

# Felületi normális becslése egyetlen pontmegfeleltetés alapján

Eichhardt Iván<sup>1,2</sup>, Baráth Dániel<sup>1,2</sup>

<sup>1</sup> MTA SZTAKI, Budapest

{eichhardt.ivan,barath.daniel}@sztaki.mta.hu

<sup>2</sup> Eötvös Loránd Tudományegyetem, Budapest

**Kivonat** Napjaink többnézetű 3D rekonstrukciós eljárásai remek eredményeket tudnak felmutatni tipikusan akkor, ha sok nézet áll rendelkezésükre. Munkánkkal elsődleges célunk, hogy robusztus módon több információt tudjunk kinyerni képekből, jelentősen kevesebb nézet felhasználásával. Cikkünkben egy eljárást mutatunk be, amely pontokban képes felületi normálist, illetve érintősíkot becsülni stereo esetben, valós jeleket 3D rekonstrukciója céljából. Az általunk bemutatott algoritmus tetszőlegesen választott kamera esetén is működik, ám a demonstráció céljából a lyukkamera-modellt választottuk egyszerűsége és hatékonysága végett. Munkánk a particle swarm optimization (PSO) eljárást használja fel geometriai és epipoláris megkötések mellett, a megfelelő gyorsaság és minőség biztosítására. Kimenete egy normálisokkal ellátott pontfelhő, amelyet kizárólag két nézet és azok pont-pont megfeleltetéseiből állít elő egy foto-konzisztencia alapú költségfüggvény felhasználásával. Eljárásunk könnyen kiterjeszthető a több nézetre is. Az algoritmust szintetikus és valós esetben is validáljuk, valamint összehasonlítjuk egy korszerű, többnézetű képfolt-alapú rekonstrukciós eljárással.

## 1. BEVEZETÉS

A felületi normális és a kapcsolódó sík-régió (patch) becslése évtizedek óta intenzíven kutatott területe a számítógépes látásnak. Cikkünkben bemutatunk egy eljárást és leírjuk a mögöttes elméletet, mellyel sikszerű térbeli felületdarabok nagy pontosságú becslését végezhetjük el 2D pontmegfeleltetések között, stereo esetben. Szemléltetjük, hogy javasolt eljárásunkkal sok esetben pontosabb eredményt kaphatunk, mint más, korszerű becselő által. Tapasztalataink szerint a legtöbb ritka vagy sűrű rekonstrukciót biztosító algoritmus a megfigyelt pontoknak jó minőségű pozíciót becsül, de az odatartozó régiók orientációjának (felületi normális) csak durva becslést ad. Jelen munkánkat ez adott motivációt.

---

Baráth, Dániel és Eichhardt, Iván (2016) A Novel Technique for Point-wise Surface Normal Estimation. In: VISIGRAPP 2016. Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. Vol. 3. VISAPP 2016. SciTePress, Setúbal, pp. 688-695.

Az algoritmus feltételezi, hogy 2D pont-megfeleltetések már adottak (2D pontpárok) egy stereo-pár két képe között. A stereo-nézet kalibrációja szintén az eljárás bemenete. Feltesszük továbbá, hogy a bemeneti pontpárok ugyan azon felületdarabhoz tartoznak. Ezen pont-megfeleltetések alapján végzett triangulációk után az algoritmusunk további információt szolgáltat majd a felületről: a triangulált 3D pontban vett felületi normálist. A folyamat kimenete tehát egy orientált ponthalmaz, a megfigyelt szintér egy ritka rekonstrukciója.

### 1.1. Kapcsolódó munkák

Az alkalmazott objektum-modell alapján négy fő kategóriába sorolhatjuk a stereo rekonstrukciós eljárásokat: voxel-alapú [5,18], deformálható poligonos [23], mélységkép-fúziós [20] és folt-alapú [7,9]. Mivel a mi megközelítésünk folt-alapú (patch-based), rövid áttekintésünkben csak ezt a témát vizsgáljuk.

Minden egyes folt a felület lokális darabkájából, illetve az ott található érintősík leírásából tevődik össze. Az érintősík becslése *közvetlen* és *közvetett* módokon végezhető. A költségfüggvény közvetlen paraméterezése esetén annak minimumhelyén találjuk az optimális érintősíkot. Közvetett esetben először megbecsüljük a relatív (affin) homográfiát a 3D folt vetületei között, majd az érintősík paraméterei e transzformáció felbontása által [6] nyerhető ki. Az alkalmazott kamera-nézetek száma és a rekonstrukció milyensége (ritka vagy sűrű) változhat a különböző közvetlen vagy közvetett eljárások esetén.

Egy **közvetett** módon működő eljárás Megyesi et al. [15] munkája, mely rektifikált képekből két lépésben végzi a rekonstrukciót. Először affin folt-illesztéssel magpontokat generál, majd azokból iteratív módon egy sor feltételnek megfelelően propagálja a becsült felületet. A magpontok létesítéséhez kimerítő keresés (exhaustive search, ES) alapú eljárásuk optimális 3 szabadsági fokú (degree of freedom, DoF) affin transzformációt választ ki a rektifikált nézetek kép-foltjai között, foto-konzisztencia alapján. A becsült transzformációk felbonthatóak diszparitássá és felületi normálissá. A szerzők több epipoláris geometria alapú megkötést alkalmaznak a keresési tér szűkítése céljából. Egy láthatóság-alapú megkötésük alapján a nem a kamera képsíkja felé mutató, vagy az optikai tengelyre közel merőleges felületi normálisokat eldobhatóak.

Több kép-alapú eljárás affin transzformációból [1] vagy homográfia felbontása [6] által számol felületi normálist. Molnár et al. [16] munkája továbbá megmutatta, hogy a felületi normális stereo nézetben közvetlenül kifejezhető affin homográfiából, felhasználva a 2D vetítőfüggvények térkoordináták szerinti gradienseit. Ezen eljárásokhoz szükséges a kamerák teljes kalibrációja, továbbá összetartozó képpontok környezetükhöz köthető (affin) homográfia-transzformáció.

Néhány eljárás [8,7,13,21] az érintősíkokat becsülő **közvetlen** módon működő eljárások közé sorolható. Habbecke et al. [8] a problémát egy 3D sík-keresési problémaként fogalmazza meg. Ez a megközelítés munkánkhoz rokonítható, azzal a különbséggel, hogy ők megfeleltetett 2D blob-ok alapján dolgoznak, a síkot pedig 3 paraméterrel írják le. Algoritmusuk foto-konzisztencia alapú Gauss-Newton optimalizáció szerint működik.

