Determination of geometrical properties of dynamic scenes by using motion- and motion-statistical features in 2D imaging sensors

Theses of the Ph. D. dissertation

László Havasi

Scientific adviser:

Tamás Szirányi, D. Sc.

Péter Pázmány Catholic University

Faculty of Information Technology

Budapest, 2006

New scientific results

1. Thesis: The high-level temporal descriptor of the structural changes of a moving non-rigid (human) object is feasible to detect the human activity (walking) and to determine information about the gait.

I give a new method to detect human activity and to determine a specific feature of walking in video sequence. I have introduced the eigenwalk space which is utilizable to detect the human walking. The proposed method is applicable to identify the leading leg, which is a descriptor of the gait.

The detection of the human activity, namely the walking is possible by classifying the extracted temporal descriptors of object. The general criterion of walking is the moving legs, by detecting this motion the walking is perceptible and the leading leg can be identified from two successive steps.

This information about the scene is suitable not only for the event-level analysis of video sequences but also for the registration of wide-baseline indoor camera configuration, which is a challenging task.

Published in [1][2][3][4][5][6][7]

Detection of human walking using the spatio-temporal patterns of horizontal symmetries.

I worked out a method, which is able to compute the near horizontal symmetry axes. I have defined the symmetry levels, from which the third is characteristic for the presence of two pair near parallel edges (legs). From this fact, the input of the method is the binarized ridge of the edge map instead of the intensity map or object silhouette. I worked out a method for the temporal tracking and processing (time continuous interpolation and dimension reduction) of symmetry segments, which is the basis for the detection of pedestrians from walk patterns. The two classes (walk and non-walk) were separated with a non-linear hyper-plane. This classification step was carried out by using the Support Vector Machine (SVM) in the eigen space of walk patterns which is called eigenwalk space. I have managed to reduce the dimensionality of walk patterns through the linear dimension reduction technique (PCA).

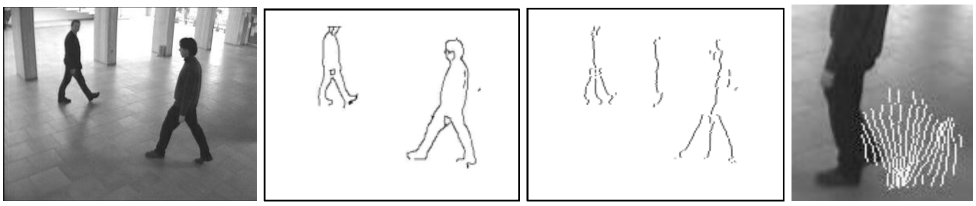

Figure 1: Overview of feature extraction steps: image from input sequence, filtered Canny edge map, third-level symmetries (L3S) and the symmetry pattern.

Identification of the leading leg.

I introduced a method for the identification of leading leg from one walk cycle (two successive steps). The non-rigid human body during a walking cycle has a useful property, which assists us in recognizing the leading leg. Depending on the 3D walk-direction, and on which is currently the leading leg, one leg or the other practically obscures the visible area between the legs. During a walk-cycle the ratio of the visible leg-opening areas, together with the 2D direction on the image-plane, can be used to identify which is the leading leg. *I have shown* experimentally that the method is reliable and accurate. *I summarized* the conditions necessary to the correct functioning.

I have listed the relationship between the leading leg and the ratio of surfaces from two successive patterns in table form. Furthermore the limitations of the method are also named. The leading leg as a discriminative feature is a novel description of the gait.

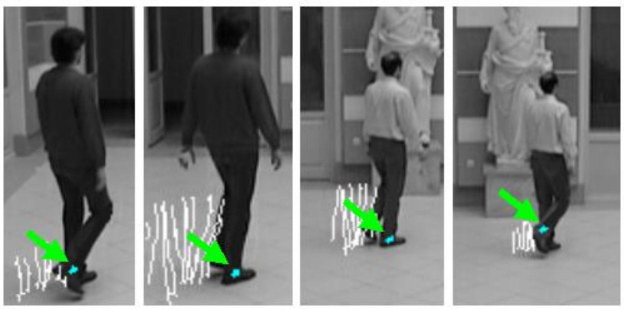

Figure 2: Tracked traces during one cycle comprising two steps. The leading leg can be identified from one walk cycle.

