CZECH TECHNICAL UNIVERSITY IN PRAGUE

FACULTY OF ELECTRICAL ENGINEERING

BACHELOR THESIS

Martin Pospíšil

Object Tracking for Robotic Formation Control

Department of Cybernetics

Supervisor: **Ing. Jan Faigl, Ph.D.**

Prague, 2012

Prohlášení autora práce

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne 30.5.2012 .... M. Pospíšil ....

podpis autora práce

Czech Technical University in Prague
Faculty of Electrical Engineering

Department of Cybernetics

BACHELOR PROJECT ASSIGNMENT

Student: Martin P o s p í š i l

Study programme: Cybernetics a Robotics

Specialisation: Robotics

Title of Bachelor Project: Object Tracking for Robotic Formation Control

Guidelines:

1. Study vision-based formation control technique [1] using entropy-based motion segmentation [2].
2. Study clustering techniques [3].
3. Implement the object-tracking algorithm [2] for IMR's sensor modules based on the Gumstix Caspa module [4] and using ROS [5].
4. Experimentally verify the algorithm in "leader-follower" task.
5. Determine operational parameters of the developed tracking module and constraints restricting motion of robots in a formation.

Bibliography/Sources:

[1] Hyeun Jeong Min, Andrew Drenner, and Nikolaos Papanikolopoulos: Vision-based leaderfollower formations with limited information. In Robotics and Automation, 2009. ICRA '09. IEEE International Conference on, pages 351-356, May 2009.
[2] Hyeun Jeong Min and N. Papanikolopoulos: Entropy-based motion segmentation from a moving platform. In Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on, pages 4559-4564, oct. 2009.
[3] Christian Bohm, Christos Faloutsos, Jia-Yu Pan, and Claudia Plant. Robust informationtheoretic clustering. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '06, pages 65-75, New York, NY, USA, 2006. ACM.
[4] Gumstix CaspaTM VL. https://www.gumstix.com/store/product_info.php?products_id=260.
[5] ROS - Robot Operating System. http://www.ros.org/wiki.

Bachelor Project Supervisor: Ing. Jan Faigl, Ph.D.

Valid until: the end of the winter semester of academic year 2012/2013

prof. Ing. Vladimír Mařík, DrSc.
Head of Department

L.S.

prof. Ing. Pavel Ripka, CSc.
Dean

Prague, December 9, 2011

České vysoké učení technické v Praze
Fakulta elektrotechnická

Katedra kybernetiky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Martin Pospíšil

Studijní program: Kybernetika a robotika (bakalářský)

Obor: Robotika

Název tématu: Sledování objektu v úloze řízení formací

Pokyny pro vypracování:

1. Seznamte se s problémem řízení formací typu ,,leader-follower" [1] založené na vizuálním sledování [2].
2. Seznamte se principem shlukové analýzy a vybranými přístupy řešení [3].
3. Implementujte algoritmus sledování objektu [2] pro senzorové moduly IMR založené na Gumstix Caspa kamerovém modulu [4] s využitím prostředí ROS [5].
4. Experimentálně ověřte činnost algoritmu v úloze sledování robotu.
5. Stanovte operační parametry realizovaného modulu sledování objektu a podmínky pro pohyb robotů ve formaci.

Seznam odborné literatury:

[1] Hyeun Jeong Min, Andrew Drenner, and Nikolaos Papanikolopoulos: Vision-based leaderfollower formations with limited information. In Robotics and Automation, 2009. ICRA '09. IEEE International Conference on, pages 351-356, May 2009.
[2] Hyeun Jeong Min and N. Papanikolopoulos: Entropy-based motion segmentation from a moving platform. In Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on, pages 4559-4564, oct. 2009.
[3] Christian Bohm, Christos Faloutsos, Jia-Yu Pan, and Claudia Plant. Robust informationtheoretic clustering. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '06, pages 65-75, New York, NY, USA, 2006. ACM.
[4] Gumstix CaspaTM VL. https://www.gumstix.com/store/product_info.php?products_id=260.
[5] ROS - Robot Operating System. http://www.ros.org/wiki.

Vedoucí bakalářské práce: Ing. Jan Faigl, Ph.D.

Platnost zadání: do konce zimního semestru 2012/2013

prof. Ing. Vladimír Mařík, DrSc.
vedoucí katedry

L.S.

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 9. 12. 2011

Abstrakt

Bakalářská práce se zabývá ověřením funkčnosti nového na entropii založeného segmentačního algoritmu pro extrakci pohybujícího se objektu z pohybující se platformy, který je popsán v článku [1]. Tento algoritmus využívá korekce obrázků založených na kombinaci optického toku a matched features, entropii z teorie informace k získání nejlépe informujícího atributu, od kterého jsou pomocí Gaussovy distribuce aplikované na histogramy vytvořeny shluky. Dále využívá porovnávání a modelování cíle.

