Novel features in object detection for image segmentation tasks

Theses of the Ph.D. Dissertation

Andrea Manno -Kovács
computer engineer

Scientific advisers:
Tamás Szirányi, D.Sc.
Zoltán Vidnyánszky, D.Sc.

fides et ratio

Faculty of Information Technology
Pázmány Péter Catholic University

MTA SZTAKI

Computer and Automation Research Institute
Hungarian Academy of Sciences

Budapest, 2013

1 Introduction and aim

Automatic detection is a very important task in several computer vision and image understanding applications. Nowadays, with the widespread availability of affordable digital imaging devices and the presence of high capacity personal computers, significance of digital image processing is increasing substantially. As the amount of digital data is huge and manual administration and operation is unmanageable, automatic processing techniques are continuously improved to solve complex challenges in various fields of interest, like video surveillance, change detection, medical image analysis, force protection and defense applications, urban area extraction and building detection in aerial images.

As the large variety of applications shows, there is a wide range of tasks to be resolved: different fields have separate concepts and methodology, therefore context-sensitive solutions should be developed to satisfy the special conditions, cope with the altering challenges and reach the exact goals. The aim of this dissertation is to present contributions in three main tasks of automatic detection. Although these tasks are related to each other, and the given contributions of different tasks can be fused to be applied for complex solutions, these tasks should be handled separately. The given solutions can all be labeled as techniques for object featuring, but the distinct aims and applications (like extraction, tracking, change detection) need different tools and developments. This thesis concentrates on three selected tasks from object featuring.

Task 1 is giving a low dimensional feature descriptor by exploiting local structure information around feature points. In this task, the usability of the local shape representation is investigated, where an image series or video is given with a fixed camera position either remaining static or rotating during scanning. Local properties are extracted by generating active contour in the small proximity of the feature point, which is then represented by low dimensional Fourier descriptors. The goal is to match feature points through this descriptor set for further post-processing (tracking, classification and change detection).

Task 2 concentrates on the detection accuracy improvement of parametric active contour algorithms for high curvature, noisy boundaries and giving an automatic initialization technique on single object images. This involves the analysis of the behavior of existing parametric active contour methods and the development of a new feature map which is able to emphasize complex boundary parts and support initialization by defining feature points simultaneously. Moreover, the proposed feature map and point set is adapted successfully for other change detection and multi-object detection applications in registered aerial and medical image series as well.

In *Task 3*, the aim is detecting built-in areas and building outlines in single airborne images by introducing orientation of the close proximity of the feature points as a novel feature. This task involves the comparison of different feature point detectors for built-in area detection, the statistical analysis of extracted orientation feature to define main directions of the urban area, and a building detection process applying the given directions for extracting the accurate shape of the buildings without any restrictions (like shape template).

2 Methods used in the experiments

The majority of my work is connected to shape analysis with Active Contour method. This approach is energy minimization, driven by different forces, representing the characteristics of the deformable curve and the image as well:

$$E = \int_0^1 \frac{1}{2}(\alpha \left|\mathbf{x'}(s)\right|^2 + \beta \left|\mathbf{x''}(s)\right|^2) + E_{\mathrm{ext}}(\mathbf{x}(s))\mathrm{d}s, \qquad (1)$$

where α and β are weighting parameters for the elasticity and rigidity components of the internal energy; $\mathbf{x'}(s)$ and $\mathbf{x''}(s)$ are the first and second order derivatives with respect to s. In my work, I concentrated on the improvement of E_{ext}, the external energy, derived from the image.

Contributions of the thesis are presented in low-level shape description with Fourier methods, efficient feature point detection and feature extraction, edge detection by shearlets (wavelets) combined with mathematical toolkits: classification and graph theory.

Images used for evaluation in the thesis are partially coming from publicly available image datasets (Brodatz, Weizmann dataset). In *Task 1* I also used a video set taken by an outdoor surveillance camera of a city police central. Magnetic Resonance Imaging (MRI) scans used in *Task 2* were provided by Péter Barsi, MR Research Center, Semmelweis University. Airborne images used in the evaluations in the different tasks were bought from the Hungarian Institute of Geodesy, Cartography and Remote Sensing (FÖMI).