Registration of partially overlapping and non-overlapping views by utilizing the detected walk patterns.

I have shown experimentally that the registration of views can be done by using the spatial position of walk patterns. I have showed that the leading leg is a stronger discriminative feature and the spatial accuracy of walk detection is sufficient for the computation of homographies.

I used known optimization procedures and I have compared the model error of these methods. In case of non-overlapping views I utilized the line homography and corresponding line fragments instead of point-to-point correspondences.

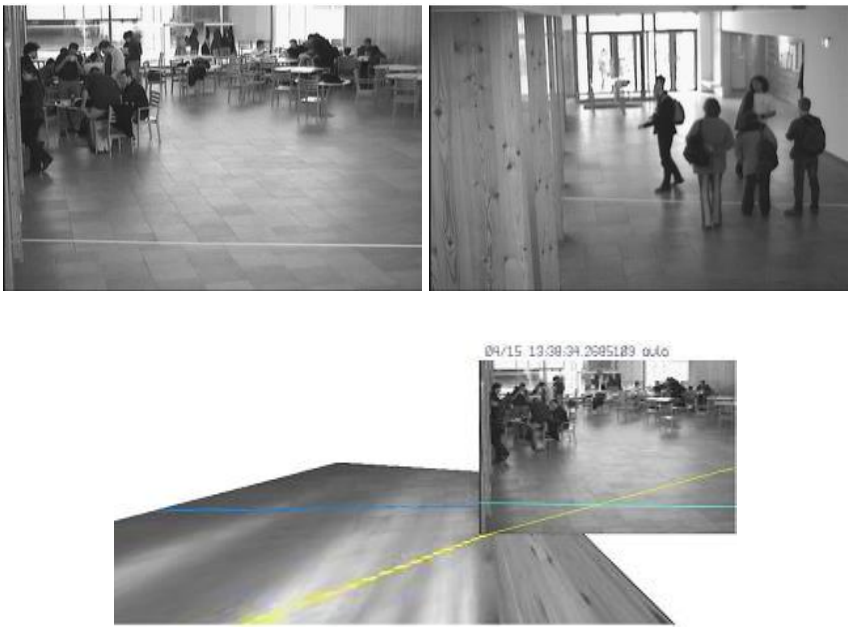

Figure 3: Result of alignment of non-overlapping views with the highlighted control lines by computing line homography.

2. Thesis: The model-based, statistical description of the perceptible changes of scene and environment is applicable to

determine the geometrical model of the plane-mirror, cast shadow and the horizon.

I have shown that the cumulative information that comes from scene changing can be modeled with Gaussian mixture and can be used in geometrical model estimation problems. I worked out a framework for parametrical processing of motion statistics and for the use in different geometrical scene analysis computations.

The detected changes in the camera plane reflect the changes of the dynamic scene and provide information about the position of camera and the geometrical properties of the scene. The vertical planar surface, or shadow casts on the ground-plane occur frequently in surveillance videos (both indoor and outdoor), and they inevitably cause problems in further image-processing steps and reduce the processing system's performance. A specific problem is the determination of horizontal vanishing line (horizon) which describes the relative camera orientation to the world coordinate system. These situations can be viewed as a geometrical optimization problem. To solve this optimization task I retrieved the set of measurements from the parametrical descriptors of motion statistics.

In summary, the investigated local co-motion statistics are feasible for the analysis of camera view in case of unknown environmental conditions.

Published in [8][9][10][11][12]

The model based processing of motion statistics allows determination of spatial features to sub-pixel accuracy.

I have justified experimentally that, the parametrical descriptors of motion statistics can be used for the robust and accurate determination of 2D position information for parameter computation tasks.

Briefly, the co-motion statistics are a numerical estimation of the concurrent motion probability (conditional probability) of different pixels in the camera plane. I investigated the theoretical background of the evaluation of such statistics and I have given the condition of the parametrical description. The analysis supports our empirical confidence in this statistical method.