Abstract

This thesis deals with verifying the novel entropy-based segmentation algorithm for extracting a moving object from a moving platform, which is proposed in article [1]. The algorithm uses image correction based on combining optical flow and matched features, entropy from information theory to determine best informative attribute, Gaussian Distributions applied to the histogram of features to building clusters and target modeling and matching.

Acknowledgements

I would like to thank my supervisor Ing. Jan Faigl, Ph.D. for his knowledge support and patience with me, my family and friends for their moral support.

Contents

List of Figures

List of Tables

Introduction

One of motivation of robotics is consting moving platforms like robots for using them to search, rescue etc., where environment is too dangerous for human. These robots can not rely on one of method communication with surrounding environment and their programme equipment must not contain unreal presumtions in this case.

One of method of interaction with surrounding environment is visual tracking of a target. Neither unreal presumptions of target as prediction models, nor unreal presumptions of platform with camera can not be performed.

When camera is moving, it can not be used background subtraction. There is necessary to recognize the target in real-time from noisy images, so for recognizing the target should be used robust clustering algorithm [2].

Clustering algorithms often need parameters based on view of camera, but staticaly given parameters are not sufficient enough for continuously changing view of camera.

Min and Papanikolopuolos introduce a new entropy-based clustering algorithm[1], where required parameters are automaticaly configured for clustering algorithm. They tested algorithm in leader-follower formations [3] using eROSI and Explorer robots [4].

This thesis deals with verifying and implementing this novel entropy-based segmentation algorithm for extracting a moving object from a moving platform. The algorithm uses image correction based on combining optical flow and matched features, entropy from information theory to determine best informative attribute, Gaussian Distributions applied to the histogram of features to building clusters and target modeling and matching. Operational parameters should be determined.

Implementation differences of algorith are described in Chapter 1 and experiments with implemented algorithm are shown in Chapter 2.

Chapter 1

Implementation

I started implementation in C++ with using SDL[5] and GSL[6] libraries, but GSL has not support for convolution and own implementation of convolution was too slow for using them in own implementation of Harris corner detector. Therefore this code was dropped and now it is using OpenCV[7] library.

Implemented part of the algorithm for estimating relative position of the leader/object consists of the following sequence of steps:

1. Two images are loaded and converted into gray scale.
2. Correction of the second image.
3. Harris corner detector [8] is applied for determining corners and edges in the images.
4. Temporal Difference method is applied providing new image TD.
5. Histograms for each axis of the TD are computed.

The input of the algorithm is a pair of two images of the scene taken in time t_1 and t_2, where $t_2 > t_1$ is sufficiently small to detect changes of the tracked objects.

Particular implementation details of the all algorithm's steps are described in the following sections together with the proposed changes to the original algorithm [1].

1.1 Image correction

Translation and scale on the image is applied for image correction. Translation and scale factors are computed from matched features differences of their positions in the images. To get matched features SURF[9] and FLANN[10] methods are used. In particular, implementations provided from the OpenCV is used. Only good matching features (i.e., those having small distances) are used for determining the required parameters of the image correction.

Only static objects are used to compute the factors. The static objects are identified as:

$$S_i = \{\vec{u} : \|f_i(\vec{u})\| < \frac{1}{N}\sum_{i=1}^{N}\|f_i(\vec{u})\|\}, \tag{1.1}$$

where $f_i(\vec{u})$ is a pair of matched features, and $\vec{u} = \begin{pmatrix} x, & y \end{pmatrix}^T$.

After that, static matched features are classified into classes $\{NE, SE, NW, SW\}$ according to differences of the $\vec{u}$ of the matched features dX and dY, see Table 1.1.

Table 1.1: Classification of matched features by orientation of motion

Orientation	$dX < 0$	$dX \geq 0$
$dY < 0$	NE	SE
$dY \geq 0$	NW	SW

An image is corrected by following affine tranformation matrix:

$$\mathbf{T} = \begin{bmatrix} \delta_x & 0 & \alpha_x - \varepsilon_1 \\ 0 & \delta_y & \alpha_y - \varepsilon_2 \end{bmatrix}, \tag{1.2}$$

where δ_x and δ_y are scaling factors, α_x and α_y tranlational factors and $\varepsilon_1, \varepsilon_2 \approx 0$, i.e.,

$$\begin{aligned} \alpha_x &= \frac{1}{|S|}\sum_{i=1}^{|S|} S_i[x] \\ \alpha_y &= \frac{1}{|S|}\sum_{i=1}^{|S|} S_i[y] \end{aligned} \tag{1.3}$$

$$\begin{aligned}\delta_x &= 1 + \frac{m_W[x] - m_E[x]}{c_x}\\ \delta_y &= 1 + \frac{m_S[y] - m_N[y]}{c_y}\end{aligned} \tag{1.4}$$

where $m_W[x]$ and $m_E[x]$ are means of the classes $\{\mathrm{NW}, \mathrm{SW}\}$ and $\{\mathrm{NE}, \mathrm{SE}\}$ on the x-axis, $m_S[y]$ and $m_N[y]$ are means of the classes $\{\mathrm{SE}, \mathrm{SW}\}$ and $\{\mathrm{NE}, \mathrm{NW}\}$ on the y-axis and c_x, c_y are experimentaly found parameters.

