2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

Application of Particle Filters to Vision-Based Orientation Estimation using Harmonic Analysis

Timo Schairer, Benjamin Huhle and Wolfgang Straßer

Abstract— **We investigate the application of particle filters to estimate the orientation of a mobile agent based on an omnidirectional video stream. By applying spherical signal analysis to sequences of low resolution input images we perform a real-time estimation of the relative camera rotation. We use normalized cross-correlation computed with a fast frequency domain approach that yields unbiased estimates which can be further processed by a particle filter. We discuss the quaternion representation of the rotational state space and evaluate methods for meaningful averaging. A prototype system is presented and experiments with real image sequences show that robust estimation of the rotational motion is achievable even when the input images are corrupted, *e.g.* due to occlusions.**

I. INTRODUCTION

The problem of estimating the relative motion of a camera between a pair of images has been studied extensively over the last years. While the approaches differ greatly in the way translation and rotation are estimated, most algorithms have been typically developed for conventional perspective images and were later adapted to different types, such as panoramic images.

The larger information content inherently present in images with a large field of view (FOV) has made omnidirectional vision become popular for different applications, *e.g.* in robot vision and scene acquisition, where spherical images alleviate the recovery of ego-motion [1] or provide for stable position referencing for mobile navigation [2]. When we constrain the motion to 3D rotations one could even consider spherical images the perfect device to reconstruct this motion since image content remains the same and the rotation equals a shift operation. Thus, estimating this shift is a key technique, *e.g.*, for ego-motion estimation or localization that can be performed robustly on omnidirectional images. Additionally, correct rotation estimation is not only important for further processing of the data, but also for visualization purposes (*e.g.* stabilizing the video of a tele-operated mobile agent).

Many applications for orientation estimation that have been developed are limited to three degrees of freedom (DOF) due to the approximation that navigation of ground-based robots is limited to planar surfaces. The more exact estimation of all six DOF is especially important with the advent of unmanned air vehicles (UAVs). Here, for orientation estimation, data from inertia measurement units (IMUs) is often used (see *e.g.* [3] for further references). However, when imaging devices are present on the agent and the data can be processed in real-time, these sensors can also be used to estimate the orientation.

The task of estimating the relative rotation from image data is typically done by detecting local salient features in the template and signal image. Examining the correspondences between these features allows for an estimation of quite large camera motions (*e.g.* using *SIFT* features). An application on panoramic images was presented by Fiala [4]. Alternatively, the optical flow can be computed to extract the motion parameters on a global scale subject to sufficiently small motion. Only recently, a technique was presented where motion estimates are computed from the optical flow at a single pair of antipodal points on the image sphere [5].

The fact that panoramic images can be mapped onto the unit sphere allows for the use of spherical signal analysis. Similar to estimating the translational movement in 2D images ([6]) we estimate rotation by calculating the correlation in spherical Fourier space. In contrast to feature-based approaches, at least for normalized signals, correlation is a very natural and meaningful similarity measure. Taking into account the whole image assures a robust estimation even in difficult situations, such as dynamic environments and low resolution or unfocused/blurred input images. The spherical harmonics approach to orientation estimation was originally introduced by Kostelec et al. [7] and several analysis and application papers followed during the last years, *e.g.* [8], [9], [10]. Friedrich et al. presented an application of this technique to the problem of robot localization [11] limited to rotations on the plane.

To predict or smooth orientation estimates over time, the Kalman filter is very popular. When dealing with image data, however, we are interested in more robust filters since in real world environments, due to occlusions or abruptly changing lighting conditions for example, single estimates can completely fail. We therefore investigate the application of a particle filter to the task of orientation estimation.

After introducing the necessary transforms from spherical signal analysis (Sec. II) and the representation of rotations that we use in the filter (Sec. III), we discuss the issues of particle filters applied to rotation estimation in Section IV. An evaluation on real image sequences is given in Section V.

