Online Next-Best-View Planning for Accuracy Optimization Using an Extended E-Criterion

Michael Trummer, Christoph Munkelt*, and Joachim Denzler
Friedrich-Schiller University of Jena, Chair for Computer Vision
{*michael.trummer, joachim.denzler*}@*uni-jena.de*

Abstract

Next-best-view (NBV) planning is an important aspect for three-dimensional (3D) reconstruction within controlled environments, such as a camera mounted on a robotic arm. NBV methods aim at a purposive 3D reconstruction sustaining predefined goals and limitations. Up to now, literature mainly presents NBV methods for range sensors, model-based approaches or algorithms that address the reconstruction of a finite set of primitives. For this work, we use an intensity camera without active illumination. We present a novel combined online approach comprising feature tracking, 3D reconstruction, and NBV planning that addresses arbitrary unknown objects. In particular we focus on accuracy optimization based on the reconstruction uncertainty. To this end we introduce an extension of the statistical E-criterion to model directional uncertainty, and we present a closed-form, optimal solution to this NBV planning problem. Our experimental evaluation demonstrates the effectivity of our approach using an absolute error measure.

1 Introduction

Performing 3D reconstruction with a controlled sensor, e. g. a camera mounted on a robot arm, provides additional information and allows purposive actions during the reconstruction process. NBV planning is an important issue, whenever the reconstruction process shall minimize some defined cost or achieve certain reconstruction qualities.

In this paper, we present an online approach for NBV planning using an intensity camera without active illumination (further referred as *passive camera*). All steps of the combined method are calculated at execution time. The reasons for focusing on passive optics are, first, that passive sensing works in cases of intense optical structure where fringe projection systems stumble, and second, the smaller price tag compared to range scanners. Using a passive camera for view planning faces special challenges, since sensing yields just images and no direct 3D information. If mere feature tracking is used while moving to some distant planned position, then one takes a lot of images in-between without actual NBV planning. During this fixed tracking sequence, features might get lost and the planning basis changes, which reduces the value of the planned position. As a solution to these problems we propose to integrate the planning step into the tracking and reconstruction framework of Guided KLT (GKLT) tracking [5]. By this means we achieve a combined online method for feature tracking, 3D reconstruction, and NBV planning for accuracy optimization. We define the extended E-criterion, which gives evidence of where to scan next to minimize the reconstruction uncertainty. Further we use a variable spherical motion model of the sensor which allows to compute the optimal sensor movement w.r.t. the extended E-criterion in closed form. Finally we evaluate our planning method using the benchmark scheme presented in [1].

The remainder of the paper is organized as follows. Section 2 outlines literature relevant to NBV planning. In Sect. 3 we introduce our combined online approach for NBV planning. The results of the experimental evaluation are given in Sect. 4, and Sect. 5 concludes the paper.

2 Literature Review

This section gives a review of previous NBV planning methods.

Pito [7] proposes a discrete positional space for surface acquisition using a range scanner. Preceding Scott [11], he formulates general planning paradigms and considers range scanning physics. Wang and Gupta [12] focus on scene exploration using a range scanner by optimizing C-space entropy. Chen and Li [2] investigate trend surfaces for complete surface reconstruction. Some of the present work includes given 3D models into the planning procedure, thus doing model-based view planning [11, 4]. Earlier, the work of Chaumette et al. [3] explores view planning for special geometric primitives. Wenhardt et al. [10] revisit the statistical E-criterion for NBV planning, but they do not use eigenvectors. For further reading about NBV

*C. M. is with the Fraunhofer Society, Optical Systems, christoph.munkelt@iof.fraunhofer.de

planning, please refer to the book of Chen et al. [9].

In summary, only the approach of Wenhardt is meant for the reconstruction of arbitrary unknown objects using a passive camera. This approach defines optimality based on the E-criterion without using the information from eigenvectors. The planning optimization is carried out as an exhaustive search over all possible camera positions, which is not a classical online scheme.

3 Online Combined Tracking, Reconstruction, and Planning for Accuracy Optimization

In this section we present our combined online approach for NBV planning. We explicitly refer to planning the next step to optimize accuracy. At this point we omit the incorporation of sensor reachability and visibility of the object, for they just constrain the pose space, but do not touch the accuracy criterion itself.