A több nézetből sűrű rekonstrukciót előállító eljárások területéről meg kell említenünk a PMVS [7]-t. Megközelítésük folt-alapú és 3-lépésből – *illesztés, kiterjesztés és szűrés* – áll, ahol az utolsó két lépést iteratívan alkalmazzák, előállítva egy folyamatosan finomodó orientált pontfelhőt, vagyis patch-ek halmazát. Az illesztés rész egy fotometrikus eltérést mérő függvényt igyekszik minimalizálni, ezáltal térben és kép-térben is optimalizálnak. A kezdeti illesztés után a patch-ek finomítására egy gradiens-eljárást alkalmaznak, a jobb orientációk reményében. A szűrés lépés során a szerzők láthatóság alapú feltételek mellett eliminálják a helytelen illesztéseket és az outliereket. Ezek a megkötések rokoníthatóak a Megyesi et al. [15]-ben alkalmazottakhoz. Összehasonlításként mi eljárásunkban található hasonló láthatósági megkötések *nem egy utófeldolgozási* lépésben fejtik ki hatásukat, hanem közvetlenül a keresési térre alkalmazhatóak. A kiterjesztés lépése szintén hasonló Megyesi et al. [15] leírásához. A PMVS utolsó lépéseként a Poisson Surface Reconstruction [11] és iteratív illesztések sorozataként sűrű háromszögelt rekonstrukciót (felületet) állít elő. Az illesztések során előtér/háttér szegmentációt igyekeznek érvényesíteni energia-optimalizáció által. Eljárásuk gyengesége a felületi normálisok inicializációjából származik, mely preferálja a nézet felé orientálódott foltokat. Az ettől való eltérés a gradiens-eljárás egyre rosszabb kimenetét eredményezi.

Néhány többlépéses, sok nézetekből dolgozó rendszerről [13,21] is szót kell ejtenünk: ezek célja a jó minőségű sűrű rekonstrukció. Első lépésük kritikus: egy kezdeti ritka, vagy kvázi-sűrű rekonstrukciót készítenek.

A Particle Swarm Optimization (PSO) [12,19] egy populáció-alapú algoritmus, arra kifejlesztve, hogy „hasznos” megoldásokat találjon egy folytonos problémára, határolt (vagy periodikus) keresési térben. Iteratív működése során a megoldásra újra és újra jelöltet vagy jelölteket állít, melyet egy raj (swarm) mozgása alapján igyekszik finomítani. Több részecske kooperatív módon egyszerre keresi a megoldásokat. PSO mellőzi a deriváltak használatát és zajos bemenettel is megbirkózik.

## 1.2. Motivációnk és céljaink

Képek közötti (stereo vagy többnézetű) összefüggések keresése nem egyértelmű, a mögöttes felület rekonstrukciója pedig inkorrekt kitűzésű (ill-posed) probléma. Számos eljárás ezt nagyszámú nézet alkalmazásával próbálja feloldani, jobban megszorítva a problémát.

Cikkünkben egy közvetlen módon működő eljárást fogalmazunk meg az affin folt-illesztés feladatára, oly módon, hogy azt egy 2 szabadsági fokú keresési térben végezzük. Ez a keresési tér azonos a folt felületi normálisának paraméterterével.

A Gauss-Newton eljárásokkal szemben a PSO-t preferáltuk, mivel az deriváltak nélkül igyekszik megtalálni a globális optimumot. Habár a PSO esetén – általában – nem bizonyított a konvergencia, esetünkben a becslt érintő sík minősége meghaladta a nagyobb számítási igényű kimerítő keresés által nyújtottat. Az epipoláris geometriát alkalmazva közvetlen megkötéseket fogalmaztunk meg

a keresési térre, ezáltal algoritmusunk egy újszerű utat nyit kalibrálatlan képi nézetek nagy minőségű rekonstrukciójára.

A cikkünk és az algoritmusunk nem foglalkozik a teljes, sűrű rekonstrukcióval. Célja a különálló foltok érintősíkjait minél nagyobb minőségben megbecsülni, remek alapot adva egy jövőbeli többnézetű rekonstrukciós algoritmusnak. Lehetséges alkalmazásai a 3D rekonstrukció terén: magpontok generálása felület-propagációhoz, további nézetek csatolása nagy pontossággal.

Habár a hasonló feladatok esetén a legtöbbit használt deformáló (warp) függvény a homográfia, mi az affin transzformációt választottuk. Megközelítésünk előnye, hogy *tetszőleges* kameramodell esetén alkalmazható, mivel azt a kamera vetítőfüggvényének differenciálása [16] által hozzuk létre:

1. A bemeneti fotókról a kameratorzítás eltüntetése szükségtelen.
2. A képfoltok deformálása során annak minden pixelére véve a kameratorzítás kiértékelése szükségtelen. Az affin deformáló mátrix még a pixelek mintavételezése előtt kiértékelhető, így egy képfolt torzítása egy olcsó affin deformáció.
3. Bár itt nem tárgyaljuk, de kézenfekvő és egyszerű ennek kiterjesztése tetszőleges centrális kamerára (pl.: omnidirekcionális kamerák).

Bár algoritmusunk mutat hasonlóságokat egy másik eljárás [15] magpontokat előállító lépésére, az kimerítő keresés alapján működik, amit mi megfogalmazásunkban a PSO-ra tudtuk helyettesíteni. A bemeneti képek rektifikációja a fent taglaltak szerint szükségtelen. Ismereteink szerint a PSO rekonstrukciós problémákhoz való alkalmazásai [3] között a cikk témáját illetőleg újdonság. Továbbá újszerű láthatóság alapú megkötéseink nagymértékben csökkentik a PSO keresési terét.

Kísérleteinket szintetikus és valós adatokon végeztük, eredményeinket pedig összehasonlítottuk egy korszerű eljárással [7], stereo képpár esetén alkalmazva, ahogyan az a 4.1. részben is látható. Az eljárásunk által szolgáltatott orientált pontfelhő Poisson felületrekonstrukciója [11] minden további utófeldolgozás vagy felület-propagáció [15] nélkül is kielégítő minőséget biztosít.

Cikkünk fennmaradó része a következőkből áll: a 2. szakaszban leírjuk a szükséges jelöléseket és a geometriai háttérrel, a 3. szakaszban az algoritmus lényegi részéről olvashatunk, a geometriai reprezentációról és a keresési téren alkalmazott újszerű megkötésekről. Végezetül a 4. részben bemutatjuk, hogy algoritmusunk jól működik szintetikus és valós körülmények között, munkánkat pedig a 5 pontban foglaljuk össze.