Both the empirical and the theoretical results confirm that the method is robust and is fairly insensitive to inaccuracy of the motion-mask. Based on our investigations, the length (frame count) of video sequence necessary for the robust extraction of correspondences may be estimated. The parameters in this formula are the estimated motion-intensity (which is a descriptor of the scene dynamics), and the detection error-rate of the motion-detector algorithm.

Using global optimum search method for the determination of geometrical model of camera-mirror scenes and cast shadow.

I proved that the geometrical model of planar mirror and cast shadow can be described with a skew-symmetric (auto epipole) fundamental matrix. This matrix determines a point-to-line transformation and it is formed from the position of vanishing point (2D mirror pole).

I have defined an objective function from the geometrical features and statistical characteristics. The model parameters are the arguments of the objective function in its maximum, thus it leads to a global optimum search task. I have shown experimentally the robustness and accuracy of the proposed approach in both indoor and outdoor environmental conditions.

I have justified experimentally that the sub-pixel accuracy can be achieved. The reduced spatial resolution of input data does not affect the precision of extracted features and the model parameters but enhances the running capabilities of the implementation.

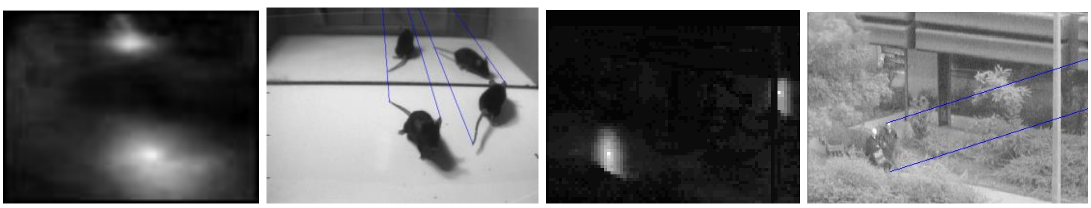

Figure 4: Sample local co-motion statistics and the computed mirror-pole where the collinearities are observable.

Determination of the horizontal vanishing line – which determines the orientation of camera – by using the height information of objects extracted from the motion statistics.

I have introduced a method which based on the statistical error propagation and the measurement transformation into the model parameter space. I showed that the computation of horizon can be originated in the same optimization problem such the previous section but it operates in the parameter space of lines namely in the Hough-space. It assumes that the data representation is continuous instead of the discrate accumulator array in the Hough-space, thus the formula of propagated error is expressed by Gaussian function. *I have shown experimentally* that the estimated height of objects can be used for horizon determination.

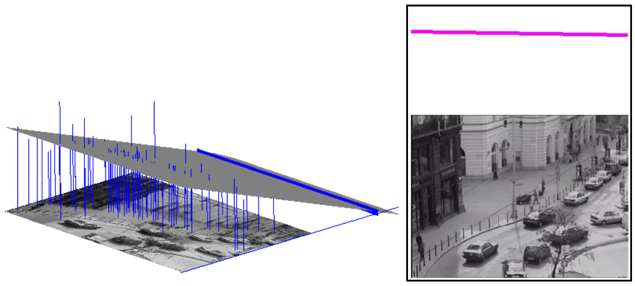

Figure 5: Odd sample from height estimation and the result of horizon computation.

3. Thesis: Use of geometrical model for improved extraction of foreground image mask in case of reflection and cast shadow.

I worked out video segmentation methods with the integration of geometrical information into the decision process. I have shown experimentally that the resulted foreground image mask is more accurate than the mask without using the geometrical knowledge about the scene content.

During the processing of video sequences the basic feature extraction step is the perception of changes and foreground objects. Reflections and cast shadows in surveillance videos usually cause problems in image analysis. This is because they appear in the foreground mask extracted by using an adaptive background model. In turn, the inaccurate mask reduces the performance of the further image-processing steps. Consequently, techniques for the avoidance of such disturbances constitute an active current research area.