II. CORRELATION ON THE SPHERE

A. Spherical Harmonics

Images captured with a catadioptric sensor, or, as in our case, with a spherical camera system can be mapped onto a sphere given the intrinsic parameters of the camera system. Therefore, these omnidirectional images can be considered

The authors are with the Department of Graphical Interactive Systems WSI/GRIS, University of Tübingen, Germany.
{schairer,huhle,strasser}@gris.uni-tuebingen.de

a function $f(\theta, \phi) = f(\omega)$ on the 2-sphere where the angle θ denotes the colatitude angle in the range $[0, \pi]$ and ϕ the azimuth defined in $[0, 2\pi)$. Driscoll and Healy [12] showed that the spherical harmonic functions Y_l^m form a complete orthonormal basis over the unit sphere and any square-integrable function $f(\omega) \in L^2(S^2)$ can be expanded as a linear combination of spherical harmonic functions (Spherical Fourier Transform, SFT)

$$f(\omega) = \sum_{l \in \mathbb{N}} \sum_{|m| \leq l} \hat{f}_l^m Y_l^m(\omega). \quad (1)$$

Here, Y_l^m denotes a spherical harmonic function of degree l and order m and is given by

$$Y_l^m(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\theta) e^{im\phi}, \quad (2)$$

where $\hat{f}_l^m \in \mathbb{C}$ are the complex expansion coefficients and $P_l^m(\cos\theta)$ denote the associated Legendre polynomials. Our input data, the spherical functions $f(\omega)$, are defined on a uniformly sampled equiangular grid. A perfect reconstruction from a $2B \times 2B$ grid is possible when bandlimiting $f(\omega)$ to B.

B. Fourier Transform on $SO(3)$

Similar to the phase correlation method on planar images, Kostelec and Rockmore [7] presented a method to estimate the alignment of images defined on the sphere using cross-correlation as similarity measure. They showed that the correlation between two images g and h as a function

$$C(R) = \int_{S^2} g(\omega)\, \Lambda(R) h(\omega)\, d\omega \quad (3)$$

of rotations can efficiently be evaluated in the Fourier domain. Here, $\Lambda(R)$ denotes the rotation operator corresponding to the rotation $R = R(\alpha, \beta, \gamma)$ where α, β, γ are the Euler angles (in *ZYZ* representation) defining the rotation. Further, the spherical harmonic functions Y_l^m form an orthonormal basis for the representations of $SO(3)$ and the $SO(3)$ Fourier transform (*SOFT*) coefficients of the correlation of two spherical functions can be obtained directly by calculating the bandwise outer product (denoted by $\diamond$) of their individual SFT coefficients. Taking the inverse *SOFT*

$$C(R) = \text{SOFT}^{-1}\left(\hat{g} \diamond (\hat{h})^\star\right), \quad (4)$$

where $(\hat{h})^\star$ denotes the complex conjugate of $\hat{h}$, yields the correlation $C(R)$ evaluated on the $2B \times 2B \times 2B$ grid of Euler angles G and its maximum value ideally indicates the rotation separating the two images. The accuracy of the rotation estimate $\tilde{R} = \arg\max_{(\alpha,\beta,\gamma) \in G} C(R(\alpha, \beta, \gamma))$ is directly related to the resolution of the likelihood grid which in turn is specified by the number of bands used in the SFT. Given images of bandwidth B, the resolution of the likelihood grid implicates an inaccuracy of up to $\pm(\frac{180}{2B})^\circ$ in α and γ, and $\pm(\frac{90}{2B})^\circ$ in β. The cubic computational cost when evaluating the grid, in practice, restricts this method to bandwidths in the order of $B = 256$.

Makadia and colleagues [8] showed that it is also possible to compute the same correlation function with a three-dimensional discrete Fourier transform (3DFT).

C. Spatially Normalized Cross-Correlation

When acquiring omnidirectional images, in practice, the sensors do not cover the whole sphere and the images have limited support. It is therefore necessary to normalize the correlation function with regard to the intersection of the support of both image functions, resulting in

$$NC(R) = \frac{\int_{S^2} f'_W(\omega)\, [\Lambda(R) p'_W(\omega)]\, d\omega}{\sqrt{\int_W [f'_W(\omega)]^2\, d\omega \int_W [p'_W(\omega)]^2\, d\omega}}, \quad (5)$$