We base our NBV planning on the GKLT tracking framework [5] which performs feature tracking using knowledge about camera parameters. By this means we achieve a combined method for feature tracking, 3D reconstruction, and NBV planning. For each 3D point estimate, we establish its covariance matrix and model the directional uncertainty. We present the extended E-criterion to find the NBV and a closed-form, optimal solution. Our final planning cycle performs online, realtime planning based on the up-to-date information from tracking and reconstruction.

3.1 GKLT Tracking for NBV Planning

Guided KLT tracking, as depicted in [5], performs feature tracking based on the well-known KLT framework, but uses known camera parameters. Additionally, the method provides a robust estimate $\widehat{\mathbf{P}}_s$ of the 3D feature position $\mathbf{P}$ after each tracking step $s > 0$, where $s = 0$ is the index of the feature initialization frame. For a sequence of observations $\langle \widehat{\mathbf{P}}_s \rangle_{s=1,2,\dots,n}$ with $n > 1$, we determine the covariance matrix of the 3D position of feature $\mathbf{P}$, $\mathbf{\Sigma}_{\widehat{\mathbf{P}}_n} = \left(\sigma_{ij}^{\widehat{\mathbf{P}}_n}\right)_{1 \le i,j \le 3}$ with the covariances $\sigma_{ij}^{\widehat{\mathbf{P}}_n}$ of the 3D coordinates. Let

$$\mathbf{\Sigma}_{\widehat{\mathbf{P}}_n} \mathbf{V}_{\widehat{\mathbf{P}}_n} = \mathbf{V}_{\widehat{\mathbf{P}}_n} \mathbf{\Lambda}_{\widehat{\mathbf{P}}_n} = \left(\mathbf{v}_1^{\widehat{\mathbf{P}}_n} \mathbf{v}_2^{\widehat{\mathbf{P}}_n} \mathbf{v}_3^{\widehat{\mathbf{P}}_n} \right) \begin{pmatrix} \lambda_1^{\widehat{\mathbf{P}}_n} & 0 & 0 \\ 0 & \lambda_2^{\widehat{\mathbf{P}}_n} & 0 \\ 0 & 0 & \lambda_3^{\widehat{\mathbf{P}}_n} \end{pmatrix} \tag{1}$$

denote the eigen decomposition of the matrix $\mathbf{\Sigma}_{\widehat{\mathbf{P}}_n}$ with eigenvalues $\lambda_1^{\widehat{\mathbf{P}}_n} \ge \lambda_2^{\widehat{\mathbf{P}}_n} \ge \lambda_3^{\widehat{\mathbf{P}}_n}$ and corresponding perpendicular eigenvectors $\mathbf{v}_i^{\widehat{\mathbf{P}}_n}$ $(i = 1, 2, 3)$. From this point, we omit the point index $\widehat{\mathbf{P}}_n$, if there it is unquestionable which point estimation is addressed.

3.2 Camera Movement

For the illustration of our solution to the NBV problem, we need to describe our model for camera movement. As other authors have done before, we use a spherical motion model. That means that the optical center of the camera is situated on a sphere $\mathcal{S}$ with center $\mathbf{C}$ and radius r, and the optical axis points at $\mathbf{C}$. Two angles, azimuth α and height β, define the camera position $\mathbf{c}$ on the sphere. We allow changes of sphere size and position, which makes the model a dynamical sphere model. In this formulation, the model can describe any position and angle of view of the camera with a normalized roll angle.

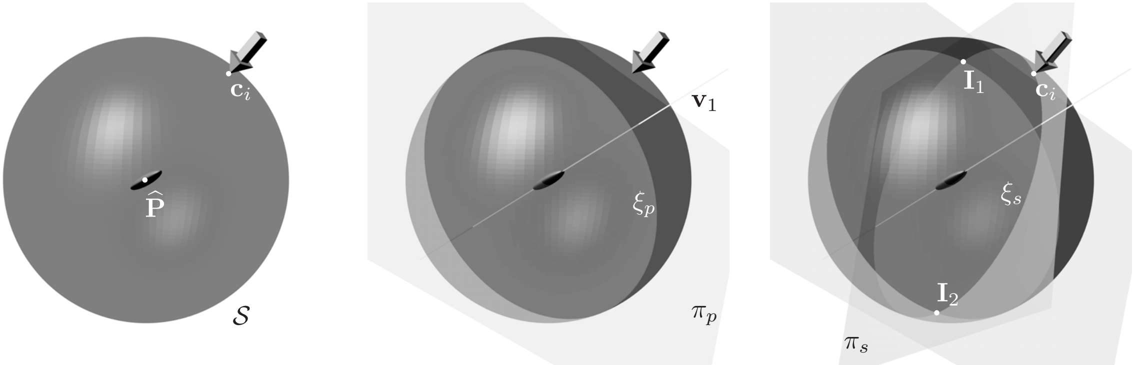
3.3 The Extended E-Criterion and a Closed-Form, Optimal Solution