1.2 Entropy-based Clustering

1.2.1 Choosing the principal attribute

Entropy of each attribute can be rewrited as

$$E(a_i) \equiv H(B|A = a_i) = -\sum_{j=1}^{k} p_{B|A}(b_j|a_i)\ \log_2 p_{B|A}(b_j|a_i), \tag{1.5}$$

where A is a set of attributes, B is a set of possible values and $p_{B|A}(b_j|a_i)$ is the probability function of a_i for each b_j estimated from the from histogram.

The best informative attribute has minimum entropy. The authors of original article use equation **??**, but there exists optimal solution to search attribute with minimum entropy:

$$\begin{aligned}a_{max} =& \arg\max_{a_i \in A} \left(1 - \frac{E(a_i)}{\sum\limits_{j=1}^{n} E(a_j)} \right) = \arg\max_{a_i \in A} \left(-\frac{E(a_i)}{\sum\limits_{j=1}^{n} E(a_j)} \right) = \\ &= \arg\min_{a_i \in A} \frac{E(a_i)}{\sum\limits_{j=1}^{n} E(a_j)} = \arg\min_{a_i \in A} E(a_i)\end{aligned} \tag{1.6}$$

Adding of constant (In this case is 1.) or dividing with same sum do not change results in this case.

Chapter 2

Experiments

The implementation of algorithm was tested using images grabbed from Gumstix Caspa^{TM} VL module [11] in offline mode. It means that images were saved as files and later were used.
The images with similar properties combines into dataset.

2.1 Static camera dataset

All images of this dataset have resolution 752x480, they are unsharp. Example of time-near images is in Figures 2.1.

(a) t_1. (b) t_2.

Figure 2.1: An example of grabbed images.

The images are converted into gray-scale and matched features are found using SURF and FLANN. Matched features are shown in Figure 2.2, where small circles are found SURF that are matched using FLANN.

Figure 2.2: Matched features.

Then, the parameters of transformation matrice in Equation 1.2. Because camera is not moving in this dataset, δ_x, δ_y in Equation 1.4 should be close to 1 and α_x, α_y near to 0.

Detected edges and corners (found by Harris corner detector) are visualized in Figures 2.3 as white segments.

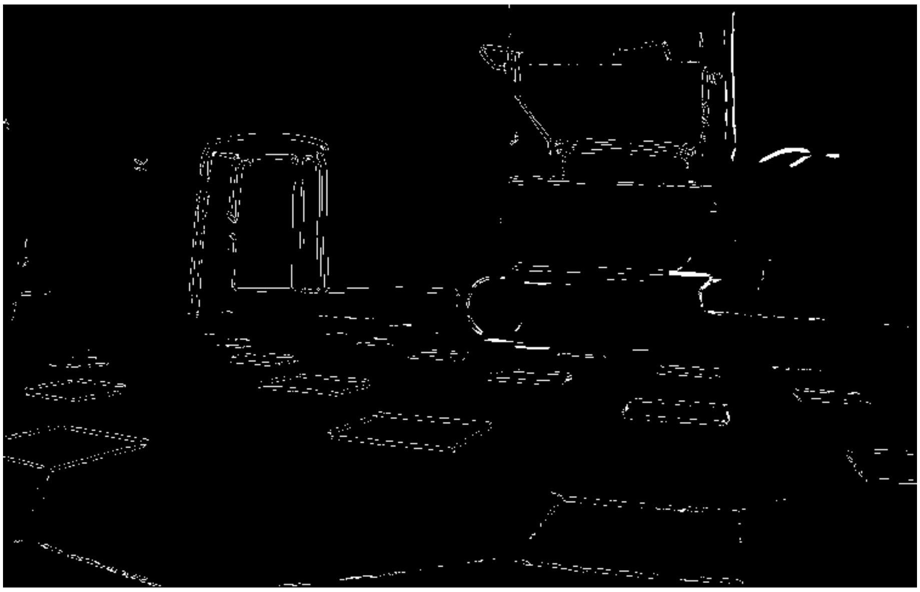
(a) At time $t1$.

(b) At time $t2$.

