Evaluating the Trackability of Natural Feature-Point Sets

Lukas Gruber[1], Stefanie Zollmann[2], Daniel Wagner[3] and Dieter Schmalstieg[4]

Graz University of Technology

ABSTRACT

In this work we present a novel idea of evaluating natural feature-point based tracking targets. Our main objective is to evaluate the inherent characteristics of natural feature-point sets with respect to vision-based pose estimation algorithms. Our work attempts to break new ground by concentrating on evaluating complete tracking *targets*, rather than evaluating tracking *methods* or single features. This allows deriving indications on how to improve the trackability of natural feature point sets.

KEYWORDS: Natural Feature Tracking Target Design, Augmented Reality, Tracking Simulation

INDEX TERMS: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems - Artificial, augmented, and Virtual Realities; I.4.8 [Image Processing and Computer Vision]: Tracking, Image processing, Noise measurement;

1 INTRODUCTION

Much work has gone into improving algorithms for detecting feature-points and estimating a 6DOF camera pose from a set of point correspondences, but to our knowledge, no work focused so far on the analysis and improvement of the tracked objects themselves.

Common feature-point based approaches such as SIFT [4], SURF [1] or Ferns [7] work under the assumption that a point can be characterized by its close proximity, called the support area. These methods aim at describing the support area in a way that it is robust to affine transformations, so that a feature-point can be recognized from different viewpoints. Typical descriptors have 64 or even 128 dimensions and can therefore theoretically discriminate large amounts of feature-points. However, in practice repetitive features are common and the resulting similar descriptors lead to ambiguities.

Ambiguous feature-points are often suppressed by monitoring such cases in real-time, rejecting those correspondences that are not unique enough – typically by empirically choosing a threshold. In many situations, an offline analysis of the tracked object would be desirable, since it can give early feedback on the tracking robustness and accuracy that can be expected at runtime. In some applications, the user may even be able to influence the appearance of the tracked object (e.g.: advertisement). The right toolset can enable a user to tune tracking for optimal performance.

2 RELATED WORK

The previous work can roughly be divided into the analysis of feature sets, feature set improvement and the evaluation of descriptors and tracking algorithms.

Shi and Tomasi [8] used feature set analysis for the selection and monitoring of features during the tracking process to optimize the Kanade-Lucas-Tomasi (KLT) tracker. To support feature set analysis Eckmann and Boult [2] introduced new quality measures for features: The spatial-temporal consistency that describes how consistently a feature has been tracked and the distributivity of features. For feature set improvement Knapek et al. [3] explore methods for selecting promising features from an image by testing their methods empirically and simulating the possible changes of a feature window. But all this work is focused on the analysis and improvement of the feature sets instead of the tracked objects themselves.

An evaluation system for local descriptors was presented by Mikolajczyk and Schmid [6], which has become a standard framework in the computer vision community. To evaluate feature detectors and descriptors under more realistic conditions Moreels and Perona [5] used a computer-controlled turntable and real objects to match 3D object features under different viewpoint and lighting conditions. The most existing evaluation systems are either using exclusively simulated data with simple image transformations and lighting changes or real image data, with an expensive production process. With our simulation pipeline we try to provide a way for cost-effective production of more realistic image material.

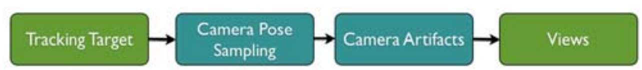

Figure 1. Simulation pipeline.

3 EVALUATION AND IMPROVEMENT METHODOLOGIES

In this paper we introduce two main ideas. The first idea is to create a method for estimating the *trackability* of feature-point sets. We define trackability as the tracking behavior, which a user will experience at runtime. The second aim is to find methods to improve the trackability of such data sets. To achieve these two strongly related goals we use the following methods.

3.1 Evaluation of Trackability

To quickly provide reliable statements about the trackability of a certain feature-point set, we study its inherent characteristics. In our first approach we investigate the following three characteristics of 2D image based natural feature-point data sets.

Spatial Feature Distribution: By analyzing the spatial distribution of a target's feature, we want to determine if a feature set is uniformly distributed over the area covered by the target or if the features are clustered. Our assumption is that well distributed features are more likely to be detected and hence lead to more accurate and robust pose estimations.

Feature Texturedness: The texturedness of a feature provides information how strongly the image intensity in the feature window varies. Whereas Shi and Thomasi [8] used the eigenvalue of the Second Moment Matrix to determine the texturedness of features, we use the standard deviation of intensities in the feature window. Our assumption is that well textured features are more likely to be found than poorly textured features.

[1]lgruber, [2]zollmann, [3]daniel, [4]schmalstieg@icg.tugraz.at

IEEE International Symposium on Mixed and Augmented Reality 2009
Science and Technology Proceedings
19 -22 October, Orlando, Florida, USA
978-1-4244-5419-8/09/$25.00 ©2009 IEEE

Feature Uniqueness: The feature uniqueness estimates how many similar features exist in the same feature set. For each feature descriptor we calculate the Euclidean distances to all other feature descriptors. Pairs with a distance below a certain threshold are treated as similar.

To prove the significance of these heuristics, we target to apply statistical methods, which require an adequate database of test scenarios. Since a manual acquisition of statistically relevant data implies a very high effort, we synthetically generate the required test data. For that purpose, we designed a simulation pipeline (see Figure 1) that allows the creation of a large amount of realistic test data.

