

Benedek Csaba „Képi alapú többszintű környezetelemzés (Image based multi-level environment analysis)” című MTA doktori értekezésének bírálata

A disszertáció témájában a képfeldolgozáshoz tartozik, elsősorban a Markov mezők és pontfolyamatok eszköztárának alkalmazásával ért el új tudományos eredményeket, amelyeket közvetlenül alkalmazni lehet a környezetelemzésben, azaz például rövid és hosszú távú változás detektálásra, mozgó objektumok követésére, nyomtatott áramköri lapok minőségellenőrzésére, stb. A Markov-i megközelítés a gépi látás egyik robusztus és hatékony technikája, amelyet számos előnye mellett a relatíve nagyobb számításigény jellemzi. *Ezért kérem a Jelöltet, hogy helyezze el az eredményeit a feldolgozási idő, hardver igény skáláján is, és mutassa be, hogy mely algoritmusa alkalmas valós-idejű, interaktív, illetve off-line feldolgozásra.*

A dolgozat egy bevezető fejezetből, egy az alapokat bemutató fejezetből, és négy, az új eredményeket taglaló és a tézisek rendjét követő fejezetből áll, amit a tézisek összefoglalása és függelék egészítenek ki. A bevezető és az alapokat bemutató fejezet célirányosan és formálisan összefoglalja a tudomány jelenlegi állását a markovi képfeldolgozás területén, arra is rávilágítva, hogy az új eredmények milyen ismert módszerekre épülnek, és azokat hogyan alkalmazzák és fejlesztik tovább.

Az algoritmusok az energiaminimalizálási módszerek családjába tartoznak, amelyek lényege, hogy az elvárt megoldás tulajdonságait, azaz egy kívánság listát adunk meg, majd egy optimalizálási eljárásra, tipikusan a szimulált lehűtés valamely változatára bízunk, hogy a kívánságainkat a leginkább kielégítő megoldást megtalálja. A Markov véletlen mezők esetén az energia a kép elemeitől és a szomszédokkal fennálló viszonyaitól függ. A képelemek tipikusan a pixelek, a szomszédság pedig a pixelszomszédságból származtatható. A szomszédság kezelése a kritikus pont, mert így a becslést nem a különálló pontokra végezzük el, hanem feltételezünk valamilyen folytonosságot, azaz egy szűrést végzünk szofisztikált módon. A Bayes megközelítés a priori eloszlásokat is beépíthet a keresésbe. A kevert (mixed) Markov módszer a szomszédságot dinamikusan kezeli. A pont folyamat az elemeket választja ki dinamikusan, így azok nem feltétlenül a statikusan definiált pixelek, hanem az algoritmus futása során változó számú geometriai objektumok.

A bevezető röviden kitér a mélytanulás alapú megközelítésekre és elválasztja a saját módszereket azoktól. Jó lett volna röviden a jellemző alapú (pl. SIFT) megközelítéseket is hasonló módon elhatárolni.

A harmadik fejezet az első tézis kifejtése, amely a több szintű fúziós modelleket tárgyalja, két konkrét alkalmazást szem előtt tartva, a mozgó járművek azonosítását, illetve az épületek, növényzet változását légi fotók alapján. Az első alkalmazás rövid időléptékű, azaz feltételezhető, hogy a háttér nem változik, a második hosszú időléptékű, azaz valamekkora változásra mindenhol fel kell készülni. A bemutatott módszerek általánosak, nem csak ezekre az alkalmazásokra jók. Bár azt jónak tartom, hogy a Jelölt az általános módszert az elejétől kezdve ezen konkrét alkalmazásokon keresztül mutatta be, mert így könnyű a döntéseket megérteni, de szerencsés lenne kiemelni azokat az általánosítható fogásokat és alkalmazási kört, amit a Jelölt ezzel a módszerrel javasol megoldani.

A rövidtávú modell a képregisztráció után két jellemzőt fuzionál, az intenzitás különbséget és a környezetek keresztkorrelációját. *Kérdésem, hogyan lehetne a színinformációt felhasználni, annak*

elhanyagolása ugyanis jelentős információvesztést jelent. A Jelölt demonstrálja, hogy javasolt modell mind a két fuzionált módszernél, mind pedig a naív fúziónál lényegesen jobb eredményt ad. Megjegyzem, hogy a széleskörű elemzéshez egy saját adatbázis felépítésére is szükség volt.

A hosszú távú változásdetektáláshoz a Jelölt még egy újabb leíró alkalmaz, amely a többi leíró lokális megbízhatóságát mutatja. Ezt szép eredeti ötletnek tartom, az eredmények ismét jól alátámasztottak és meggyőzőek. *A hosszú távú modellhez kapcsolódóan a regisztráció módjára kérdezek rá és arra, hogy annak pontossága mennyire veszélyezteti a javasolt algoritmus pontosságát.*