with

$$f'_W(\omega) = f(\omega) - \overline{f_W} \quad (6)$$

$$p'_W(\omega) = p(\omega) - \overline{p_W}\,, \quad (7)$$

where $W = W(R)$ denotes the intersection of the support of both functions and $\overline{f_W}$ and $\overline{p_W}$ their respective means with regard to W. Note that when the term *normalized cross-correlation* is used, it sometimes only refers to a normalization with regard to the global energy of both functions (*e.g.*in [8]). In our previous work [13] we showed that the the function $NC(R)$ can be expanded in terms of simple correlations and therefore can be computed with multiple applications of the inverse *SOFT* transform. Introducing two mask functions that indicate the support of both signals, the NCC grid ($NC(R)\ \forall\ R \in G$) can be computed from six cross-correlation evaluations involving the images, their squared versions and the mask functions. In the remainder of the paper we use this function as the similarity measure to estimate the orientation.

D. Refinement of the Grid-based Estimates

As described in section II-B, the accuracy of the rotation estimates depends on the bandwidth of the involved Fourier transforms. However, the bandwidth of the SFT, in theory, does not determine the accuracy of the following rotation estimation. Also the low-frequency components carry enough information of the relative rotation, only limited by the quantization of the coefficients. Fehr and colleagues [10] therefore proposed their so-called *zero-padding*, where the bandwidth of the inverse Fourier transform (they use a 3DFT-approach instead of SOFT; see above) is enhanced by padding the outer product of the SFT coefficient vectors with zeros. However, the computation time of the inverse Fourier transform still increases with $O(N^3 \log^2(N))$ and limits this technique. A different way to refine the orientation estimates is averaging the rotations, *i.e.* interpolation in the spatial domain. We incorporate this when determining the final output of the particle filter (see Section IV).

Makadia et al. proposed to use a direct non-linear optimization method starting at the grid cell of maximum correlation [8]. Whereas they minimize the difference (SSD) between the SFT coefficient of the reference image and

the rotated pattern, we use our spatially normalized cross-correlation function as cost function. The maximum of the resulting function can be computed with standard optimization methods (we use gradient descent with line search). Following an idea by Makadia et al. [8], the rotation can be divided into two parts:

$$R(\alpha,\beta,\gamma) = R_2(\beta+\pi,\frac{\pi}{2},\gamma+\frac{\pi}{2})R_1(\alpha+\frac{\pi}{2},\frac{\pi}{2},0). \quad (8)$$

This way, the computational cost is significantly reduced since the argument of the Wigner-d matrix $d^l_{km}(\beta)$ is constant and it has to be evaluated for $\beta=\frac{\pi}{2}$ only:

$$f(R) = \sum_l \sum_{m=-l}^{l} \sum_{p=-l}^{l} \sum_{k=-l}^{l} \hat{g}^m_l \, (\hat{h}_{pl})^\star \, d^l_{pk}\left(\frac{\pi}{2}\right) \, d^l_{km}\left(\frac{\pi}{2}\right) \exp\left(-i(p\gamma' + k\beta' + m\alpha')\right), \quad (9)$$

where $\alpha' = \alpha+\frac{\pi}{2}, \beta' = \beta+\pi, \gamma' = \gamma+\frac{\pi}{2}$.

III. THE ROTATION GROUP

Rotations in 3D space can be represented in various forms. By far the most popular parametrization is in terms of Euler angles. However, not only the well-known gimbal locks are problematic, this representation also suffers from other shortcomings: Hornegger and Tomasi [14] introduced the terminology of a *fair representation* which describes the sensitivity of a transform with regard to changes in the parameters. In contrast to the Euler angles, the unit quaternion representation of rotations is *fair* following this definition, as well as the combined axis-angle (a.k.a. rotation vector [15]) representation. The latter one is also a good choice when it comes to interpolation or averaging of rotation estimates. To gain a better understanding, we introduce this three-vector representation from a different point of view and provide the formula for correct weighted averaging. For a more thorough treatise see for example [16]. The unit quaternions, representing the group of rotations $SO(3)$ in 3D, lie on a smooth three-dimensional manifold in 4D space, namely on the 3-sphere S^3. Since also their multiplication and inversion are smooth maps, the unit quaternions form a Lie group. To every Lie group one can associate a Lie algebra with the underlying vector space being the tangent space to the manifold at the identity element of the group. In the case of the unit quaternions we apply the *log map* (the *exp map* being its inverse) and obtain a 3D tangential space that preserves all distances going through the origin. This tangential space, also called the *exponential chart*, is identical (up to a scaling difference) to the rotation vector representation, mentioned earlier.

