

- (a) Az egyenletekben szereplő integrálok általános 3D felszínek esetén felszíni integrálok lesznek. Ezek hatékony kiszámítására levezettem két rekurzív numerikus formulát (egy egzakt és egy approximációs) háromszögfelszíni hálós reprezentáció esetén. A javasolt keretrendszert nagy méretű szintetikus adathalmazon teszteltem, vékony fémlemez spline transzformációt (TPS) használva. Empirikusan igazoltam a módszer robusztusságát szegmentálási hibákkal szemben, és az eredményeket összehasonlítottam két korábban publikált módszer eredményeivel [2, 3]. Végül a keretrendszer valós alkalmazhatóságát különböző emberekről készített 3D arcfelvételek illesztésével mutattam be.
- (b) Térfogati objektumok illesztése esetén az integrálok felírhatók volumetrikus integrálokként is. Ehhez kétféle reprezentációt vizsgáltam: a voxel alapú és a zárt háromszögfelszíni hálókkal megadott objektumokat. A voxel alapú megoldásnál a [1] cikkben publikált elméleti eredmény felhasználásával levezettem egy hatékony numerikus számítási eljárást polinomiális transzformációk alkalmazásakor. A zárt felszíni hálós megoldásban az általános felszínekre felírt egzakt számítási formulát követve egy rekurzív, tetraéder alapú megoldást vezettem le TPS transzformációkra. A módszerek hatékonyságát nagy méretű szintetikus teszthalmazon igazoltam. A módszer robusztusságát szegmentálási és modell túlillesztési hibákkal szemben kísérleti úton igazoltam. A kapott

eredményeket összehasonlítottam két korábban publikált módszer eredményeivel [2, 3]. A módszer valós alkalmazhatóságának igazolására tüdő CT felvételek és agyi felszínek regisztrációja került bemutatásra.

II.) Kétdimenziós alakzatok regisztrációja

A téziscsoport összefoglalja az eredményeim bináris és többszintű képek regisztrációjában. Első témában kitalált és hiányos bináris alakzatok regisztrációjával foglalkoztam, ami egy gyakori probléma valós környezetben készített képek illesztésénél. Ezt követően a bináris alakzatok nem-lineáris regisztrációja során tapasztalható alulhatározottság problémájára javasoltam egy megoldást. Végül kidolgoztam egy ad-hoc kamerahálózatok kalibrációjának meghatározására alkalmas algoritmust.

- (a) A [4, 6] cikkekben bemutatott, bináris alakzatok affin illesztésére alkalmas keretrendszer adaptálható kitalált alakzatok regisztrációjára, az integrálási tartomány alkalmas megválasztásával. A javasolt algoritmusban az alakzatokat poligonokként reprezentálva, iterációnként meghatároztam a kitalált területeket. Majd ezen területeket integrálási tartományként használva meghatároztam a legjobb affin transzformációt az alakzatok között. A módszer hatékonyságát nagy méretű szintetikus adathalmazon igazoltam. Az eredményeket összehasonlítottam két korábban publikált módszer eredményeivel [2, 7]. A javasolt módszer valós alkalmazhatóságát városi körülmények között készített statikus és dinamikus takarásokat tartalmazó képeken mutattam be.
- (b) A [1] cikkben bemutatott módszer tovább regularizálható szürkeárnyalatos képek alkalmazásával. Ehhez kidolgoztam egy olyan formalizmust, amelyben a geometriai és intenzitás információk egy közös egyenletrendszerben jelennek meg. A módszer hatékonyságát empirikus úton igazoltam nagy elemszámú szintetikus generált halmazon. Kísérleti úton megmutattam az eljárás robusztusságát különböző szórással generált normális eloszlású intenzitás zaj mellett. Összehasonlítottam a kapott eredményeket két korábban publikált módszer eredményeivel [9, 10]. Végül valós körülmények között készített képekkel és egy publikusan elérhető tesztadatbázison igazoltam a módszer alkalmazhatóságát.
- (c) Az alacsony rangú textúrákkal (TILT) [12] kapcsolatos eredmények felhasználhatók kamerahálózatok kalibrációjának meghatározására. Kidolgoztam egy algoritmust, amely TILT jellemzőkből kinyert 2D homográfiák segítségével meghatározza egy kamerahálózat és az egyes kamerák abszolút helyzetét a világ koordináta-rendszerben. A javasolt algoritmus hatékonyságát és stabilitását nagy elemszámú szintetikus halmazon és mobil kamerákkal készített

valós képekkel is igazoltam.

	I		II		
	a	b	a	b	c
[Sánta and Kato, 2012a]		•			
[Sánta and Kato, 2012b]		•			
[Sánta and Kato, 2013a]		•			
[Sánta and Kato, 2016a]	•				
[Sánta and Kato, 2018]	•	•			
[Sánta and Kato, 2014]			•		
[Sánta and Kato, 2016b]				•	
[Sánta and Kato, 2013b]					•

1. táblázat. A tézispontokhoz kapcsolódó publikációk.

Publikációs lista

Folyóiratcikkek

[Sánta and Kato, 2018] Z. Sánta and Z. Kato, “Elastic Alignment of Triangular Surface Meshes”, *International Journal on Computer Vision*, 2018, Accepted.

Konferencia-kiadványokban megjelenő cikkek

[Sánta and Kato, 2016a] Z. Sánta and Z. Kato, “3D Face Alignment without Correspondences”, in *Proceedings of the ECCV Workshop on 3D Face Alignment in the Wild*, Amsterdam, Netherlands, Oct. 2016, pp. 521–535. DOI: 10.1007/978-3-319-48881-3_36.