I introduced two different methods based on the estimated geometrical models; one for the removal of strong shadow pixels and an other for the removal of reflected pixels related to a valid foreground object, which can lead to better performance.

Published in [13][14][15]

The removal of object's reflection from the foreground image mask based on Bayes decision rule.

I have introduced the integration of the geometrical model and statistics into a foreground-extraction method which is more reliable than previous approaches. Based on the model, only the fundamental constraint can be used, which is a point-line transformation. Because the co-motion statistics store information about the position of concurrent points, the inclusion of the appropriate component of the statistics into a class-conditional density function conveniently solves the identification of the reflection related to an arbitrary point.

I have shown experimentally – using both indoor and outdoor videos – that the detection error rate of foreground segmentation process can be reduced by taking into account the presence of reflective surface and by using the proposed post-processing method.

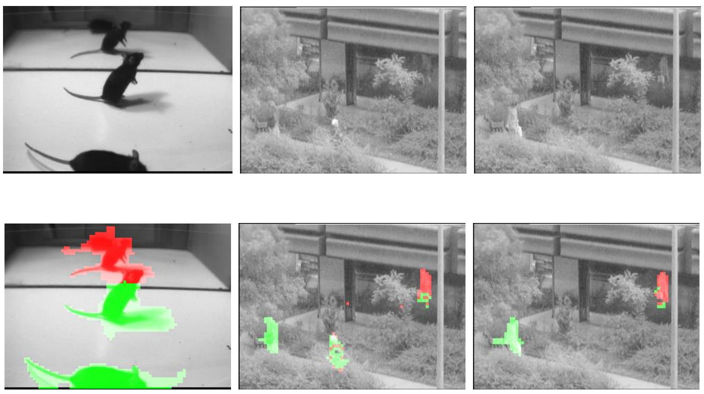

Figure 6: Results of the proposed classification process which supports the removal of reflections from the foreground mask.

The removal of moving cast (strong) shadow from foreground image mask.

I give a novel approach for the identification of cast shadow related to foreground objects. The applied Bayesian iteration scheme is able to handle the a prior information about the object-shadow configuration. The proposed method is capable to remove the cast shadow regions from foreground image mask in case of strong shadow too.

In case of shadow the motion statistics can be used for the estimation of geometrical model only. This lack of spatial information is compensated by using the Bayesian iteration completed with the knowledge of the geometrical model. The proposed method uses the color based shadow segmentation in the initialization stage.

7. ábra: Experimental results on strong shadow. The final foreground mask is the output of the proposed classifier.

Applications of the results

All the developed algorithms and implementations offer solutions for real application problems.

The most important utilization of the methods is their integration into surveillance systems. These systems need algorithm with the capability of real-time functioning and robust operation.

The walk detection introduced in the first thesis is a useful procedure to scene analysis and event detection in both indoor and outdoor configurations. These tasks were implemented in the PPKEyes digital video surveillance system which is operating in the university campus.

In the second thesis the presented geometrical model estimation provides the necessary information for improving the foreground image mask. It can be used to remove the pixels related to a reflective surface. This situation is often occurred in public places. This preprocessing step is important before using the foreground mask in some higher level processing (object detection, feature extraction etc.).

The approaches presented in the last thesis are the applications of the geometrical information determined in the previous thesis. Both classification methods are novel and do not use a prior assumptions.

Acknowledgements

I would like to say thanks

To my professor, prof. Tamás Szirányi, who helped and supported me in many ways.

To prof. Tamás Roska for his support during my study and work.

To my Family.

To my colleagues, with whom I discussed ideas and question: Zoltán Szlávik, Csaba Benedek, István Kóbor and Levente Kovács.

To prof. Michael Rudzsky and Miklós Rásonyi, who provided me with fruitful discussion, ideas and help.

To the PPKE-ITK and the MTA-SzTAKI for the support of my Ph.D. studies.

I would like to acknowledge the financial support of the

Hungarian National Research and Development Program: TeleSense NKFP 2001/02/035, NoE MUSCLE project of the European Union, Economic Competitiveness Operative Programme of Hungary, Video Indexing project, Grant AKF388.