Kim and colleagues [16] showed that a correct blending of quaternion curves is possible in this space and the resulting formula is indeed identical to Shoemake's *slerp* operation [17]. The same technique can be applied in order to calculate the weighted mean $\overline{q}$ of a set of rotations M [18]. Let $*$ denote the quaternion multiplication. We iteratively compute

$$\overline{q}^{(j+1)} = \overline{q}^{(j)} * expMap\left(err_t^{(j)}\right), \quad (10)$$

where the actual averaging occurs in the exponential chart (centered at $\overline{q}^j$):

$$err^{(j)} = \frac{1}{\sum_{i\in M} w^i} \sum_{i\in M} w^i \; logMap\left(q^i * (\overline{q}^{(j)})^{-1}\right). \quad (11)$$

Convergence is achieved when the norm of err^j is negligible.

Note that a quite important detail regarding the implementation is to perform a consistency check of the set M prior to every iteration. Since the unit quaternions cover the space of rotations twice ($q=(\alpha,\vec{v})$ is equivalent to $-q=(-\alpha,-\vec{v})$), we have to ensure that all quaternions belong to the same halfspace of S^3. This is easily done by substituting q^i with $sign(q^i \bullet \overline{q}^{(j)})\, q^i$ (where $\bullet$ denotes the dot product).

IV. PARTICLE FILTERING ROTATIONS

Particle filtering is a robust and versatile technique for estimating system states in a probabilistic framework. It can be adapted to a large number of applications, *e.g.* to robot localization or visual tracking. Particle filters perform well even in presence of non-Gaussian noise and can handle multiple hypotheses.

Following the classification and notation of [19], we apply a *sampling importance resampling* (SIR) particle filter to the problem of rotation estimation. For the model underlying the filter we assume a smooth rotational movement of the agent which is characterized by a slowly changing angular momentum. In the following, we address the issues that arise when applying a generic SIR particle filter to this problem.

Defining x_t as the state vector at time t, the fundamental idea of particle filtering is to approximate the probability density function (pdf) over x_t by a weighted sample set $S_t = \{s^i_t | i = 1,\ldots,N\}$. The i'th sample (particle) at time t is represented by $s^i_t = (x^i_t, w^i_t)$ with state vector x^i_t and weight w^i_t. In our case, the state of each particle is defined as $x^i_t = [q^i_t, \Delta q^i_t]$, where q^i_t is a unit quaternion representing the rotation in the global reference frame and Δq^i_t denotes the angular momentum in quaternion representation of particle i at time t.

In a standard "kidnapped robot" scenario, the particles should initially be distributed uniformly in the state space. However, since in our case the NCC grid can be thought of as a quantized yet global solution, we spawn the N particles at the rotations corresponding to the N grid cells with highest correlation value and set the angular momentum to zero.

We apply the straight forward *multinomial approach* to resampling [20] that has been shown to be comparable to other techniques such as residual, stratified and systematic resampling [21].

To simulate the process noise in the prediction step we apply noise from a three dimensional normal distribution with zero mean and standard deviation σ_{vel} to the angular momentum in the exponential chart

$$\Delta q^i_t = expMap\left(\mathcal{N}([0,0,0]^T, \sigma_{vel})\right) * expMap\left(0.5 \; logMap\left(\Delta q^i_{t-1}\right)\right). \quad (12)$$

Finally, the rotation is updated as $q^i_{t+1} = \Delta q^i_t * q^i_t$.

We implemented two different approaches for weighting the particles in the measurement (sensing) step. In the first variant, the NCC grid G is evaluated in Fourier space for all possible Euler angles (see Sec. II-B). For each particle the rotation state representation q_t^i is converted to a ZYZ Euler sequence and weighted according to the correlation value c_t^i stored in the associated grid cell. While this approach suffers from the quantization, it can be evaluated efficiently for a large number of particles. In the alternative variant the NCC is evaluated in Fourier space according to q_t^i for each particle explicitly. This way, the correlation values are free from quantization artifacts, however, the computational cost increases with the number of particles.

