

Semantička segmentacija RGB-D slike unutrašnjih prostora

Šimić, Kristijan

Master's thesis / Diplomski rad

2015

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:551948>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-15**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
ELEKTROTEHNIČKI FAKULTET

Sveučilišni diplomski studij računarstva

SEMANTIČKA SEGMENTACIJA RGB-D SLIKE
UNUTRAŠNJIH PROSTORA

Diplomski rad

Kristijan Šimić

Osijek, 2015.

SADRŽAJ

1. UVOD	1
2. SEMANTIČKA SEGMENTACIJA SLIKE PUTEM UVJETNIH SLUČAJNIH POLJA	2
2.1. Struktura uvjetnog slučajnog polja.....	3
2.1.1. Minimalno razapinjuće stablo grafa.....	4
2.2. Potencijali uvjetnog slučajnog polja	6
2.2. Kd-stabla	10
2.3. Algoritam K najbližih susjeda.....	12
2.3. Zaključivanje i dekodiranje	14
3. SEGMENTIRANJE SLIKE NA SUPERPIKSELE	18
3.1. RGB prostor boja	18
3.2. CIE L*a*b* prostor boja	19
3.2.1. Transformacija iz RGB u CIE L*a*b* prostor boja.....	19
3.3. SLIC Algoritam.....	20
3.3.1. Postprocesiranje grupa piksela.....	23
3.3.2. Programska implementacija SLIC algoritma.....	23
4. ZNAČAJKE SUPERPIKSELA.....	26
4.1. 3D značajke	27
4.1.1. Pronalaženje dominantne ravnine RANSAC metodom	28
4.1.2. Računanje 3D značajki	29
4.2. Značajke slike	31
5. EKSPERIMENTALNI REZULTATI	34
6. ZAKLJUČAK	42
LITERATURA	43
SAŽETAK.....	44
ABSTRACT	45

SAŽETAK

U ovom radu je analiziran jedan pristup rješavanju problema semantičke segmentacije dubinskih slika unutrašnjih prostora. Proučeni su algoritmi korišteni u predloženoj metodi, kao i njihove programske implementacije. Rješavanjem problema semantičke segmentacije slike znatno bi se poboljšala sposobnost robota za samostalan rad u nepoznatim okruženjima. Problem se svodi na svrstavanje dijelova slike u neke unaprijed određene kategorije. U ovom slučaju te kategorije su pod, strukture, namještaj i rekviziti. Predložena metoda kategorizira dijelove slike koristeći uvjetna slučajna polja. Slika se najprije segmentira na superpiksela, koji služe kao čvorovi grafa uvjetnog slučajnog polja. Potencijali uvjetnog slučajnog polja određeni su vizualnim i geometrijskim značajkama superpiksela. Struktura grafa je minimalno razapinjuće stablo, pri čemu je težina brida određena na temelju udaljenosti centroida dvaju superpiksela u trodimenzionalnom prostoru. U okviru ovog rada proučena je jedna metoda zaključivanja i dekodiranja uvjetnih slučajnih polja koja omogućuje pronalaženje optimalne segmentacije. Učinkovitost predložene metode ispitana je eksperimentima na ispitnom skupu dubinskih slika unutrašnjih prostora.

Ključne riječi: računalni vid, semantička segmentacija, uvjetna slučajna polja, minimalno razapinjuće stablo, algoritam k najbližih susjeda, propagacija vjerojatnosti, RGB-D slika, SLIC superpikseli

ABSTRACT

This thesis considers an approach for semantic segmentation of RGB-D images of indoor environments. The algorithms used in this approach are described and their implementation in the provided source code is explained. Solving the problem of semantic RGB-D image segmentation would significantly improve the capacity of robots for autonomous problem solving in unknown environments. The considered problem is to classify parts of an image into a set of predefined categories. In this case, the categories are ground, structure, furniture and props. The discussed method does this using the conditional random fields. The image is first segmented into superpixels, which serve as nodes in a CRF graph, where the CRF potentials are defined by visual and geometric features of the superpixels that the image is divided into. The structure of the CRF graph is a minimum spanning tree, with the weight of the graph edges being defined by the distance between the superpixel centroids in three-dimensional space. One method for CRF inference and decoding, which provides an optimal segmentation, is examined. The efficacy of the considered approach is tested on a set of RGB-D images of indoor environments.

Key words: computer vision, semantic segmentation, conditional random fields, minimum spanning trees, k-NN, belief propagation, RGB-D, SLIC superpixels