The software design and implementation was performed in Matlab environment. The thesis and my corresponding publications were written in LaTeX.

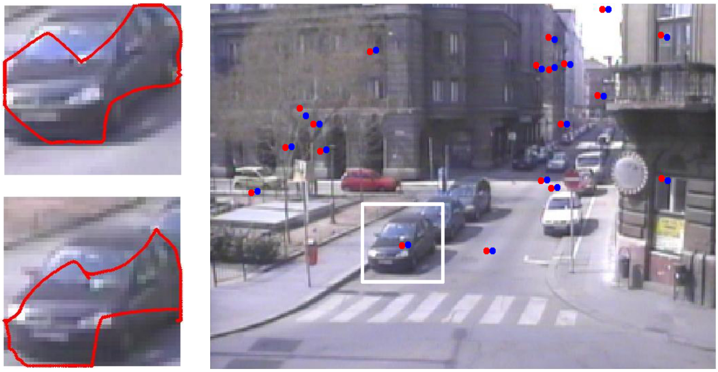
3 New scientific results

1. Thesis: *I have proposed a novel, active contour-based descriptor set for characterizing the neighborhood for scale-invariant feature points. I have experimentally shown that local contour generated around a feature point efficiently represents the main characteristics of the local neighborhood and the local contour descriptor, retrieved as a low dimensional interpretation of the local contour, is an efficient compressed descriptors which can be adapted for computer vision tasks (e.g. tracking, classification and change detection).*

Related publications: [4],[6], [7], [12], [14].

Describing local patches to register image feature points (keypoints) is an important task for many applications in computer vision. When searching for an efficient descriptor, the task is twofold: features must describe the characteristic patches at a high efficiency, while the dimensionality should be kept at a manageable low value.

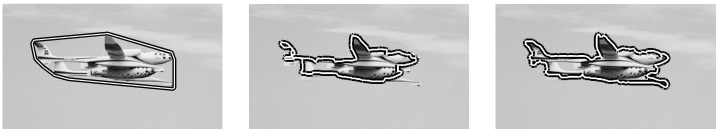
By investigating the potential applicability of methods in which some formal meaning of the local properties is represented at a reduced dimension, I have found that active contours generated around keypoints (called local contours) can be applied for designing an efficient descriptor set for comparing image regions. However, as the dimensionality of contour descriptor was high, I have introduced a feature set with reduced dimension, by characterizing the contour with Fourier descriptors and keeping the main coefficients. I have shown that the novel, low dimensional descriptor set can be efficiently adapted for different computer



(a) (b)

Figure 1: Local contour results on different frames for a coherent point: The contour represents the local structure and preserves the main characteristics, which can be adapted for point matching [Thesis 1].

vision tasks. Point matching is performed by using a distance metric for the comparison of introduced descriptors. For texture classification, I have introduced dynamic radius by investigating the variance of Fourier coefficients, to find the optimal size of point surroundings where the local contour has to be calculated. For change detection, I have presented an application to detect structural changes between registered aerial image pairs based on local contour descriptors. Test results confirmed that



(a) (b) (c)

Figure 2: Improved feature map for parametric VFC method. Automatic initial contour and the detection results of the original and the improved method [Thesis 2].

local contour descriptors can be comparable features against compressed descriptors, while the meaningful interpretation can help to design better keypoint descriptors.

2. Thesis: *I have introduced a novel feature map and experimentally confirmed that it can be applied efficiently in the energy term of parametric active contour methods for detecting noisy and high curvature object contours. The introduced method is based on my proposed modification of the traditional Harris detector's characteristic function.*

Related publications: [1], [2], [3], [5], [8], [9], [10], [12], [15].

Deformable active contour (snake) models are efficient tools for object boundary detection. Existing alterations of parametric models have reduced sensitivity to noise, parameters and initial location, but high curvatures and noisy, weakly contrasted boundaries cause difficulties for them.