Additionally, we investigated the effect of rating also the angular momentum as part of the state of the particle. This is achieved by weighting with a Gaussian kernel applied to the difference of the particle's angular momentum Δq_t^i and the relative rotation of the refined rotation estimate from the previous time step $\tilde{q}_{t-1}$ to the one of the current time step $\tilde{q}_t$:

$$\delta_t^i = \exp\left(-\frac{\left|\left|logMap\left(\Delta q_t^i\right) - logMap\left(\tilde{q}_t * (\tilde{q}_{t-1})^{-1}\right)\right|\right|_2^2}{\sigma_L}\right) \tag{13}$$

The influence of the angular momentum in the sensing step can be controlled with the parameter σ_L. The final weight is then computed as $w_t^i = w_{t-1}^i\, \delta_t^i\, c_t^i$.

Once the particles have been rated, there are several options to find the most likely estimate that should be output by the system, *e.g.*, for use in higher level applications. We tested two different methods to compute this state $\tilde{x}_t = [\tilde{q}_t, \Delta\tilde{q}_t]$ given the current sample set S_t. The first one is a *winner takes all* (WTA) approach where we just search for the particle s_t^j with highest weight $w_t^j = max(w_t^k)$: $i = 1, \ldots, N$ and assign the state vector associated with this particle ($\hat{x}_t := x_t^j$) to $\tilde{x}_t$. The second option is what we call *averaging in the tangential space*. In general, the particles might represent a multimodal distribution. In order to choose the most likely mode, we restrict the averaging to a region in the exponential chart of $\hat{q}_t$, which denotes the rotation estimate of the best rated particle. This means, we select a subset M of particles whose rotation states q_t^i deviate from $\hat{q}_t$ by an angle less than κ, which has to be chosen according to the grid quantization. Applying the iterative scheme described in Section III to the subset M with $\hat{q}_t$ as the initial estimate provides the mean rotation $\overline{q}_t$ after few iterations. Finally, $\tilde{q}_t = \overline{q}_t$ denotes the most likely rotation.

If real-time computation is not obligatory, a further refinement is possible by numerically optimizing the normalized cross-correlation function with regard to the orientation as described in Section II-D.

V. EXPERIMENTS AND RESULTS

Our prototype implementation is realized in C++ and uses the S2Kit [22] and SOFT library [7]. The refinement procedure makes use of the recurrence relations of the Wigner-d-matrices proposed by Dachsel [23] and FFTWlib [24].

A. Ground-Truth Evaluation

In the first experiment, we capture data with a spherical camera mounted on a high-precision pan-tilt unit (PTU) and compare the rotation estimates to the encoder readings.

We use the mean squared errors (MSE) of the overall angle separating the estimated rotation $\tilde{q}_t$ from the reference rotation q_t^{ref} as error measure. Note that the origin of the camera does not coincide with the rotation center and therefore, an additional translational movement (up to ≈ 50cm) is induced. Correspondingly, a correct rotation estimate in this experiment also approves a certain robustness to small translational movements.

In Table I we show the MSE for a sequence of 180 frames where the PTU rotates stepwise around both axes simultaneously from $(0, -50°, -90°)$ to $(0, +50°, +90°)$ in the ZYZ-Euler representation. The simultaneous rotation around two axes is challenging, since due to the incomplete spherical field-of-view of the camera (about $360° \times 130°$ of the entire sphere) a great amount of new image content is introduced. We applied the particle filter with different refinement methods in comparison to the unfiltered estimate relying solely on the maximum grid value. For all experiments we use a bandwidth of 16 which corresponds to an image size of 32×32. The particle filter is evaluated using 200 particles for the explicit evaluation and 2000 particles for the grid-based approach. Note that an explicit evaluation of 5 particles would result in about the same frame rate as the grid-based approach. The threshold κ was set to $28°$ and we chose 10 for the parameter σ_L. Since the additional weighting of the particles with regard to their angular momentum showed consistently improved performance, we chose to use this method in all experiments.