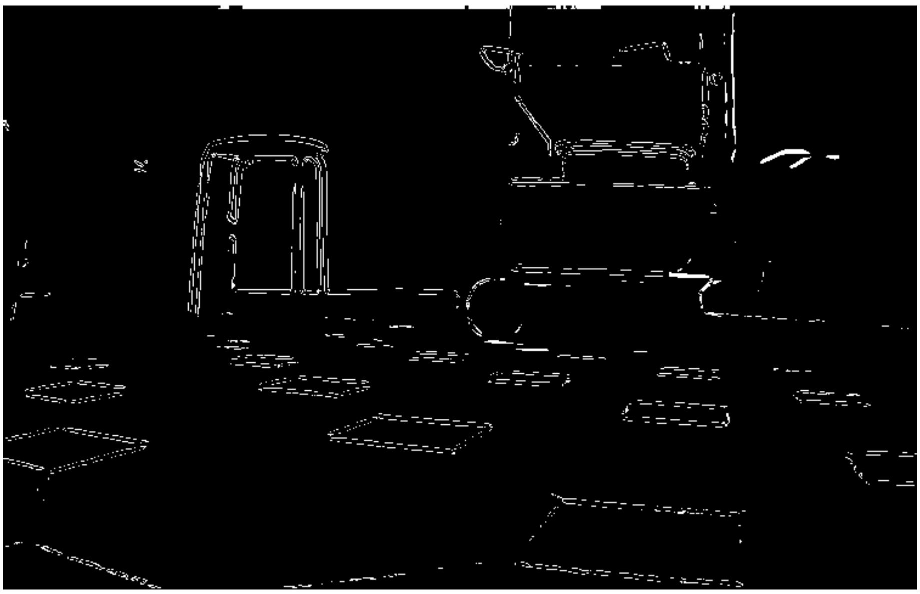
(c) Corrected image at time $t2$.

Figure 2.3: Images with features.

After that, TD image is built from features extracted by Harris corner detector. The process is visualized in Figures 2.4 and 2.3 where Figure 2.4(a) is created from Figures 2.3(a) and 2.3(b) and Figure 2.4(b) is created from Figures 2.3(a) and 2.3(c). because camera is not moving.



(a) Without correction.



(b) With correction.

Figure 2.4: Temporal Difference.

Chapter 3

Conclusion

On figures 2.4 should not be detected edges of static objects. Maybe the autors of the original algorithm uses filter.

Next parts of the algorithm was not implemented due time wasting caused by deadlocks in dead ends, a lot of tries with correction and personal problems. So operation parameters could not be determined.

Bibliography

[1] Hyeun Jeong Min and N. Papanikolopoulos. Entropy-based motion segmentation from a moving platform. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 4559 –4564, oct. 2009. doi:10.1109/IROS.2009.5354832.

[2] Christian Böhm, Christos Faloutsos, Jia-Yu Pan, and Claudia Plant. Robust information-theoretic clustering. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '06, pages 65–75, New York, NY, USA, 2006. ACM. URL: http://doi.acm.org/10.1145/1150402.1150414,doi:10.1145/1150402.1150414.

[3] Hyeun Jeong Min, Andrew Drenner, and Nikolaos Papanikolopoulos. Vision-based leader-follower formations with limited information. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 351 –356, may 2009. doi:10.1109/ROBOT.2009.5152261.

[4] M. Walter, M. Anderson, I. Burt, and N. Papanikolopoulos. The design and evolution of the erosi robot. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 2984 –2989, april 2007. doi:10.1109/ROBOT.2007.363925.

[5] SDL - Simple Directmedia Layer. [cited 28 May 2012]. URL: http://www.libsdl.org/.

[6] GSL - GNU Scientific Library. [cited 28 May 2012]. URL: http://www.gnu.org/software/gsl/.

[7] OpenCV - Open Source Computer Vision. [cited 28 May 2012]. URL: http://opencv.willowgarage.com/wiki.

[8] C Harris and M Stephens. *A Combined Corner and Edge Detection*, volume 15, pages 147–151. 1988. URL: `http://www.csse.uwa.edu.au/~pk/research/matlabfns/Spatial/Docs/Harris/A_Combined_Corner_and_Edge_Detector.pdf`.

[9] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359, June 2008. URL: `http://dx.doi.org/10.1016/j.cviu.2007.09.014`, `doi:10.1016/j.cviu.2007.09.014`.

[10] Marius Muja and David G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *International Conference on Computer Vision Theory and Application VISSAPP'09)*, pages 331–340. INSTICC Press, 2009.

[11] Gumstix caspaTM VL. [cited 28 May 2012]. URL: `https://www.gumstix.com/store/product_info.php?products_id=260`.

Content of CD

Included CD contains text of this thesis in PDF format with source code of complete text for LaTeX. In following table is described structure of CD.

Table 1: Content of CD

Directory	Description
`src`	source code of implemented program
`doc`	source code of text of this thesis
`thesis.pdf`	text of this thesis