

Csapattagok:

Atkári György

Bata Kristóf

Filep Tamás

Rafai Norbert

LÉPCSŐ- ÉS AJTÓFELISMERÉS VIDEÓFELVÉTELEKEN

1. MÉRFÖLDKŐ

1. SZAKIRODALOM FELDOLGOZÁSA

1.1 BEVEZETÉS

A feladat alapötlete egy SSIP konferenciás Zernike momentumokat használó feladat megoldásán alapul [1], a célfeladat ki lett bővítve, mivel nemcsak a képek csoportba sorolását kell elvégezni, hanem videófelvételen külön-külön meg kell címkézni az egyes ajtókat illetve a lépcsőket. A problémával külön-külön több publikáció is megjelent [1-10], ezért a problémát két részre bontva oldjuk meg külön foglalkozunk az ajtó és a lépcső detektálásokkal, valószínűleg az eltérő geometriai paraméterek miatt nincs jelen a szakirodalomban a két detektálás egyidejű megoldása. A probléma egyébként a robotnavigációs terület egyik sokat kutatott területe, több publikáció is megjelent ajtó detektálással foglalkozó robot navigációs rendszerekről [2 - 7] illetve lépcső detektálással foglalkozó navigációs rendszerekről [8-10], éppen ezért is választottuk ezt a témát, mint mérnökinformatikus mesterképzéses hallgatók. A szakirodalomban lévő publikációk alapján a 90%-os detektálási hatékonyság már jónak mondható.

1.2 ELŐFELDOLGOZÓ MŰVELETEK

A szakirodalomban található cikkek a képeken az alábbi képfeldolgozó műveletek végzik el:

- videó képkockákra való bontása
- él detektálás (Canny éldetektorral)
- Hough transzformáció
- képátméretezés
- küszöbértékes szűrés

A küszöbérték szűrés a későbbi felhasználástól függően történik, sok esetben egyszerű bináris képpé alakítják az előfeldolgozás során a képes, egyes algoritmusok azonban a kép részletes eloszlást is figyelembe veszik, aminek előnye a több információ, hátránya pedig a nagyobb számítási igény.

1.3 AJTÓ DETEKTÁLÁS SORÁN FIGYELEMBE VETT PARAMÉTEREK

1.3.1 OLDALAK ARÁNYA

Az ajtók ezenkívül meghatározott vertikális és horizontális arány intervallummal rendelkeznek, amely több információ kiegészítve jól felhasználható a detektálásukra. [5, 7]

1.3.2 FALTÓL VALÓ ELTÉRÉS

A faltól való megkülönböztetésre hisztogram alapú módszereket is használnak. Ezenkívül egyes esetekben kihasználható, hogy az ajtó nem párhuzamos a fal felülettel, hanem a tokozási kialakítás függően egy kicsit bentebb vagy kintebb helyezkedik el, szerencsés esetekben a tok és a küszöb vonala is detektálható [5].

1.3.3 KILINCS

Speciális geometriai adatként figyelembe lehet venni, hogy minden ajtó rendelkezik kilinccsel. Kétféle alapvető kilincs típust különböztethetünk meg a kerek alakú körkilincs, illetve a fogantyúval rendelkező szögletes. Ezek a különböző kilincsek detektálhatóak egy adott ROI-n belül [2].

1.4 LÉPCSŐ DETEKTÁLÁS

1.4.1 OLDALAK ARÁNYA

A lépcsők jól lehatárolható vertikális és horizontális arány intervallummal rendelkeznek, ami további információkkal kiegészítve jól felhasználható a detektálásukra. [4] Egy lépcsőtől mindenképpen elvárható, hogy a szélessége mindképpen nagyobb legyen, mint a magassága és a mélysége.

1.4.2 LÉPCSŐFOKOK PÁRHUZAMOSSÁGA

A lépcsőktől elvárható, hogy a lépcsőfokok egymással párhuzamosak legyenek, és a valóságban általában közelítőleg azonos mélységgel rendelkeznek [4,8].