We see, that, especially for the WTA approach, the gain in accuracy from the explicit evaluation of the correlation function is significant. However, this gain diminishes compared to the grid-based evaluation if an additional averaging is performed on the weighted particles. Furthermore, the computation time differs significantly.

When comparing the SOFT-based evaluation of the correlation grid and the method proposed by Makadia et al. (3DFT, [8]) the error for the 3DFT for a bandwidth of 16 is quite large. This is due to the fact that using 3DFT the correlation grid is computed for Euler angles in the range of $[0, 2\pi]^3$ which covers the space of rotations twice resulting in only half of the accuracy in the rotation around the Y-axis compared to SOFT. We therefore included an estimation using a bandwidth of 32 in the table. When measuring only the time spent in 3DFT and SOFT, the computation times of both methods are comparable, since both make use of a fast implementation of the Fourier transform. Including the time used by the SFT, however, leads to longer execution times for a comparable accuracy of the rotation estimation.

B. Alignment of image sequences

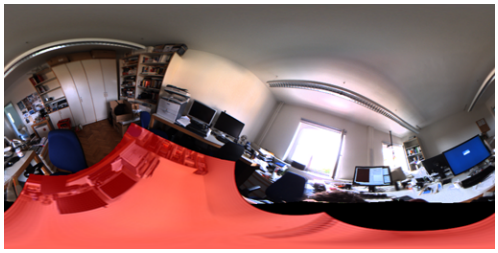
We recorded an omnidirectional video (300 frames) where the camera was moved smoothly including small translations. To visually analyze the performance of the different settings

Filter	Sensing	Refinement	MSE [$\mathbf{deg}^2$]	frames/second
unfiltered			17.08	5.68
unfiltered (3DFT, BW 16)			91.39	11.30
unfiltered (3DFT, BW 32)			31.60	1.00
particle filter	grid-based	WTA	40.22	5.83
particle filter	explicit evaluation	WTA	12.70	0.51
particle filter	grid-based	averaging	8.35	5.72
particle filter	explicit evaluation	averaging	6.62	0.51
particle filter	grid-based	averaging + optimization	4.13	2.36
particle filter	explicit evaluation	averaging + optimization	5.52	0.45

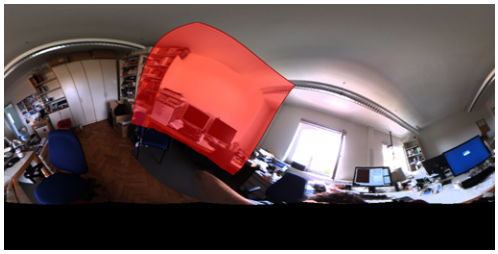
TABLE I

MSE AND FRAME RATE FOR CONTROLLED ROTATION USING DIFFERENT CONFIGURATIONS OF THE PARTICLE FILTER.

we de-rotate the image sequence according to the estimated orientation of the camera. Note that an elegant and efficient way is to apply this de-rotation using the graphics card: The images are texture-mapped onto a sphere that is inversely rotated according to the estimates. Positioning the camera inside the sphere, a perspective view is achieved by a standard perspective projection of this virtual scene. A video showing the results of this experiment as well as the results of the template tracking application is included in the proceedings and available at `http://www.gris.uni-tuebingen.de/people/staff/tschairer/`. In the sequence we simulate twelve situations where a great part of the FOV is occluded by some object resulting in a misleading maximum of the correlation grid. The particle filter, however, compensates this effect successfully. Additionally, due to its local nature the gradient based refinement starting at the particle filter estimate converges to the nearest local maximum of the correlation function resulting in an accurate rotation estimate. In Figure 1 we show an illustration of the estimated orientations over time.



(a) Result of grid-based unfiltered estimation.



(b) Result of estimation using the particle filter and optimization.

Fig. 3. Single frames of the template tracking sequence.