To address the limitation of initialization and curvature sensitivity, I have investigated the energy minimization process of the active contour theory and introduced a novel feature map for Gradient Vector Flow (GVF) and Vector Field Convolution (VFC) methods applied in the external energy part of the energy function (Eq. 1). The proposed modification is based on the Harris detector's characteristic function and it is able to emphasize high and low curvatures steadily. The experimental results confirmed that the proposed methods outperforms previous active contour models and detects high curvatures more accurately.

2.1. I have shown that the feature points of the introduced map can be applied for accurate object localization and initialization of iterative contour detection. I have improved the method given in the thesis, to handle multiple objects simultaneously by separating point sets representing different objects, adapting the method for multi-target tracking.

Curve initialization is a challenging task, existing representations either take shape information into account or extract the focus area to define the region of interest, but in case of the detection of randomly shaped objects, the initial outline is usually defined with human interaction.

I have used local maxima of the introduced feature map as feature points and generated the convex hull of the point set to initialize a starting curve around the object. I have extended the introduced method to handle multiple objects simultaneously, by separating the feature point using graph methodology.

Tests, aiming to localize small objects in noisy background, showed that the given technique can successfully reach the required goals.

2.2. By combining the feature map with the local contour descriptors, introduced in Thesis 1, I have given a model for detecting structural changes in image pairs scanned with long time difference. I have tested the introduced method on single channel brain MRI image pairs for detecting appearing malignant lesions.

Change detection is a crucial step for monitoring applications, where the changes may refer to important actions. The challenge resides in the altering image characteristics, which makes registration and detection more difficult.

I have given an automatic structural change detection method for long time-span image pairs, using the introduced feature map for robust difference image calculation. The local maxima of the difference image are change keypoint candidates and local contour descriptors are generated in their surroundings to measure the change rate and separate misregistration errors and real changes.

The method has been tested on single channel brain MRI image pairs to focus the radiologist's attention to appearing malignant lesions. Comparison of the introduced method with previous lesion detections on artificial and real images, confirmed the advantages of the proposed model.

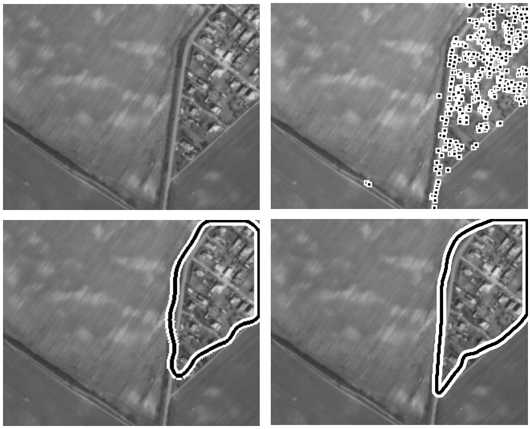


Figure 3: Orientation sensitive urban area extraction with the proposed feature point set. Upper left: Original image; Upper right: Extracted feature point set; Lower left: Detected urban area applying the non-oriented process [28]; Lower right: Detected urban area applying the improved, orientation sensitive process [Thesis 2.3, 3.1].

2.3. Based on the analysis of many airborne images, I have revealed that the proposed feature point set represents built-in areas more precisely, than other point sets extracted by existing feature and corner point detector methods.

Automatic detection of urban areas in optical aerial images means a great support in a wide range of applications, like urban development analysis, map updating, disaster management. I have applied the introduced feature point set for representing urban areas. I have built a probability map based on this point set and performed a decision-making step to identify urban areas. In the experimental part, I have demonstrated that the introduced feature point set enhances the detection accuracy versus other interest point detectors.

3. Thesis: *I have shown that orientation is an efficient feature for urban area characterization in airborne images. I have developed novel, orientation sensitive models for enhancing the localization of built-in areas and detection of building contours without shape templates, by estimating the main directions of the area surrounding feature points.*

Related publications: [3], [11], [13].