## 2. Jelölések és geometriai háttér

A cikkben a mátrixokat félkövére nagybetűkkel jelöljük, mint például  $\mathbf{K}$ , vagy  $\mathbf{R}$ . Vektorokat aláhúzással emeljük ki, pl.:  $\underline{T}$ ,  $\underline{p}$ ,  $\underline{q}$ ,  $\underline{x}$ . Továbbá a  $\underline{p}$  legtöbbször egy 3D térbeli pontot,  $\underline{p}$  2D, képen lévő (vetületi) pontot jelöl. Utóbbi homogén koordinátás alakja:  $\underline{\tilde{p}}$ .

A perspektív kamera projekciós mátrixa a következő  $\mathbf{P} = \mathbf{K} \cdot [\mathbf{R} \mid \underline{T}] \in \mathbb{R}^{3 \times 4}$ , ahol  $\mathbf{K} \in \mathbb{R}^{3 \times 3}$  a belső kamera-paraméterek mátrixa,  $\mathbf{R} \in \mathbb{R}^{3 \times 3}$  a forgatási mátrix és  $\underline{T}$  az eltolás vektora.



Egy  $\underline{x} \in \mathbb{R}^3$  irányvektor gömbi koordinátás alakja:

$$\text{Sph}(\underline{x}) = \text{Sph}([X, Y, Z]^T) = \left[ \arccos\left(\frac{Z}{\|\underline{x}\|}\right) \arctan_2(Y, X) \right]^T.$$

Legyen  $\underline{a}, \underline{b} \in \mathbb{R}^2$  úgy, hogy  $\forall i \in \{1, 2\} : a_i \leq b_i$ . Ekkor  $\text{rect}[\underline{a}, \underline{b}]$  az a minimális befoglaló téglalap, mely tartalmazza az  $\underline{a}$  és  $\underline{b}$  vektorokat.

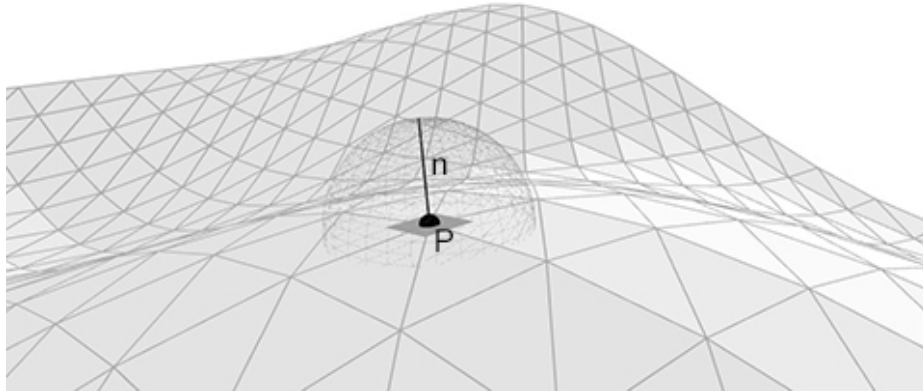
### 3. Iteratív érintősík-becslés

Ebben a részben bemutatjuk, hogy a felületi normális robusztus módon becsülhető egyetlen pontmegfeleltetésből és e pontpárok lokális képi környezetéből. Javasolt eljárásunkat a következőkben Iterative Tangent Plane Estimation-nak, röviden **ITPE**-nek nevezzük.

#### 3.1. Az algoritmus alapjai

Eljárásunk alapötlete igen egyszerű: ha adott a vizsgált felület egy pontja, a pontban vett érintősík leírható annak normálisával (lásd: 1. ábra). Ez pontosan azt jelenti, hogy a sík fennmaradó szabadsági foka 2, a felületi normális és így a sík is paraméterezhető  $(u, v)$  gömbi koordináták segítségével (elhagyva a radiális komponensét).

Jelöljük a stereo rendszer kameráinak projekciós mátrixát  $\mathbf{P}_1$ -el és  $\mathbf{P}_2$ -vel, az  $i$ -ik 3D pontot  $\tilde{P}^i$ -vel, homogén alakban. A pont vetületei a képsíkokon:  $p_1^i$  és  $p_2^i$ . Ezek alapján a vizsgált pontban vett érintősík leírható a pont és az  $\underline{n}$  normális használatával. A problémát így kétdimenziós, a gömbi koordináták szerint megszorított periodikus problématerű optimalizációra redukáltuk, ahol  $u$  és  $v$  ismeretlenek – a megoldás létezése pedig garantált.



1. ábra. A felület  $P$  pontja, az abban vett  $n$  normális és az érintősík.

Az  $\underline{n}$ -el felparaméterezett érintősík minőségét súlyozott zéró-átlagú normalizált keresztkorreláció (ZNCC) [14] alapján mérjük. Az alkalmazott ablak pixeljei Gauss-eloszlás szerint vesznek részt a korreláció kiszámításában, így a belső pixellek  $\underline{P}^i$  közelében nagyobb hangsúlyt kapnak. Két különböző nézetű folt hasonlóságának kiértékelése előtt azokat közös koordinátarendszerbe hozzuk. Habár erre a leggyakrabban alkalmazott technika a forráspixelek homográfiával való transzformálása, mi inkább a  $\underline{p}_1^i, \underline{p}_2^i$  pontokban deriváltak alapján kiszámítható ((2) egyenlet) affin közelítést használunk. Tapasztalataink alapján valós körülmények között használatával pontosabb eredményeket kapunk (lásd a 1. táblázatban) mint homográfiával, illetve affin közelítésünk alkalmazható tetszőleges kameramodell esetén. Az algoritmus az optimalizáció során a következő kifejezés maximumát keresi:

$$A_1(u, v) = \text{ZNCC} \left[ \tau_{\underline{p}_1} \circ \mathbf{A}(u, v), \tau_{\underline{p}_2} \right]. \quad (1)$$