[Sánta and Kato, 2016b] Z. Sánta and Z. Kato, “An Algebraic Framework for Deformable Image Registration”, in *Proceedings of International Conference on Pattern Recognition*, Cancun, Mexico, Dec. 2016, pp. 3792–3797. DOI: 10.1109/ICPR.2016.7900225.

[Tanács et al., 2015] A. Tanács, A. Majdik, L. Hajder, J. Molnár, Z. Sánta, and Z. Kato, “Collaborative Mobile 3D Reconstruction of Urban Scenes”, in *Proceedings of the ACCV Workshop on Intelligent Mobile and Egocentric Vision*, C. V. Jawahar and S. Shan, Eds., ser. Lecture Notes in Computer Science, vol. 9010, Singapore: Springer International Publishing, 2015, pp. 486–501, ISBN: 978-3-319-16633-9. DOI: 10.1007/978-3-319-16634-6_36.

[Sánta and Kato, 2014] Z. Sánta and Z. Kato, “Affine Alignment of Occluded Shapes”, in *Proceedings of International Conference on Pattern Recognition*, IAPR, Stockholm, Sweden: IEEE, Aug. 2014, pp. 2155–2160. DOI: 10.1109/ICPR.2014.375.

- [Sánta and Kato, 2013a] Z. Sánta and Z. Kato, “Correspondence-Less Non-Rigid Registration of Triangular Surface Meshes”, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Portland, Oregon: IEEE, Jun. 2013, pp. 2275–2282. DOI: 10.1109/CVPR.2013.295.
- [Sánta and Kato, 2013b] Z. Sánta and Z. Kato, “Pose Estimation of Ad-hoc Mobile Camera Networks”, in *Proceedings of International Conference on Digital Image Computing: Techniques and Applications*, Hobart, Tasmania, Australia: IEEE, Nov. 2013, pp. 88–95. DOI: 10.1109/DICTA.2013.6691514.
- [Sánta and Kato, 2012a] Z. Sánta and Z. Kato, “A Unifying Framework for Non-linear Registration of 3D Objects”, in *Proceedings of IEEE International Conference on Cognitive Infocommunications*, Kosice, Slovakia: IEEE, Dec. 2012, pp. 547–552. DOI: 10.1109/CogInfoCom.2012.6422041.
- [Sánta and Kato, 2012b] Z. Sánta and Z. Kato, “Elastic Registration of 3D Deformable Objects”, in *Proceedings of International Conference on Digital Image Computing: Techniques and Applications*, Fremantle, Western Australia: IEEE, Dec. 2012. DOI: 10.1109/DICTA.2012.6411674.

Referenciák

- [1] C. Domokos, J. Nemeth, and Z. Kato, “Nonlinear Shape Registration without Correspondences”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 943–958, 2012. DOI: 10.1109/TPAMI.2011.200.
- [2] A. Myronenko and X. Song, “Point Set Registration: Coherent Point Drift”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2010.46.
- [3] B. Jian and B. Vemuri, “Robust Point Set Registration Using Gaussian Mixture Models”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2010.223.
- [4] A. Tanács, J. Lindblad, N. Sladoje, and Z. Kato, “Estimation of linear deformations of 2D and 3D fuzzy objects”, *Pattern Recognition*, vol. 48, no. 4, pp. 1391–1403, 2015, ISSN: 0031-3203. DOI: 10.1016/j.patcog.2014.10.006.
- [5] B. Zitová and J. Flusser, “Image registration methods: A survey”, *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003, ISSN: 0262-8856. DOI: 10.1016/S0262-8856(03)00137-9.
- [6] C. Domokos and Z. Kato, “Parametric Estimation of Affine Deformations of Planar Shapes”, *Pattern Recognition*, vol. 43, no. 3, pp. 569–578, 2010. DOI: 10.1016/j.patcog.2009.08.013.

- [7] J. Zhu *et al.*, “Robust affine iterative closest point algorithm with bidirectional distance”, *Computer Vision, IET*, vol. 6, no. 3, pp. 252–261, 2012, ISSN: 1751-9632. DOI: 10.1049/iet-cvi.2011.0178.
- [8] P. Burt and E. Adelson, “The Laplacian Pyramid as a Compact Image Code”, *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, Apr. 1983, ISSN: 0090-6778. DOI: 10.1109/TCOM.1983.1095851.
- [9] B. Glocker, A. Sotiras, N. Komodakis, and N. Paragios, “Deformable Medical Image Registration: Setting the State of the Art with Discrete Methods”, *Annual Review of Biomedical Engineering*, vol. 13, no. 1, pp. 219–244, 2011. DOI: 10.1146/annurev-bioeng-071910-124649.
- [10] H. Lombaert *et al.*, “Spectral Log-Demons: Diffeomorphic Image Registration with Very Large Deformations”, *International Journal on Computer Vision*, vol. 107, no. 3, pp. 254–271, 2014. DOI: 10.1007/s11263-013-0681-5.
- [11] M. Salzmann, R. Hartley, and P. Fua, “Convex Optimization for Deformable Surface 3-D Tracking”, in *Proceedings of International Conference on Computer Vision*, Oct. 2007, pp. 1–8. DOI: 10.1109/ICCV.2007.4409031.
- [12] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma, “TILT: Transform Invariant Low-Rank Textures”, *International Journal on Computer Vision*, vol. 99, no. 1, pp. 1–24, 2012, ISSN: 0920-5691. DOI: 10.1007/s11263-012-0515-x.