Having a 3D point estimate $\widehat{\mathbf{P}}$ and a corresponding covariance matrix $\mathbf{\Sigma}$, we visualize directional uncertainty by assuming normal distribution. In this case, $\mathbf{\Sigma}$ as in (1) defines an equiprobable curve as an ellipsoid with semiaxes along $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. The lengths of the semiaxes are $\lambda_1 \ge \lambda_2 \ge \lambda_3$ up to global scale.

EXTENDED E-CRITERION (EEC) Using the notation from above, we can formulate the extended E-criterion as sensing perpendicular to $\mathbf{v}_1$, which is the direction of the largest uncertainty. As the known E-criterion from statistics [8], this formulation minimizes the largest eigenvalue λ_1, but additionally incorporates the corresponding eigenvector $\mathbf{v}_1$.

Regarding the motion model described in Sect. 3.2, we are able to compute closed-form camera motions for the NBV that are optimal w.r.t. EEC. Basically, we determine the motion direction for the camera, since the step size is constrained by the feature tracker used. Figure 1 illustrates the following geometrical considerations.

First, let us assume a motion sphere with fixed size and fixed center $\mathbf{C} = \widehat{\mathbf{P}}$. That means that we want to know how to adapt only the position parameters α_i and β_i (corresponding to $\mathbf{c}_i$ in Fig. 1), such that the camera takes the shortest path to a position, from where its optical axis is perpendicular to the direction $\mathbf{v}_1$ of the largest uncertainty of $\widehat{\mathbf{P}}$. All camera positions on the sphere fulfilling this condition are situated on a great circle ξ_p of the sphere surface. The great circle ξ_p is given by intersecting the sphere with the plane $\pi_p(\mathbf{X}) : \mathbf{v}_1^{\mathrm{T}}(\mathbf{X} - \widehat{\mathbf{P}}) = 0$, i. e. the plane through $\widehat{\mathbf{P}}$ with normal vector $\mathbf{v}_1$. The shortest way to reach ξ_p, starting at the current camera position $\mathbf{c}_i$, leads along another great circle ξ_s that is perpendicular to ξ_p. The great circle ξ_s lies on the plane $\pi_s(\mathbf{X}) : \mathbf{n}_s^{\mathrm{T}}(\mathbf{X} - \widehat{\mathbf{P}}) = 0$ defined by the two points $\widehat{\mathbf{P}}$ and $\mathbf{c}_i$ and the normal direction perpendicular to $\mathbf{v}_1$. Intersecting the great circles ξ_p and ξ_s



(a) Camera position $\mathbf{c}_i$ lies on sphere $\mathcal{S}$. Point estimate $\widehat{\mathbf{P}} = \mathbf{C}$ is shown together with its covariance ellipsoid.

(b) Vector $\mathbf{v}_1$ indicates direction of largest uncertainty of $\widehat{\mathbf{P}}$. Plane π_p through $\widehat{\mathbf{P}} = \mathbf{C}$ is perpendicular to $\mathbf{v}_1$. Intersection of π_p and $\mathcal{S}$ yields great circle ξ_p on $\mathcal{S}$.

(c) Shortest way from $\mathbf{c}_i$ to ξ_p leads along great circle ξ_s on π_s perpendicular to ξ_p. Intersecting the great circles ξ_s and ξ_p provides points $\mathbf{I}_1, \mathbf{I}_2$. Point $\mathbf{I}_1$ is closer to $\mathbf{c}_i$. Camera is moved on ξ_s towards $\mathbf{I}_1$.

Figure 1. Computing the optimal camera movement w.r.t. EEC. For details on the illustrated geometry, please refer to Sect. 3.3.