The author's publications

Journal publications

[16] **László Havasi**, Zoltán Szlávik and Tamás Szirányi: "Detection of gait characteristics for scene registration in video surveillance system", *IEEE Transactions on Image Processing,* 2006, in print

[17] **László Havasi** and Tamás Szirányi: "Estimation of vanishing point in camera-mirror scenes using video", *Optics Letters,* 2006

[18] **László Havasi**, Zoltán Szlávik and Tamás Szirányi: "Higher order symmetry for non-linear classification of human walk detection", *Pattern Recognition Letters,* 2006,

[19] Zoltán Szlávik, Tamás Szirányi and **László Havasi**: "Stochastic view registration of overlapping cameras based on arbitrary motion", *IEEE Transactions on Image Processing*, 2006, accepted

[20] Zoltán Szlávik, **László Havasi** and Tamás Szirányi: "Video camera registration using accumulated co-motion maps", *ISPRS Journal of Photogrammetry and Remote Sensing*, 2006, accepted

International conference publications

[21] **László Havasi**, Zoltán Szlávik, Tamás Szirányi: "Use of human motion biometrics for multiple-view registration", ACIVS 2005, Lecture Notes in Computer Science, Antverpen

[22] Zoltán Szlávik, **László Havasi**, Tamás Szirányi, „Estimation of common groundplane based on co-motion statistics", ICIAR 2004, Porto, Portugal, LNCS

[23] **László Havasi** and Tamás Szirányi: „Extraction of horizontal vanishing line using shapes and statistical error propagation", PCV 2006, accepted

[24] **László Havasi** and Tamás Szirányi: „Use of motion statistics for vanishing point estimation in camera-mirror scenes", ICIP 2006, accepted

[25] **László Havasi**, Tamás Szirányi and Michael Rudzsky: „Adding geometrical terms to shadow detection process", EUSIPCO 2006, accepted

[26] Zoltán Szlávik, **László Havasi** and Tamás Szirányi: „Analysis of dynamic scenes by using co-motion statistics", IEEE International Workshop on Visual Surveillance, 2006

[27] Csaba Benedek, **László Havasi**, Tamás Szirányi, Zoltán Szlávik, "Motion-based Flexible Camera Registration", in Proc. of IEEE Advanced Video and Signal-Based Surveillance, AVSS'05, pp. 439-444, 2005

[28] Zoltán Szlávik, **László Havasi**, Tamás Szirányi, Csaba Benedek, "Random motion for camera calibration", European Signal Processing Conference, EUSIPCO, Antalya, 2005

[29] **László Havasi**, Zoltán Szlávik, Tamás Szirányi: „Eigenwalks: walk detection and biometrics from symmetry patterns", ICIP 2005, Genova

[30] Zoltán Szlávik, Tamás Szirányi, **László Havasi**, Csaba Benedek, "Optimizing of searching co-motion point-pairs for statistical camera calibration", IEEE International Conference on Image Processing, ICIP, pp. 1178-1181, Genova, 2005

[31] **László Havasi**, Zoltán Szlávik, Csaba Benedek, Tamás Szirányi, "Learning human motion patterns from symmetries", International Conference on Machine Learning Workshop on Machine Learning for Multimedia, Bonn, 2005, (on CD-ROM)

[32] **László Havasi**, Zoltán Szlávik, Tamás Szirányi: „Pedestrian detection using derived third-order symmetry of legs", ICCVG 2004, Warsaw, Poland, Kluwer, Computational imaging and vision

[33] **László Havasi**, Csaba Benedek, Zoltán Szlávik, Tamás Szirányi: „Extracting structural fragments of overlapping pedestrians", VIIP 2004, Marbella, Spain

[34] **László Havasi**, Zoltán Szlávik, „Symmetry feature extraction and understanding”, CNNA'04

[35] Zoltán Szlávik, **László Havasi**, Tamás Szirányi, „Image matching based on co-motion statistics”, 3DPTV'04, Thessaloniki, Greece

[36] **László Havasi**, Tamás Szirányi: „Motion Tracking Through Grouped Feature Points”, ACIVS, Ghent, Belgium