A  $\tau_{\underline{p}_j}$  kifejezés a  $\underline{p}_j$  pont lokális pixel-környezetét (subpixelesen mintavételezve) jelöli, illetve a  $\circ$  operátor az  $\mathbf{A}$  affin transzformáció alkalmazását jelöli a mintavételezés során. Ezzel az  $\mathbf{A}$  transzformációval deformáljuk a  $\tau_{\underline{p}_j}$  foltot.  $\mathbf{A}$  egy függvény, ahogyan azt a (2) egyenletben láthatjuk, melyet Molnár et. al. [16] formulációja alapján adtunk meg. Ezáltal megkapjuk az  $(u, v)$ -vel paraméterezett felületi normálishoz tartozó érintősík két vetülete közötti relatív kép-transzformációjának elsőrendű közelítését:

$$\mathbf{A}(u, v) = \frac{1}{|\nabla x_1 \underline{n} \nabla y_1|} \begin{bmatrix} |\underline{n} \nabla y_1 \nabla x_2| & |\underline{n} \nabla x_2 \nabla x_1| \\ |\underline{n} \nabla y_1 \nabla y_2| & |\underline{n} \nabla y_2 \nabla x_1| \end{bmatrix}, \quad (2)$$

ahol a következő derivált mennyiségeket használjuk:

$$\begin{aligned} \nabla x_i &= \frac{1}{s_i} \left( \mathbf{P}_i|_{(1,1:3)} - x_i \cdot \mathbf{P}_i|_{(3,1:3)} \right) \\ \nabla y_i &= \frac{1}{s_i} \left( \mathbf{P}_i|_{(2,1:3)} - y_i \cdot \mathbf{P}_i|_{(3,1:3)} \right) \\ s_i &= \mathbf{P}_i|_{(3,1:4)} \cdot \begin{bmatrix} P \\ 1 \end{bmatrix} \end{aligned} \quad (3)$$

továbbá  $|\underline{a} \underline{b} \underline{c}|$  a három  $\underline{a}, \underline{b}, \underline{c} \in \mathbb{R}^3$  vektor vegyes szorzata,  $\mathbf{P}_i|_{(k,l:m)}$  pedig a  $\mathbf{P}_i$  mátrix  $k$ -ik sorából és  $l$ -től  $m$ -ig terjedő oszlopaiból álló almátrix. Az  $\underline{n}$  felületi normális  $u, v$  szerint a következőképpen paraméterezett:

$$\underline{n} = \begin{bmatrix} \cos(u) \sin(v) \\ \sin(u) \sin(v) \\ \cos(v) \end{bmatrix}^T. \quad (4)$$

A feladat első megközelítése szerint a (1) kifejezés maximumát kell megtalálni, ami az első kép transzformált pixeljeinek értékeit hasonlítja össze a második képével, mindezt a felületi normális szerint paraméterezve. A probléma egy egyszerű keresésnek tűnik, de az affin transzformációban rejlő többértelműségek és

a foto-konzisztenciái hiányosságok miatt több megkötést kell adnunk a rendszerhez a nagyobb robusztusság érdekében. Továbbá, a stereo probléma szimmetrikus természete miatt a fenti egyenlet kiegészíthető a második foltra vett inverz affin transzformáció szerint:

$$A_2(u, v) = \text{ZNCC} \left[ \tau_{p_1}, \tau_{p_2} \circ \mathbf{A}(u, v)^{-1} \right]. \quad (5)$$

Végezetül egyesítve a (1) és (5) egyenleteket, a legjobban illeszkedő felületi normálist a következő alapján keressük:

$$J'' = \arg \max_{(u, v)} (A_1(u, v) A_2(u, v)). \quad (6)$$

Tapasztalataink szerint a keresztkorrelációhoz adott Gauss-súlyok javítják a rekonstrukció minőségét.

### 3.2. A keresési tér megkötései

Kellő minőségű rekonstrukció elérése érdekében egy sor új megszorításra van szükségünk az optimalizáció során.

**Láthatósági megkötés** Több geometriai megkötést tehetünk a PSO működésére, mivel az optimális érintősík 2 szabadsági fokú keresési térben található, ami a felületi normális  $(u, v)$  gömbi koordinátás reprezentációja. A következő pontokban bemutatunk néhány – többnézetű – láthatóság alapú, a paraméterterre alkalmazható megkötést.

1. Első megkötésünk a gömbi koordináták periodikus mivoltát veszi figyelembe:

$$\begin{aligned} u &\in \left[ u^0 - \frac{\pi}{2}, u^0 + \frac{\pi}{2} \right] \\ v &\in \left[ v^0 - \pi, v^0 + \pi \right] \end{aligned} \quad (7)$$

ahol  $(u^0, v^0) = \text{Sph}(\underline{w})$  és a  $\underline{w} := -\frac{\mathbf{K}^{-1}\tilde{p}}{\|\mathbf{K}^{-1}\tilde{p}\|_2}$  vektor a nézőpont felé mutató vektor. Általánosan megfogalmazva a koordinátarendszer középpontja:

$$(u_i^0, v_i^0) = \text{Sph}(\underline{w}_i) = \text{Sph} \left( -\mathbf{R}_i^T \mathbf{K}_i^{-1} \tilde{p}_i \right). \quad (8)$$

2. Érvénytelen egy olyan érintősík, amely nem a nézőpont felé néz. A keresési tér ezért megfelelezhető:

$$v \in \left[ v_0 - \frac{\pi}{2}, v_0 + \frac{\pi}{2} \right]. \quad (9)$$

3. Ez a terület felfogható egy olyan, a gömbi koordinátarendszerben lévő *tég-lalapként* mely  $\underline{a}^0 = (u^0 - \frac{\pi}{2}, v^0 - \frac{\pi}{2})$ -t és  $\underline{b}^0 = (u^0 + \frac{\pi}{2}, v^0 + \frac{\pi}{2})$ -t foglalja

magába. Jelöljük ezeket a téglalapokat  $\text{rect}_i := \text{rect} [\underline{a}^i, \underline{b}^i]$ -vel az  $i$ -ik nézetből vizsgálva. Feltételezve, hogy minden érintősíkot több nézetből vizsgáljuk, a minden  $i$ -re vett  $\text{rect}_i$ -k metszete megszorítja a lehetséges normálisok keresési terét:

$$\text{rect} \left[ \min_{u,v}, \max_{u,v} \right] := \bigcap_i (\text{rect}_i). \quad (10)$$

A keresési tér ilyen módon legfeljebb egy  $\frac{\pi}{2}$  oldalhosszúságú minimális négyzetre csökkenthető. A következő pont leírja ezt a jelenséget, és egyúttal megfogalmazzuk az utolsó geometriai megkötést.

4. Ha létezik egy olyan  $j$  nézőpont, melyre  $(u_j^0, v_j^0) \notin \text{rect} [\min_{u,v}, \max_{u,v}]$ , akkor a keresési tér is *üres*:

$$\begin{aligned} \exists j : (u_j^0, v_j^0) \notin \text{rect} \left[ \min_{u,v}, \max_{u,v} \right] &\iff \\ \exists j \exists i : (u_j^0, v_j^0) \notin \text{rect}_i &\iff \\ \exists j \exists i : \langle \underline{w}_j, \underline{w}_i \rangle < 0. & \quad (11) \end{aligned}$$

Ez azt jelenti, hogy a két irány közötti szög nagyobb, mint  $\frac{\pi}{2}$ , melyből az következhet, hogy egy vagy több nézőpont outlierok közé sorolható.

5. Érdeemes némileg kiterjeszteni minden  $\text{rect}_i$  határait a lehetséges kalibrációs hibák miatt.