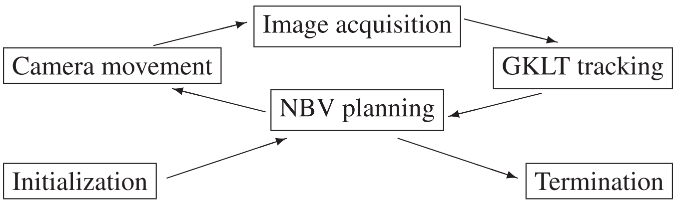


Figure 2. The NBV planning cycle as used for this work.

yields two intersection points $\mathbf{I}_1$, $\mathbf{I}_2$. W.l.o.g. let $\mathbf{I}_1$ be the intersection point being closer to $\mathbf{c}_i$. Since we use feature tracking and cannot jump directly to $\mathbf{I}_1$, we apply a predefined step size to optimally move the camera on ξ_s towards $\mathbf{I}_1$.

Second, we consider the dynamical sphere model for camera motion which provides variable sphere position $\mathbf{C}$ and radius r. Further we drop the assumption that the point estimate $\widehat{\mathbf{P}}$ is situated in $\mathbf{C}$. It is important to notice that these positional aspects do not influence the optimal viewing direction. As a consequence, the optimal camera movement on the sphere described above still holds. The only matter of the dynamic model is to move the camera center $\mathbf{C}$ towards $\widehat{\mathbf{P}}$. Again, the step size of this movement is constrained by the feature tracker.

3.4 The Planning Cycle

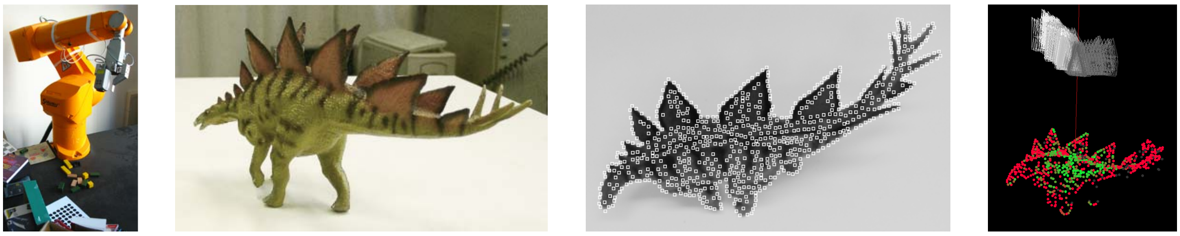
Up to now, we concerned optimizing the camera motion for estimating one point. In the following we want to give an overview over our planning cycle and mention a strategies for optimizing the accuracy for more than one point.

As illustrated in Fig. 2, the planning procedure demands an initialization, since for each point the covariance matrix is necessary to compute the directional uncertainty. So the initialization consists of taking at least three images from random positions for feature detection and two tracking steps to get two 3D estimates and, thus, a covariance matrix for each point. The actual NBV planning step computes the optimal camera motion as depicted in Sect. 3.3. After moving the camera and taking the next image, GKLT performs feature tracking and 3D reconstruction, which leads to updated covariance matrices and a new planning step. Depending on the planning goals, the cycle may terminate, for instance, for small error changes or a maximum number of views.

For more than one point a suitable planning strategy defines the regarded direction of uncertainty. Our current strategy is to consider the worst point only, which we define as $\mathbf{P}^* = \arg\max_{\mathbf{P}} \operatorname{trace}(\mathbf{\Lambda}_{\mathbf{P}})$. Then, the optimization only depends on the direction $\mathbf{v}_1^*$ of the currently worst estimated point. Nonetheless, the estimations of all other points being tracked are also improved.

4 Experimental Results

In the following we evaluate the results of our combined online approach for view planning. As proposed in [1], we assess the reconstruction accuracy by computing the distances between reconstructed points and a registered ground-truth model of the object. Here, the ground-truth is given as a high-accuracy 3D reconstruction of the dinosaur model by using a fringe-projection system [6], which produces a measurement error smaller than $70\mu m$. We perform the planning cycle from Fig. 2 and plot the median error distance of all reconstructed points after each view. Starting from



(a) Robotic arm and camera.

(b) Dinosaur figurine for 3D reconstruction.

(c) 700 features for GKLT tracking.

(d) View of the resulting 3D reconstruction and camera positions.

Figure 3. Experimental setup.