3.1. I have developed an orientation sensitive model, by improving the method used in Thesis 2.3. with inserting the orientation of feature point surroundings. By applying the improved model for different feature point extraction methods, I have revealed that using orientation as a feature enhance the accuracy of built-in area detection. I have experimentally shown that the improved model applying the feature point set introduced in Thesis 2.1. combined with the novel feature outperforms previous techniques.

Orientation has an important role when detecting residential areas. Orientation information of feature points is calculated based on the edges of the local neighborhood. I have shown that, inserting this information into the previously used probability model and introducing a novel orientation-sensitive voting map system, increases the accuracy of urban area detection.

Experiments showed that orientation-sensitivity is an efficiently adaptable feature and improves the performance of multiple feature point detector. Tests have also confirmed that the feature point set introduced in Thesis 2.1. fused with the orientation feature obtains the highest detection accuracy compared to previously used point detectors both using and missing orientation information.

3.2. I have given a novel, orientation sensitive model for detecting object contours without shape templates in airborne images. I have shown that by building a statistical model using the orientation information extracted from the surrounding of feature points, the main directions, representing the objects in the image, can be defined and more specific local features can be gained. I have experimentally confirmed the benefits of the proposed approach over previously used building detection methods, either detecting purely location or using shape templates.

A small urban area has buildings with connected orientation, therefore this feature can also be efficient for building detection. In most cases, houses are oriented according to some bigger structure (e. g. the road network), therefore orientation of such structure should be analyzed and used for complex tasks.

I have given an orientation sensitive building detection model. First, I have extracted the orientation information based on the introduced feature point set. Then, main directions representing an urban region are determined with bimodal Gaussian function fitting based on the orientation distribution. Edges in the defined directions are enhanced and for this issue, shearlet based edge detection is used. By fusing point and edge information, building candidates are initialized and active contour is applied to detect accurate boundaries.

The proposed model have been compared to previously used algorithm, giving either purely object location or adapting shape templates (like rectangles). Experiments showed that besides it's technical advantages (no templates and accurate detecting), the performance of the method is better than previous approaches.

4 Application of the results

The introduced algorithms and methods can all be integrated in complex systems for different computer vision tasks, like surveillance and change detection, medical and airborne image analysis, force protection and defense applications.

The selected tasks were corresponding to either ongoing research projects or cooperation with other institutes.

The aim of project No. 76159 of the Hungarian Scientific Research Fund (OTKA) is the analysis of structural information in the space of sensor networks, to measure and extract valuable information which can be used as a feature set for definite problems, like detecting the important changes in a dynamic scene.

OTKA project No. 80352 is about coherent attributes for interpreting the visual world and its perception, by addressing the fundamental problem of automatic extraction of visual information from raw sensory data, giving coherent models for sensory understanding the visual world by investigating the human vision for solving special tasks like the analysis of medical or aerial images.

The aim of the cooperation with MR Research Center, Semmelweis University was to give an automatic support for Multiple Sclerosis lesion detection by focusing the radiologist's attention with providing a list of hypothetical lesions, corresponding to significant, but not anatomical changes.

Flying target detection and recognition is an important task for defense applications. The goal of the Multi Sensor Data Fusion Grid for Urban Situational Awareness (MEDUSA) project of the European Defence Agency was to realize a multi-sensor data fusion grid to improve situational awareness and Command & Control in the context of force protection in the urban environment. MEDUSA has also analyzed, developed and applied algorithms that facilitate usage across a range of different types of sensors, fusing the information obtained from them.

5 Acknowledgements

Due to the electrical engineering family background, I have always been keen on mathematics and computer science. This enthusiasm led me since my childhood, cherished by Sándor Dobos in Fazekas Mihály Secondary School. However, the past years required continuous effort and commitment. This is why I am deeply grateful to my supervisor Tamás Szirányi, for accepting my supervision and supporting me in so many ways during my studies. His guidance and advices meant a great help in progressing with my work, as well as his invitation to join his research laboratory. I would also like to thank my other supervisor, Zoltán Vidnyánszky for his encouraging and giving the opportunity to cooperate with the Semmelweis University.