**Az affin transzformációra tett megkötés** Nyilvánvaló, hogy a vetítés során a tükrözés nem lehetséges, tehát kikötjük, hogy  $\det(\mathbf{A}^i) > 0$ . Az affin transzformációt komponenseire bontva további extrém eseteket szűrhetünk ki, mint: túlzottan alacsony skála és túlzottan nagy nyírás. Tapasztalati alapon a skála alsó-, illetve a nyírás felső határának rendre 0.2-t és 2.0-t választottunk.

### 3.3. Globális optimalizáció

Megfogalmazásunkban a költségfüggvény optimumát a geometriai kényszerek mellett ((10) egyenlet) globálisan kell megtalálnunk. Mivel néhány affin transzformációt nem veszünk figyelembe (lásd: 3.2. rész), a függvény nem folytonos és igen zajos képi interpolációs hiányosságok és diszkrét képi tulajdonságok miatt. Ilyenkor kézenfekvő a teljes keresési téren kimerítő keresést végezni, ám ez a kellő minőségért cserében igen lassú.

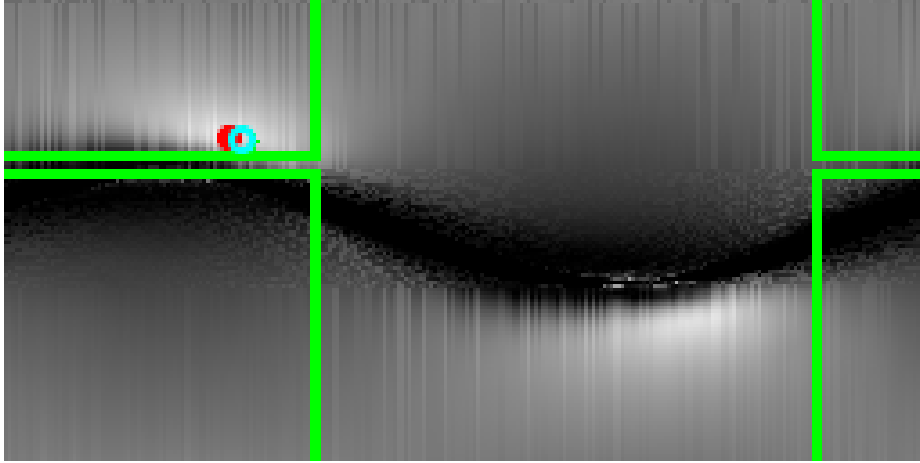
Megközelítésünk szerint a PSO-t alkalmaztuk, az iteratív algoritmust addig futtatva, amíg a kiválasztott optimum értékének abszolút változása  $k$  lépés során  $\theta$  alá nem csökken. Kísérleteinkben ezeknek az értékeknek rendre 5-öt és  $10^{-9}$ -et választottunk. A rajt inicializálás során egyenletesen osztottuk el a keresési térben. Közöttük véletlenszerűen kommunikációs kapcsolatokat hoztunk létre. Ez a beállítás gyorsnak és kellő pontosságúnak bizonyult. A minősége vetekszik egy igazán részletes kimerítő keresésével (exhaustive search– ES), ám annál sokkal gyorsabb (lásd: 2. táblázat).

A részecskék maximum számát  $n_{\max} = 100$ -nak határoztuk meg, de ezt a számot adaptívan csökkentettük  $n_{opt}$ -ra, ahogy a keresési tér mérete (3.2 rész) is csökkent.

$$n_{opt} := \min \left( 4, \left\lceil 100 \frac{|\text{rect}[\min_{u,v}, \max_{u,v}]|}{\pi^2} \right\rceil \right) \quad (12)$$

ahol  $|\text{rect}|$  a keresési tér mérete. Ez a lépésünk egy enyhe csökkenést eredményezett futásidőben, precizitás-csökkenés nélkül.

A 2. ábra mutatja a periodikus paraméterteret és a mintavételezett költségfüggvényt, a vízszintes  $u$  és függőleges  $v$  tengelyekkel. (A bal-felső sarok a  $[0, 0]^T$ , a jobb alsó sarok pedig a  $[2\pi, \pi]^T$  pontot jelenti.). A zöld téglalap a lecsökkentett keresési teret jelöli, hála a láthatósági megkötéseknek. Vörös és kék pontok jelölik a kimerítő keresés, illetve a PSO által talált optimumokat. Ahogyan azt vártuk is, a zöld keresési tér tartalmazza az optimumot. A kapott koordinátapár (kék pont) közel található a kívánt megoldáshoz. A kép közepén látható érvénytelen (fekete) rész, az affin transzformációra tett megkötéseket jelenti. Ezáltal számos értéket egyszerűen elhagytunk, 0 függvényértéket választva annak. Látható hogy a kép-foltok hasonlóságát mérő függvény nem konvex a két látható csúc és a fekete rész miatt. Megjegyezzük, hogy előzetes tesztheink alapján az érvénytelen fekete részek is tartalmaznak magas értékű csúcsokat, ám azok érvénytelen normálisok lennének. Sajnos a megszorított keresési térben sem szolgálna jó megoldással egy gradiens-módszer, mivel ott sem konvex a keresési tér.



2. ábra. A hasonlóságot leíró függvény egy kísérletünk során, valós adaton, a vízszintes  $u$  és függőleges  $v$  tengelyekkel. A zöld téglalap a lecsökkentett keresési teret jelöli, hála láthatósági megkötéseknek. A csökkentett keresési tér a hasonlósági függvény periodicitás miatt tört több részre a képen. Vörös és kék pontok jelölik a kimerítő keresés, illetve a PSO által talált globális optimumokat.

## 4. Tesztek

Ebben a szakaszban szemléltetjük, hogy eljárásunk helytáll fél-szintetikus tesztek során és valós fényképek esetén is. Tesztjeink során a 2D foltok vizsgálatához használt ablakméretet  $s = 70$ -nek választottuk, továbbá  $\sigma = \frac{s}{2}$  a normális eloszlás szórása.

### 4.1. Fél-szintetikus tesztek

Három különböző szintetikus jelenetet generáltunk a Processing<sup>1</sup> szoftver segítségével: egységgömb, egység-oldalú kocka, illetve egy összetett szintér. Az utóbbi két merőleges síkból, egy gömbből és egy paraméteres felületből áll. A belső és a külső paramétereket a későbbi lépések során ismertnek tekintettük. Különböző irányból készítettünk képeket a jelenetekről (lásd 3. ábra). Az ASIFT [22] eljárást használtuk jellemző képpont-párok kereséséhez. Ezután algoritmusunkat lefuttatunk minden pontpárra a felületdarabkák kiszámításához. A mért és a ground truth normálisok hibáját az átlagos szöghibával jellemeztük.