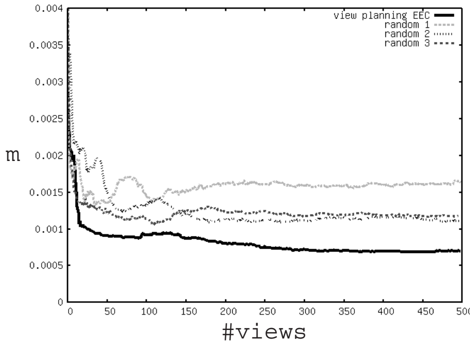


Figure 4. Median error distances after each view of one planned and multiple random sequences. Planning w.r.t. EEC clearly reduces the reconstruction error.

the same camera position and using the same 700 initial features, we repeat the measurement with random camera movements. The experimental setup is illustrated in Fig. 3. We restrict our planning procedure to 500 views including five views for initialization.

The error plot in Fig. 4 shows that the reconstruction accuracy benefits from NBV planning w.r.t. EEC. Instantly after the initialization phase, the planned camera movements produce lower reconstruction errors than all random movements tested. After 500 views, the median reconstruction error with planning is 0.697 mm and 1.649 mm / 1.118 mm / 1.166 mm without planning, cf. Fig. 3(d) and Fig. 4; error mean and standard deviation are $(0.975, 0.917)$ mm with planning and $(3.494, 12.730)$ mm / $(4.906, 26.326)$ mm / $(2.933, 12.801)$ mm without planning. The runtime of one planning cycle is about 5 s, including movement planning, robot movement, image acquisition, and GKLT feature tracking. At least 75% of that runtime are due to network traffic, robot movement, and image acquisition. Using the setup described above, the actual planning is done in a few ms on a Core2 Duo 2.4 GHz with 4 GB RAM.

5 Conclusion

In this paper we presented a combined online approach to next-best-view planning for optimizing the accuracy of the 3D reconstruction of arbitrary objects. Since we used a passive camera, we integrated the planning procedure into the GKLT tracking framework which itself is designed for feature tracking and 3D reconstruction inside a controlled environment. We formulated the extended E-Criterion for accuracy optimization and presented an optimal solution for the camera movement. Our experimental evaluation gave evidence that the camera movements using our planning approach endow a clearly more accurate 3D reconstruction than without planning. Future work should deal with the evaluation of different strategies on computing the direction of uncertainty. A further promising planning aspect is integrating surface estimation to handle occlusions.

References

[1] C. Munkelt et al. Benchmarking 3D Reconstructions from Next Best View Planning. *MVA*, 1:552–555, 2007.

[2] S. Y. Chen and Y. F. Li. Vision Sensor Planning for 3D Model Acquisition. *IEEE SMC – B*, 35(4):1–12, 2005.

[3] F. Chaumette et al. Structure From Controlled Motion. *IEEE PAMI*, 18:492–504, 1996.

[4] F. Prieto et al. CAD-based range sensor placement for optimum 3D data acquisition. *Proc. of Int. Conf. on 3-D Digital Imaging and Modeling*, 1:128–137, 1999.

[5] M. Trummer et al. Combined GKLT Feature Tracking and Reconstruction for Next Best View Planning. *Proc. of DAGM, LNCS, Springer*, 5748:161–170, 2009.

[6] P. Kuehmstedt et al. 3D shape measurement with phase correlation based fringe projection. In *Optical Measurement Systems for Industrial Inspection V*, volume 6616, page 66160B. SPIE, 2007.

[7] R. Pito. A Solution to the Next Best View Problem for Automated Surface Acquisition. *IEEE PAMI*, 21(10):1016–1030, October 1999.

[8] F. Pukelsheim. *Optimal Design of Experiments*. John Whiley & Sons Inc., New York, 1993.

[9] S. Chen et al., editor. *Active Sensor Planning for Multiview Vision Tasks*. Springer Berlin Heidelberg, 2008.

[10] S. Wenhardt et al. An information theoretic approach for nbv planning in 3-d reconstruction. In *ICPR*, 2006.

[11] W. R. Scott. Model-based View Planning. *MVA*, 20:47–69, 2007.

[12] P. Wang and K. Gupta. View planning for exploration via maximal C-space entropy reduction for robot mounted range sensors. *Advanced Robotics*, 21(7):771–792, 2007.