I am very thankful to Tamás Roska and Péter Szolgay for providing me the opportunity to spend my Ph.D. years at the Doctoral School of Interdisciplinary Sciences and Technology in Pázmány Péter Catholic University (PPCU), and broadening my view in many fields of interest.

The support of the Computer and Automation Research Institute of the Hungarian Academy of Sciences (MTA SZTAKI) is gratefully acknowledged for employing and awarding me with the Young Researchers' Academic Grant to proceed with my research. I thank my present and former colleagues in MTA SZTAKI, especially to István Méder and the members of the Distributed Events Analysis Research Laboratory for helping me with all of their professional and non-professional advices: Mónika Barti, Csaba Benedek, László Havasi, Anita Keszler, Ákos Kiss, Levente Kovács, Zoltán Szlávik and Ákos Utasi.

Thanks to Eszter Nagy, Virág Stossek and Jenő Breyer for arranging all my scientific visits 'to infinity and beyond'.

I thank the support of Zsuzsa Vágó at PPCU in my teaching and the help of PPCU Students' Office, Financial Department and Viktória Sifter from the Library.

Thanks to former and present, older any younger fellow Ph.D. students, especially to Judit Rónai, Petra Hermann, Ádám Balogh, Ádám Fekete, László Füredi, András Gelencsér, Zoltán Kárász, László Kozák, Vilmos Szabó, Kálmán Tornai, Balázs Varga, András Kiss and Dániel Szolgay.

I thank the reviewers of the dissertation for their work and valuable comments.

I would like to thank the financial support of the Hungarian Scientific Research Fund under Grant No. 76159 and 80352. The help of the Hungarian Institute of Geodesy, Cartography and Remote Sensing (FÖMI) and Péter Barsi at the Semmelweis University MR Research Center (MRKK) is acknowledged for providing special images for my research.

I wish to gratefully thank my friends and my family: my grandparents, my brother, my father and especially my mother for giving me a good foundation, teaching me all the qualities and supporting me in all possible ways. I am also thankful to my in-laws I got during my studies, especially to my father-in-law, for his wise and peaceful guidance.

And most of all for my loving, encouraging and patient husband Balázs for believing in me, thank you with all my heart and soul.

6 Publications

6.1 The author's journal publications

[1] **A. Kovacs** and T. Sziranyi, "Harris function based active contour external force for image segmentation," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1180–1187, 2012.

[2] L. Kovacs, **A. Kovacs**, A. Utasi, and T. Sziranyi, "Flying target detection and recognition by feature fusion," *Optical Engineering*, vol. 51, no. 11, pp. 117002–1–13, 2012.

[3] **A. Kovacs** and T. Sziranyi, "Improved Harris feature point set for orientation sensitive urban area detection in aerial images," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 796–800, 2013.

6.2 The author's international conference publications

[4] **A. Kovacs** and T. Sziranyi, "Local contour descriptors around scale-invariant keypoints," in *Proceedings of IEEE International Conference on Image Processing*, (Cairo, Egypt), pp. 1105–1108, 2009.

[5] **A. Kovacs** and T. Sziranyi, "High definition feature map for GVF snake by using Harris function," in *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science 6474*, (Sydney, Australia), pp. 163–172, 2010.

[6] **A. Kovacs** and T. Sziranyi, "New saliency point detection and evaluation methods for finding structural differences in remote sensing images of long time-span samples," in *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science 6475*, (Sydney, Australia), pp. 272–283, 2010.

[7] **A. Kovacs** and T. Sziranyi, "Shape detection of structural changes in long time-span aerial image samples by new saliency methods," in *ISPRS Workshop on Modeling of Optical Airborne and Space Borne Sensors*, vol. XXXVIII-1/W17, (Istanbul, Turkey), 2010.