A 3. ábra a „Sphere”, „Cube” és „Complex” jelenetek két-két nézetét mutatja.<sup>2</sup> Az összehasonlítás kedvéért a jelenetet PMVS-el [7] is rekonstruáltuk, illetve a standard LS Plane [10] ponthalmazokra alkalmazható, legkisebb négyzetes síkillesztési eljárást is kiértékeltek.

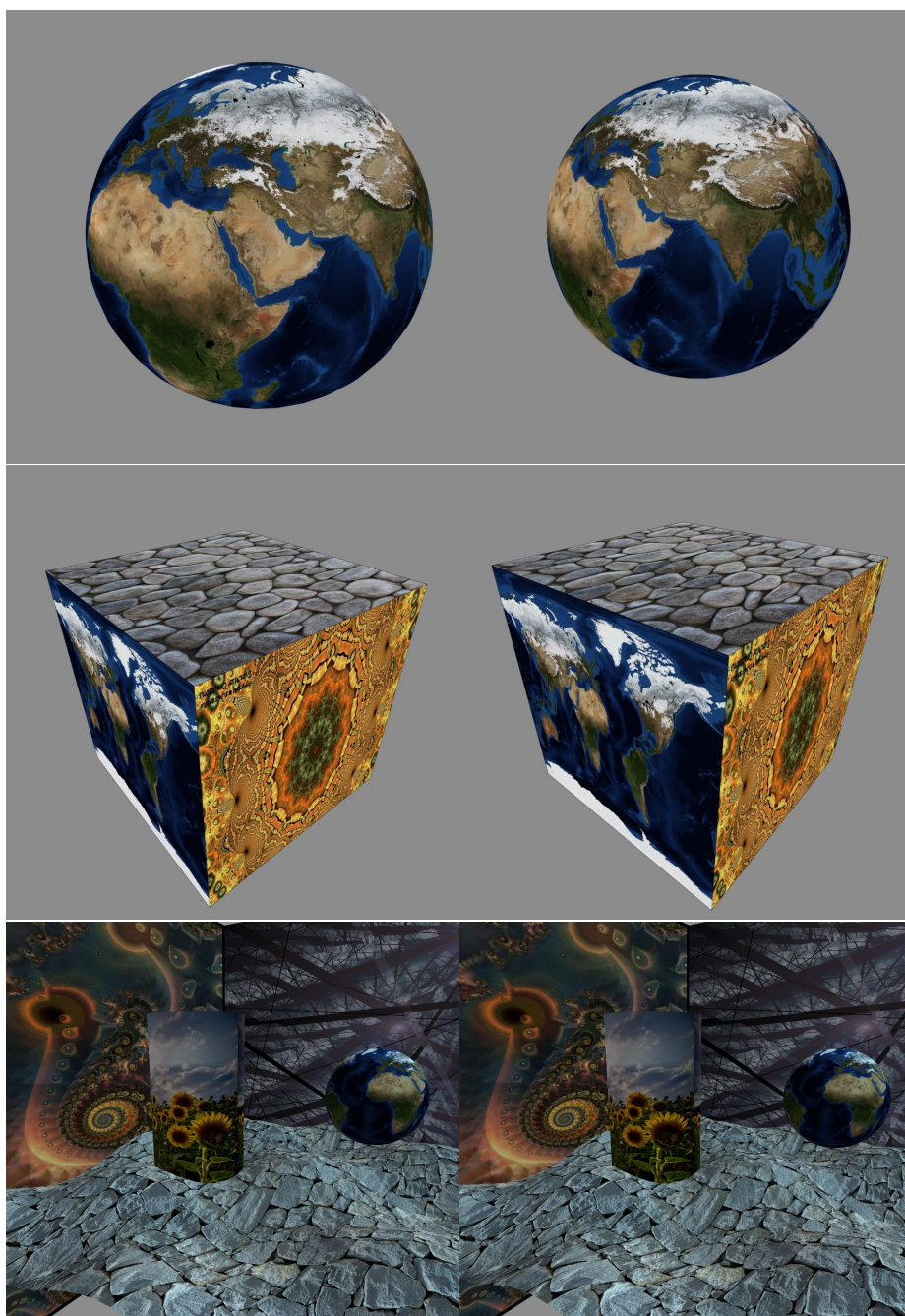
1. táblázat. Eredmények  $\sigma = 50$  szórással és 100 px ablakmérettel.

		#points	Avg. ang. err.	Med. ang. err.	Avg. dist. err.	Med. dist. err.
Sphere	ITPE-PSO	9492	5.5225°	3.4042°	<b>0.0310</b>	<b>0.0321</b>
	ITPE-ES		<b>5.5021°</b>	<b>3.3994°</b>		
	LS Plane		30.5298°	22.1780°		
	PMVS		12658	16.6978°		
Cube	ITPE-PSO	9960	2.0883°	1.1481°	<b>0.0581</b>	<b>0.0585</b>
	ITPE-ES		<b>2.0767°</b>	<b>1.1352°</b>		
	LS Plane		25.1969°	29.6932°		
	PMVS		13376	24.6775°		
Complex	ITPE-PSO	15343	6.3756°	3.4440°	<b>0.0181</b>	<b>0.0158</b>
	ITPE-ES		<b>6.3461°</b>	<b>3.4280°</b>		
	LS Plane		22.0623°	11.0703°		
	PMVS		47114	12.0152°		

A 1. táblázatban négy eljárást értékeltünk ki: ITPE-PSO, ITPE-ES, PMVS és LS Plane. Látható, hogy javasolt eljárásunk (ITPE-PSO) kisebb, mint 6.5°-os átlagos szöghibával teljesít bármely esetben és a hibák mediánja pedig rendre

<sup>1</sup><https://processing.org/>

<sup>2</sup>A 3D rekonstrukciós eredményeket Meshlab-ben [4] vizualizáltuk. A tesztkonfigurációról: Core(TM) i7-3610QM CPU at 2.30GHz, 8 hw szál és 8192MB RAM.



3. ábra. A szintetikus tesztek bemeneti képpárjai. A jelenetek rendre: „Sphere”, „Cube” és „Complex”.

3.5°, 1.2°, és 3.5° alatt van a gömb, kocka és összetett jelenetek esetén. Ez azt jelenti hogy a becült normálisok fele a „Cube” esetén közelebb van a ground truth értékekhez, mint 1.2°. A tesztek azt is mutatják, hogy a rivális eljárások megközelítőleg rendre 3-szor, 10-szer, illetve 2-szer pontatlanabbak a három teszt esetén. A 2. táblázatban a ITPE-PSO és ITPE-ES eljárások futásidejét láthatjuk. Hasonló eredmények (1. táblázat) mellett, ám közel 10-szeres gyorsasággal produkálja az eredményeket a PSO keresési heurisztikával ellátott eljárásunk, mint a kimerítő keresés.