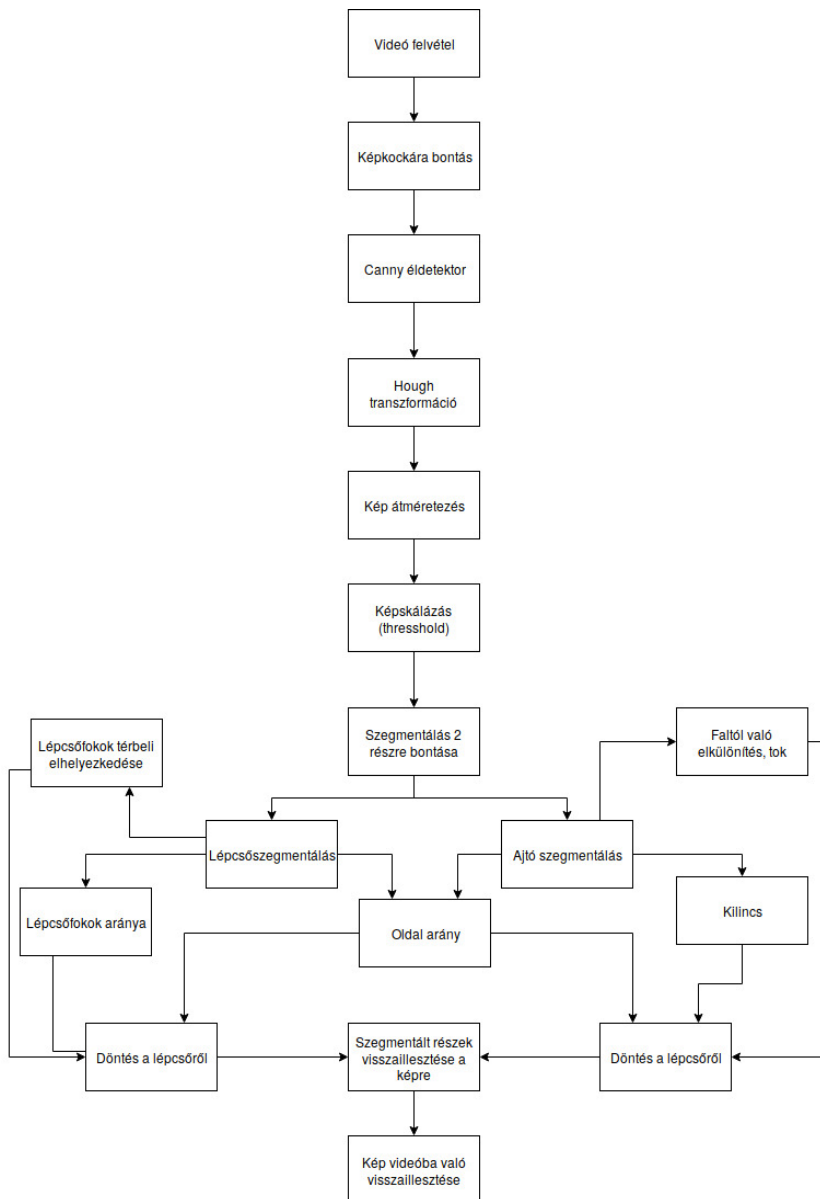
1.4.3 LÉPCSŐFOKOK TÉRBELI ELHELYEZKEDÉSE

A lépcsőktől elvárható, hogy a valóságban azonos magasságúak legyenek, tehát a lépcső szegletekre illeszthető legyen egy állandó meredekségű görbe (egyenes). Ezenkívül az egyes lépcső fok hisztogramjaiból is nyerhető ki a lépcsőfokokra vonatkozó információ, amelyből a lépcső iránya is meghatározható. [9]

2. VÁLASZTOTT MÓDSZEREK

2.1. IMPLEMENTÁCIÓ OPENCV-BEN

Az áttekintett szakirodalom alapján az egyik lehetséges választott módszer folyamatábrája a következő:



Egyes megoldások nemcsak az adott kép geometriáját veszik figyelembe, hanem egy tanító adatbázis alapján határozzák meg a szegmentálásnál használt paraméterek értékeit [1, 3, 5, 10], ha nem kapunk elég jó eredményt, akkor ilyen statisztika vizsgálatokkal, illetve a paraméterek megfelelő súlyozással kiegészítjük az algoritmust.