[8] **A. Kovacs**, C. Benedek, and T. Sziranyi, "A joint approach of building localization and outline extraction," in *IASTED International Conference on Signal Processing and Pattern Recognition*, (Innsbruck, Austria), pp. 721–113, 2011.

[9] **A. Kovacs**, A. Utasi, L. Kovacs, and T. Sziranyi, "Shape and texture fused recognition of flying targets," in *Proceedings of Signal Processing, Sensor Fusion, and Target Recognition XX, at SPIE Defense, Security and Sensing*, vol. 8050, (Orlando, Florida, USA), pp. 80501E–1–12, 2011.

[10] **A. Kovacs** and T. Sziranyi, "Improved force field for vector field convolution method," in *Proceedings of IEEE International Conference on Image Processing*, (Brussels, Belgium), pp. 2853–2856, 2011.

[11] **A. Kovacs** and T. Sziranyi, "Orientation based building outline extraction in aerial images," in *ISPRS Annals of Photogrammetry, Remote Sensing and the Spatial Information Sciences (Proc. ISPRS Congress)*, vol. I-7, (Melbourne, Australia), pp. 141–146, 2012.

[12] **A. Kovacs** and T. Sziranyi, "Automatic detection of structural changes in single channel long time-span brain MRI images using saliency map and active contour methods," in *Proceedings of IEEE International Conference on Image Processing*, (Orlando, Florida, USA), pp. 1265–1268, 2012.

[13] **A. Kovacs** and T. Sziranyi, "Multidirectional building detection in aerial images without shape templates," in *ISPRS Workshop on High-Resolution Earth Imaging for Geospatial Information*, (Hannover, Germany), 2013. accepted.

6.3 The author's other publications

[14] **A. Kovacs** and T. Sziranyi, "Detecting boundaries of structural differences in long time-span image samples for remote sensing images and medical applications," in *8th*

Conference of the Hungarian Association for Image Processing and Pattern Recognition, (Szeged, Hungary), 2011.

[15] **A. Kovacs** and T. Sziranyi, "Új típusú, Harris függvény alapú tulajdonságtérkép és ponthalmaz objektumok körvonalának megkeresésére," in *9th Conference of the Hungarian Association for Image Processing and Pattern Recognition*, (Bakonybél, Hungary), 2013.

[16] A. Utasi and **A. Kovacs**, "Recognizing human actions by using spatio-temporal motion descriptors," in *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science 6475*, (Sydney, Australia), pp. 366–375, 2010.

6.4 Selected publications connected to the dissertation

[17] M. Kass, A. P. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[18] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[19] C. Xu and J. L. Prince, "Gradient vector flow: A new external force for snakes," in *Proceedings of Conference On Com-*

puter Vision and Pattern Recognition, (San Juan, Puerto Rico), pp. 66–71, 1997.

[20] B. Li and T. Acton, "Active contour external force using vector field convolution for image segmentation," *IEEE Trans. Image Processing*, vol. 16, no. 8, pp. 2096–2106, 2007.

[21] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference*, (Manchester, UK), pp. 147–151, 1988.

[22] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[23] B. Jähne, *Digital image processing (5th revised and extended edition)*. Springer-Verlag, 2002.

[24] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[25] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 4, pp. 1156–1167, 2009.

[26] C. Benedek, X. Descombes, and J. Zerubia, "Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics," *IEEE*

Trans. Pattern Anal. Mach. Intell., vol. 34, no. 1, pp. 33–50, 2012.

[27] H. J. Seo and P. Milanfar, "A non-parametric approach to automatic change detection in MRI images of the brain," in *IEEE Int. Symp. on Biomedical Imaging (ISBI)*, (Boston, MA, USA), pp. 245–248, 2009.

[28] B. Sirmaçek and C. Ünsalan, "Urban area detection using local feature points and spatial voting," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 146–150, 2010.

[29] S. Yi, D. Labate, G. R. Easley, and H. Krim, "A shearlet approach to edge analysis and detection," *IEEE Trans. Image Processing*, vol. 18, no. 5, pp. 929–941, 2009.