2. táblázat. Pontonkénti futásideje az ITPE-PSO és ITPE-ES eljárásoknak (ablakméret: 100 px).

	Sphere	Cube
ITPE-PSO	<b>0.0265 sec</b>	<b>0.0283 sec</b>
ITPE-ES	0.1884 sec	0.2035 sec

Szintetikus tesztjeink megmutatták, hogy az algoritmus jól működik, alacsony futásidővel. A becült felületi normálisok átlagos szöghibája a gömb esetén 5.5°, a kocka esetén pedig 2.09°.

**Szabadformájú felületek** A javasolt algoritmust kézi kamerával készített fotókra is alkalmaztuk. A kamerákat sakktáblás módszerrel külön kalibráltuk. Egy stereo pár esetén relatív elhelyezkedésüket esszenciális mátrix felbontással számítottuk ki, felhasználva az OpenCV [2] könyvtárat. A jellemző pontokat ASIFT [17] segítségével nyertük ki.

A következő lépésként ITPE-PSO eljárást alkalmaztuk minden pontpárra. Hogy eredményeinket a látvány szempontjából is kiértékelhessük, a kimenetként kapott orientált pontfelhőre a Poisson felületrekonstrukciót [11] alkalmaztuk.

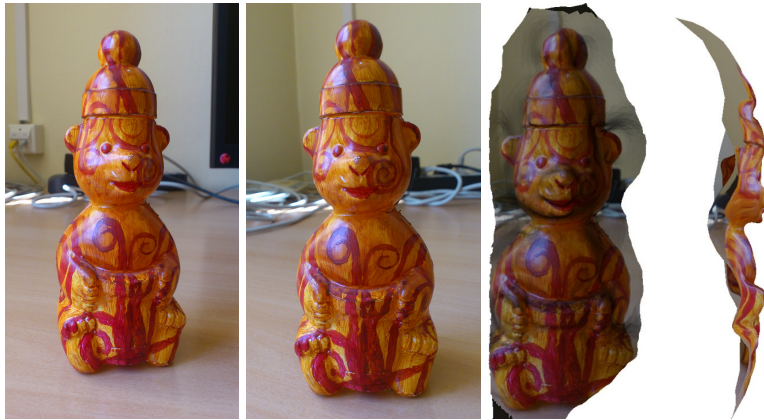
A 4. ábrán egy mintásra befestett medve figura látható, mely formájából adódóan nagy görbületeket is tartalmaz. Az ábra első két képe a stereo-nézet, melyet a rekonstrukcióhoz használtunk. A következő két kép pedig a létrehozott felület két különböző nézete. Ahogyan az látható, algoritmusunk kimenete pontosan követi az objektum formáját annak ellenére, hogy azt csak szórt pontokban vizsgálja. A visszaállított felület jó minősége legjobban a medve orra körül figyelhető meg. A teszt során levontuk a következtetést, hogy módszerünk a nagy görbületek ellenére sokkal finomabb kimenetet produkált, mint a MeshLab-ben használt LS Plane síkillesztő.

Összehasonlítottuk a kapott kimeneteket a PMVS<sup>1</sup> [7] kimeneteivel is. A fountain-P11<sup>2</sup> adathalmaz két képét választottuk ki bemenetnek és alkalmaztuk rá a rivális eljárásokat. Mindkét esetben a ritka rekonstrukcióból felületet állítottunk elő a Poisson eljárással, azonos paraméter-beállítás mellett.

<sup>1</sup><http://ccwu.me/vsfm>

<sup>2</sup><http://cvlabwww.epfl.ch/data/multiview/>





4. ábra. Medve: Az első két kép a bemeneti stereo képpár. Az utóbbi két kép pedig az algoritmusunk majd a Poisson felületrekonstrukció alkalmazásával kapott háromszögháló két nézete.

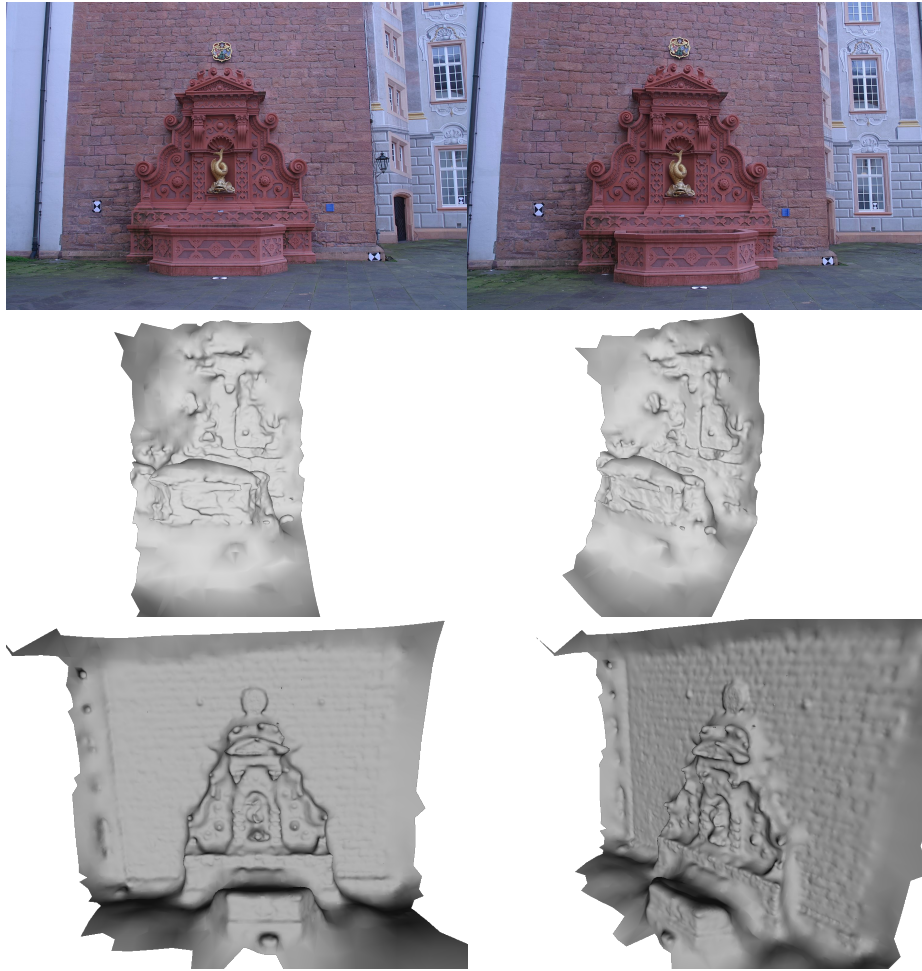
A 5. ábra bemutatja az ITPE és PMVS eljárások kimenetére alkalmazott Poisson felületrekonstrukciók minőségét. Az első sor a bemeneti stereo nézetet tartalmazza. Az ábra második és harmadik sorai pedig a PMVS, illetve az ITPE eredményeinek különböző nézőpontjait mutatja. Nyilvánvaló, hogy eljárásunk sokkal részletesebb kimenetet ad, azonos paraméter-beállítás mellett a Poisson rekonstrukcióhoz. Megjegyezzük, hogy a PMVS egy összetett eljárás, mely a miénkkel szemben sűrű rekonstrukciós, illetve felület-propagációs lépéseket is tartalmaz. Mindennek ellenére, az ITPE sokkal jobban közelíti a felületet azonos bemenet mellett.