C. Template tracking on the sphere

The spatial normalization (see Sec. II-C) allows to test an application of tracking a small template on an omnidirectional image sequence. The combination of a low bandwidth of the SFT transform and a small template image (appr. 12% of the entire sphere in this case) can often lead to erroneous estimates (we counted 32 frames on a 300 frames test sequence). The particle filter successfully compensates these outliers. Figure 2 shows an illustration of a part of the sequence of the orientation estimates. The unfiltered grid-based approach yields erroneous estimates quite frequently, whereas the result of using the particle filter and optimization technique leads to a smooth estimated trajectory. In Figure 2(c) the single particles are visualized on the unit sphere and multiple modes of the pdf can be observed. Two single frames showing the estimated rotation of the templates can be seen in Figure 3.

VI. CONCLUSION

We discussed the application of particle filters to orientation estimation and presented a novel approach to integrating rotation estimates based on harmonic analysis in such a filter. A brief review of the methods to compute the cross-correlation as similarity measure for the rotation estimate is given and different techniques for the refinement of the grid-based approaches are discussed. Further, we present an iterative optimization-based method to compute the spatially normalized cross-correlation using a separation of variables technique. Since the particle filter works on a state space consisting of orientations and the angular momentum of the agent, we also discuss the group of unit quaternions and their tangent space representation. The system is evaluated on real image data including a comparison with ground-truth information from a high precision pan-tilt-unit and a template tracking application.

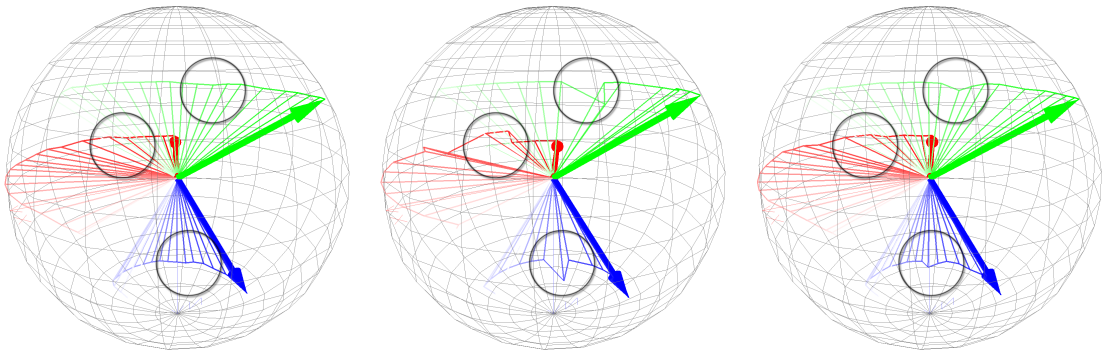
In future work we plan to use the system on a mobile agent and a technique to update the template/reference image will be included. This is necessary to estimate the orientation robustly over time for a moving agent.

REFERENCES

[1] C. Fermüller and Y. Aloimonos, "Ambiguity in structure from motion: Sphere versus plane," *Int. J. Comput. Vision*, vol. 28, no. 2, pp. 137–154, 1998.

[2] H. I. Christensen and G. D. Hager, "Sensing and estimation," in *Springer Handbook of Robotics*. Springer, 2008, pp. 87–107.

[3] M. Euston, P. Coote, R. E. Mahony, J. Kim, and T. Hamel, "A complementary filter for attitude estimation of a fixed-wing uav." in *IROS*. IEEE, 2008, pp. 340–345.

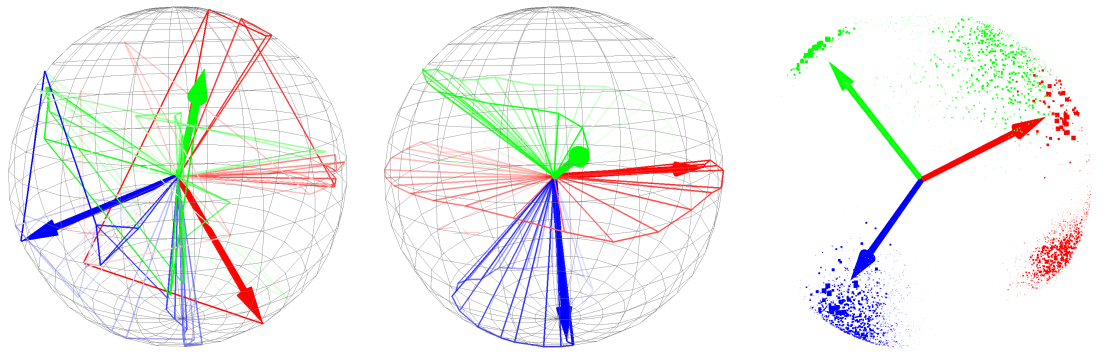


(a) Estimation on unoccluded sequence using particle filtering and optimization.