A negyedik fejezet, azaz a második tézis, a Markov pontfolyamatok időtartománybeli kiegészítésével foglalkozik, amit egyrészt épületek megjelenésének és eltűnésének a hosszú távú követésére alkalmaz, másrészt pedig ISAR radar alapú objektum követésre. A klasszikus Markov pontfolyamatok objektumok listáit generálják statikus képek alapján, a listák pedig az optimalizációs megoldás során, tehát pszeudó időben változnak. A jelölt erre épített időbeli követést, ahol a valódi időbeli fejlődést vizsgálta, azaz nem egy statikus listát számol ki, hanem időfüggvényeket. A gépi látás és egyéb inverz feladatoknál (pl. dinamikus tomográfia) hatékony eljárás, ha nem egyenként vizsgáljuk a diszkrét idejű mérési pontokat, hanem egyben tekintjük a problémát, mert ekkor egy-egy mérés feldolgozásához a teljes mérési adathalmazban rejlő információ felhasználható, és az eredmények időbeli konzisztenciája is automatikusan garantált. A javasolt módszer ennek az alapelvnek szép alkalmazása a Markov-i pontfolyamatokra.

Az első gyakorlati alkalmazásban a „pontok” az azonosított épületek, amelyeket két képen kell megtalálni a különbségeket kiemelve. A második alkalmazásban pedig hajók vagy repülő, amelyek a radar általános orientációjú szakaszoknak vagy kereszteknek lát. Az ilyen megközelítésű módszerek fő kihívása a születési/halálozási folyamatok kontrollálása és az energiafüggvény meghatározása. A Jelölt a konkrét feladatra megalapozott döntésekkel javasol megoldást. *Kérdezem, hogy milyen szintig van lehetőség általános recepteket adni ezen problémák megoldására.* Erre a fejezetre is elmondható, hogy a javasolt módszerek jóságát bőséges tesztek és demonstrációk támasztják alá.

Az ötödik fejezet, azaz a harmadik tézis, a hierarchikus Markov pontfolyamatok fogalmát dolgozza ki. Ezek olyan folyamatok, ahol a „pontok” több hierarchikus szintet alkotnak, így a viszonyaikat is több szinten vizsgálhatjuk. A felvetés elegáns, lényegében az energia függvényt kell szülő-gyermek kapcsolati tagokkal kiegészíteni. Az első példa a módszer alkalmazására az épületek követése légi fotók alapján, ahol a szülő-gyermek kapcsolat az épületek és a kémények között áll fenn, illetve a azonos állású épületeket tartalmazó csoport és tagjai között. Az orientáció ad csoportosítási lehetőséget járművek LIDAR alapú követésénél is, nyomtatott áramköröknél pedig a forrasztások orientációja és alakja jelöli ki a családot.

A hatodik fejezet, azaz a negyedik tézis, az idő és tértartománybeli (4D) környezetfelügyelettel foglalkozik, ezen belül többkamerás személy lokalizációval, LIDAR alapú személydetektálással, követéssel, és tevékenység felismeréssel, valamint LIDAR alapú környezet rekonstrukcióval, amit pl. önvezető járművek alkalmazhatnak. Ez utóbbi, ismét demonstrálja a Markov-i módszer előnyeit a különálló becslésekhez képest helyett (itt kevert Gauss), mivel képes a ponthibák kiszűrésére. A Jelölt ezt

dinamikus Markov modell segítségével tovább javítja. A gyalogos követésnél több egyidejűleg mozgó gyalogost kell megkülönböztetni, azaz nem csak a helyet, hanem a trajektóriákat visszaállítani. Végül a képek időbeli átlagképéből következtetni lehet a személy által végrehajtott tevékenységre. A LIDAR alapú környezet rekonstrukció is új, előremutató eredmény, amelynek számos alkalmazása és továbbfejlesztése is elképzelhető. Erre a fejezetre is elmondható, hogy a Jelölt komoly gondot fordított az eredmények szimulációkkal történő bemutatására és a jóság igazolására.

A dolgozatban sok helyen előjön az algoritmusok paramétereinek beállítása, amely alapvetően befolyásolhatja a konvergencia létét vagy sebességét. A javasolt módszer általában egy fajta tanítás, amikor is ismert képek alapján lőjük be automatikusan ezeket a paramétereket, amelyeket azután már az éles képekre alkalmazunk. *Milyen más lehetőség volna még?*

Összefoglalva, a dolgozat számos értékes képfeldolgozási megoldást javasol, szimulációkkal igazolja azok képességeit és a korábbi módszereknél jobb voltát, és megmutatja a gyakorlati alkalmazásukat. Az eredményeket a Jelölt a témakör rangos fórumain, folyóirataiban publikálta. A publikációk rangja, száma és a hivatkozások sokasága is tiszteletre méltó. Az eredményeket tudomány állását előrevivőnek, igazoltan érvényesnek és a gyakorlatban alkalmazhatónak, illetve alkalmazottnak tekintem. A dolgozat jó felépítésű, nagyon színvonalas munka. Kritikai megjegyzésként megemlítem, hogy a dolgozat több helyen túlságosan formális, az érthetőséget javította volna, ha képletek helyett néhány mondatban fogalmazta volna meg a mondanivalót.

A dolgozatot eredményeit elegendőnek tartom az MTA doktori cím megszerzéséhez, a nyilvános védés kitűzését javaslom, a fokozat odaítélését támogatom.

Budapest, 2020. március 25.



Dr. Szirmay-Kalos László
MTA doktora, egyetemi tanár