Az ajtónál a kilincs és a tokozás illetve a lépcsőfokok térbeli elhelyezkedésének figyelembevételétől eltekintünk abban az esetben, ha a használatuk során nem kapunk jobb eredményt vagy pedig azok nélkül is jó eredményt kapunk a teszteken.

2.2. MEGOLDÁS MATLAB-BAN, CNN SEGÍTSÉGÉVEL

A fentebb említett képfeldolgozást használó módszer mellett a szakirodalmak alapján szem előtt tartunk egy olyan módszert is, mely a gépi tanulást (deep learning) használja a címkék létrehozásához. A megoldás során a tanuló algoritmusok közül CNN-t (convolutional neural network), azaz konvolúciós neurális hálót használnánk.

Ehhez a MATLAB-hoz elérhető „Deep Learning for Computer Vision with MATLAB” példát vennénk alapul. A módszer során elengedhetetlen egy kellő számú tanuló adatbázis, mely mintaképeket tartalmaz különböző ajtó- illetve lépcső változatokról. Ezek alapján taníthatjuk az algoritmusunkat. Elérhető egy kellően nagy méretű (1,2 millió címkézett

nagyfelbontású kép, 1000 kategóriában) képi adatbázis, az ImageNet, melyet az általunk készített képekkel kibővíthetünk, és felhasználhatjuk a megoldás során.

Segítségünkre lehet továbbá a [12]-es szakirodalomban bemutatott MatConvNet, mely szintén hasonló feladatot valósít meg.

Amennyiben a neurális hálós megoldás túlzottan erőforrásigényes lenne, úgy a döntéshez használhatunk döntési fát is.

3. VÁLASZTOTT MÓDSZEREK KÖLTSÉGTERVE

3.1 HARDVER KÖLTSÉG

A módszerek egyike sem igényli különleges hardver eszköz beszerzését. A mindennapokban használt kamerák, fényképezőgépek, mobiltelefon kamerák által készített felvételek a célnak megfelelnek.

3.2 SZOFTVER KÖLTSÉG

Az implementációhoz használt OpenCV a 3 pontos (revidiált) BSD licenszt használ, a licenc részletes leírása az alábbi linken található:

<http://opencv.org/license.html>

Az operációs rendszerként használt Ubuntu illetve a használt c++ fordító pedig GNU GPL és GNU LGPL licenszek alá esnek, részletes információ:

<https://www.ubuntu.com/about/about-ubuntu/licensing>

A MATLAB-os megoldás során a költségtervbe bele kell foglalnunk a MATLAB szoftver árát, mely egyetemi hallgatóknak 35 EUR áron érhető el, ugyanakkor a standard verzió 2000 EUR:

<https://www.mathworks.com/pricing-licensing.html?prodcode=ML&intendeduse=comm>

SZAKIRODALOM

- [1] <https://ssip2016teamcblog.wordpress.com/>
- [2] http://cdn.intechopen.com/pdfs/10237/InTechApproaches_to_door_identification_for_robot_navigation.pdf
- [3] <https://arxiv.org/pdf/1301.0432.pdf>
- [4] <https://pdfs.semanticscholar.org/2d70/ab62aa5378efe9354d8e6091cdb641ee9830.pdf>
- [5] <https://pdfs.semanticscholar.org/8369/06b2ca9109760d370ee567ba4cd8b7ce9f33.pdf>
- [6] http://www.irisa.fr/lagadic/pdf/2013_icip_sekkal.pdf
- [7] http://www-ee.ccny.cuny.edu/wwwn/yltian/Publications/CVAVI2010_Yang.pdf
- [8] http://www.mrt.kit.edu/z/publ/download/harms/StairDetection/2015_Harms_StairDetection.pdf
- [9] <http://www-ee.ccny.cuny.edu/wwwn/yltian/Publications/ICME16-W155.pdf>
- [10] http://www.iaeme.com/MasterAdmin/uploadfolder/IJCET_07_03_008/IJCET_07_03_008.pdf
- [11] <https://www.mathworks.com/company/newsletters/articles/deep-learning-for-computer-vision-with-matlab.html>
- [12] <https://arxiv.org/pdf/1412.4564.pdf>