Views of feature-point sets are created by exhaustive camera pose sampling. Therefore we not only model the tracked object, but also environmental effects such as lighting and surface properties. In our current approach we support the following light scenarios. To ease setting up the simulation pipeline, we define three different lighting scenarios:

- **Ambient light:** The intensity *I* of the ambient light is varied uniformly over the entire dataset. The same influence is applied to all features equally.
- **One directional light source:** A directional light source illuminates the feature data set. Changing the viewpoint results in different intensities across the feature data set.
- **Point lights:** Point lights are arranged around the feature dataset. Depending on the material, the light intensity and the viewing angle, different amounts of specular highlights are generated, which generate an inhomogeneous lighting situation.

Additionally we emulate camera image artefacts such as camera noise and lens distortion.

In a consecutive step, all views are redirected as input to an arbitrary feature detector and pose estimator. Since the simulation pipeline provides accurate ground truth, we are able to compute the jitter of the pose estimation. These results form the basis for the statistically validation of heuristics.

3.2 Feature-point Set Improvement

The second main goal of our work is the improvement of feature-point sets with respect to their trackability. Feature-point set improvement can be beneficial whenever a tracking target is designed from scratch or allows modification. In the following, we present two concepts which support the improvement and synthesis of natural feature-point sets.

Trackability Visualization: The first concept is to create auxiliary visualizations of the trackability of a certain feature-point set. We call this visualization a footprint of the tracking target. Based on the simulation pipeline mentioned in Section 3.1, we create a rendering of the detection and pose-estimation accuracy of a set of possible views. The final result is a color coded 3D point cloud. Based on this kind of preview, a designer can modify and enhance the natural feature-point data set in an iterative workflow.

Tracking Target Synthesis: Based on the findings of the previously mentioned heuristics (see Section 3.1), we are currently investigating algorithms for automatic enhancement and synthesis of natural feature-point sets. Natural feature-point sets based on 2D images for instance could be modified by computer vision algorithms by adding or removing certain feature-points.

4 Results

Figure 2 shows a short view sequence of 20 pose samples of a planar tracking target generated by our system. To create the different views we translated the virtual camera along hemispheres with different radii. The camera itself pointed towards the center of the tracking target. This strategy allows a systematically sampling of all possible views. To avoid specular highlights we simulated a constant ambient light situation only.

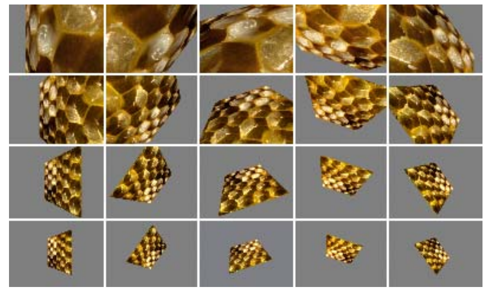

Figure 2. View sequence generated by the simulation pipeline.

5 Conclusion and Future Work

This work introduced a first attempt at simulating tracking performance of objects with natural feature-sets. Ongoing work investigates the following tasks.

Definition and evaluation of heuristics: After specifying a set of heuristics (such as “well distributed features result in robust tracking”) we will analyze them with the help of our simulation framework. In a first step we will synthesize a database for dozens of objects with hundreds of views each. Subsequently we will evaluate which heuristics hold by comparing the expected trackability with the results from the database.

Improve the simulation pipeline: The simulation pipeline will be enhanced by accounting more artifacts caused by the camera processing pipeline, such as contrast and image blur. Furthermore, we plan to perform machine based evaluation of the entire framework to prove the accuracy of the simulation pipeline.

6 Acknowledgements

This project has been funded by the Handheld Augmented Reality - Christian Doppler Laboratory (CDL).

References

[1] Bay, H., Tuytelaars, T., Gool, L.V., SURF: Speeded Up Robust Features, In Proceedings of European Conference on Computer Vision (ECCV), pp. 404-417, 2006

[2] Eckmann, M., Boult, T.E, Spatio-Temporal Consistency and Distributivity as Qualities of Features, In Proceedings of Computer Vision and Pattern Recognition Workshops (CVPRW), 2008

[3] Knapek, M., Oropeza, R., Kriegman, D., Selecting promising landmarks, In Proceedings of International Conference on Robotics and Automation, pp. 3771-3777, 2000

[4] Lowe, D.G., Distinctive Image Features from Scale-Invariant Keypoints, International Journal on Computer Vision, vol. 60, pp. 91-110, 2004

[5] Moreels, P., Perona, P., Evaluation of features detectors and descriptors based on 3D objects, International Conference on Computer Vision (ICCV), pp. 800-807, 2005

[6] Mikolajczyk, K. Schmid, C., A Performance Evaluation of Local Descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, pp. 1615-1630, 2005

[7] Ozuysal, M., Fua, P., Lepetit, P., Fast Keypoint Recognition in Ten Lines of Code, In Proceedings of Computer Vision and Pattern Recognition (CVPR), 2007

[8] Shi, J. Tomasi, C., Good Features to Track, In Proceedings of Computer Vision and Pattern Recognition (CVPR’94), pp. 593-600, 1994