(b) Grid-based unfiltered estimation of occluded sequence.

(c) Particle filter and optimization applied to occluded sequence.

Fig. 1. Illustration of the estimated camera orientation over time. Note the highlighted regions depicting the frames where occlusions were simulated. Furthermore, the quantization is also quite apparent in (b), whereas (a) and (c) exhibit smooth trajectories.



(a) Result of the grid-based unfiltered estimation in the template tracking application.

(b) Results of the particle filter and optimization in the template tracking application.

(c) Visualization of the particles at a single point in time. The point size denotes the particle weight.

Fig. 2. Illustration of the estimated orientation over time in the template tracking application.

[4] M. Fiala, "Structure from motion using sift features and the ph transform with panoramic imagery," in *Proc. Canadian conference on Computer and Robot Vision (CRV)*, 2005.

[5] J. Lim and N. Barnes, "Directions of egomotion from antipodal points," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.

[6] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 700–703, 1987.

[7] P. Kostelec and D. Rockmore, "Ffts on the rotation group," Fe Institutes Working Paper Series, Tech. Rep., 2003.

[8] A. Makadia and K. Daniilidis, "Rotation recovery from spherical images without correspondences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1170–1175, 2006.

[9] ——, "Correspondenceless structure from motion," *Internation Journal of Computer Vision*, vol. 75, no. 3, pp. 311–327, 2007.

[10] J. Fehr, M. Reisert, and H. Burkhardt, "Fast and accurate rotation estimation on the 2-sphere without correspondences," in *European Conference on Computer Vision (ECCV)*, ser. LNCS, vol. 5303, 2008, pp. 239–251.

[11] H. Friedrich, D. Dederscheck, E. Rosert, and R. Mester, "Optical rails: View-based point-to-point navigation using spherical harmonics," in *DAGM 2008*, 2008.

[12] J. Driscoll and D. M. Healy, Jr., "Computing fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, no. 2, pp. 202–250, 1994.

[13] B. Huhle, T. Schairer, and W. Straßer, "Normalized cross-correlation using SOFT," in *International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*, 2009.

[14] J. Hornegger and C. Tomasi, "Representation issues in the ml estimation of camera motion," in *International Conference on Computer Vision (ICCV)*, 1999.

[15] J. Diebel, "Representing Attitude: Euler Angles, Quaternions, and Rotation Vectors," Stanford University, Palo Alto, CA, Tech. Rep., 2006.

[16] M. soo Kim and K. won Nam, "Interpolating solid orientations with circular blending quaternion curves," *Computer Aided Design*, vol. 27, pp. 385–398, 1995.

[17] K. Shoemake, "Animating rotation with quaternion curves," in *Proc. SIGGRAPH '85*. New York, NY, USA: ACM Press, 1985, pp. 245–254.

[18] X. Pennec, "Computing the mean of geometric features – application to the mean rotation," Institut National de Recherche en Informatique et en Automatique, Tech. Rep., 1998.

[19] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, February 2002.

[20] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *Radar and Signal Processing, IEE Proceedings F*, vol. 140, pp. 107–113, 1993.

[21] R. Douc and O. Cappe, "Comparison of resampling schemes for particle filtering," in *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on*, 2005.

[22] P. Kostelec and D. Rockmore, "S2kit: A lite version of spharmonickit," Dartmouth College, Tech. Rep., 2004.

[23] H. Dachsel, "Fast and accurate determination of the wigner rotation matrices in the fast multipole method," *Journal of Chemical Physics*, vol. 124, 2006.

[24] M. Frigo and S. G. Johnson, "The design and implementation of FFTW3," *Proceedings of the IEEE*, vol. 93, no. 2, pp. 216–231, 2005, special issue on "Program Generation, Optimization, and Platform Adaptation".