## 5. Összegzés

Munkánkat motiválta, hogy jó minőségű felületi normálisokat állítsunk elő szemben a korszerű többnézetű rekonstrukciós eljárások által szolgáltatott gyengébb minőséggel. Javasolt eljárásunk jelentősen pontosabb érintősíkokat produkál, mint a rivális algoritmusok: kisebb szöghiba a normálisban, kisebb távolság-hiba a pontokban. Az affin transzformációkban rejő haszon, hogy a bemutatott elméletet és illesztési eljárást tetszőleges kameramodellek esetén alkalmazhatjuk. A bemeneti kameraképek rektifikálása, vagy kameramodellek közötti transzformációja felesleges.

Más eljárásokkal összehasonlítva [7,10,21,15,13] a mi megoldásunkban rejő újdonság három pontban foglalható össze:

1. Elméleti hozzájárulásunk: a keresési területet sikeresen szűkítettük epipoláris és geometriai kényszerek segítségével. Biztosított, hogy a megszorított keresési tér még mindig tartalmazza az optimális megoldást. A javasolt kényszerek könnyedén kiterjeszthetőek a többnézetű rekonstrukció esetére is.



5. ábra. A fountain-P11 adathalmaz rekonstrukciói. Első képpár: a bemeneti stereo nézet. Második képpár: PMVS + Poisson. Harmadik képpár: ITPE-PSO + Poisson.

2. A globális optimum gyors megtalálása érdekében a Particle Swarm Optimization eljárást alkalmaztuk. A megoldás jól párhuzamosítható, pontonkénti feldolgozási ideje 0.03 másodperc. Egy megfelelő GPU implementáció lehetővé tehet akár valós idejű rekonstrukciót is.

3. Alkalmazható számos kameramodell esetén, mint például a perspektív- vagy omnidirekcionális kamerák.

Úgy hisszük, hogy eljárásunk hatékony eszköz ritka rekonstrukciók készítésére és remek alapja lehet egy jövőbeli többnézetű rekonstrukciós eljárásnak.

## Hivatkozások

1. Daniel Barath, Jozsef Molnar, and Levente Hajder. Optimal Surface Normal from Affine Transformation. In *VISAPP 2015*, pages 305–316, 2015.
2. Gary Bradski et al. The OpenCV library. *Dr. Dobb's Journal of Software Tools*, 25(11):120–126, 2000.
3. Stefano Cagnoni. Evolutionary computer vision: a taxonomic tutorial. In *HIS'08. Eighth International Conference on*, pages 1–6. IEEE, 2008.
4. Paolo Cignoni, Massimiliano Corsini, and Guido Ranzuglia. Meshlab: an open-source 3d mesh processing system. *Ercim news*, 73:45–46, 2008.
5. Olivier Faugeras and Renaud Keriven. *Variational principles, surface evolution, pde's, level set methods and the stereo problem*. IEEE, 2002.
6. Olivier Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. Technical Report RR-0856, INRIA, 1988.
7. Yasutaka Furukawa and Jean Ponce. Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(8):1362–1376, 2010.
8. Martin Habbecke and Leif Kobbelt. Iterative multi-view plane fitting. In *Int. Fall Workshop of Vision, Modeling, and Visualization*, pages 73–80, 2006.
9. Martin Habbecke and Leif Kobbelt. A surface-growing approach to multi-view stereo reconstruction. In *CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.
10. Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. *Surface reconstruction from unorganized points*, volume 26. ACM, 1992.
11. Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, volume 7, 2006.
12. James Kennedy. Particle swarm optimization. In *Encyclopedia of Machine Learning*, pages 760–766. 2010.
13. Maxime Lhuillier and Long Quan. A quasi-dense approach to surface reconstruction from uncalibrated images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(3):418–433, 2005.
14. Jerome Martin and James L Crowley. Comparison of correlation techniques. In *International Conference on Intelligent Autonomous Systems, Karlsruhe (Germany)*, pages 86–93, 1995.
15. Zoltán Megyesi, Géza Kós, and Dmitry Chetverikov. Surface normal aided dense reconstruction from images. In *Proceedings of Computer Vision Winter Workshop. Telc:[sn]*. Citeseer, 2006.
16. József Molnár and Dmitry Chetverikov. Quadratic transformation for planar mapping of implicit surfaces. *Journal of mathematical imaging and vision*, 48(1):176–184, 2014.

17. Jean-Michel Morel and Guoshen Yu. Asift: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2(2):438–469, 2009.
18. Jean-Philippe Pons, Renaud Keriven, and Olivier Faugeras. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision*, 72(2):179–193, 2007.
19. Yuhui Shi and Russell Eberhart. A modified particle swarm optimizer. In *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, pages 69–73. IEEE, 1998.
20. Christoph Strecha, Rik Fransens, and Luc Van Gool. Combined depth and outlier estimation in multi-view stereo. In *CVPR'06 IEEE Computer Society Conference on*, volume 2, pages 2394–2401. IEEE, 2006.
21. Hoang-Hiep Vu, Patrick Labatut, Jean-Philippe Pons, and Renaud Keriven. High accuracy and visibility-consistent dense multiview stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):889–901, 2012.
22. Guoshen Yu and Jean-Michel Morel. A fully affine invariant image comparison method. In *ICASSP 2009. IEEE International Conference on*, pages 1597–1600. IEEE, 2009.
23. Andrei Zaharescu, Edmond Boyer, and Radu Horaud. Transformesh: a topology-adaptive mesh-based approach to surface evolution. In *ACCV'07*, pages 166–175. Springer, 2007.