Outdoor Human Motion Capture using Inverse Kinematics and von Mises-Fisher Sampling
Supplemental Material

Gerard Pons-Moll[1,*] Andreas Baak[2] Juergen Gall[3] Laura Leal-Taixé[1]
Meinard Müller[2] Hans-Peter Seidel[2] Bodo Rosenhahn[1]
[1]Leibniz University Hannover, Germany [2]Saarland University & MPI Informatik, Germany [3]BIWI, ETH Zurich

Abstract

This is the supplemental material for [5]. It contains a more detailed description of the closed form algorithm to compute inverse kinematics based on the Paden-Kahan subproblems. For an extended and more detailed version of [5] we refer the reader to [7].

1. Paden-Kahan subproblems

We are interested in solving the following problem:

$$\exp(\theta_1\widehat{\omega}_1)\exp(\theta_2\widehat{\omega}_2)\exp(\theta_3\widehat{\omega}_3) = \mathbf{R}_j. \tag{1}$$

This problem can be solved by decomposing it into subproblems as proposed in [4]. A comprehensive description of the Paden-Kahan subproblems applied to several inverse kinematic problems can also be found in [3]. The basic technique for simplification is to apply the kinematic equations to specific points. By using the property that the rotation of a point on the rotation axis is the point itself, we can pick a point $\mathbf{p}$ on the third axis ω_3 and apply it to both sides of Eq. (1) to obtain

$$\exp(\theta_1\widehat{\omega}_1)\exp(\theta_2\widehat{\omega}_2)\mathbf{p} = \mathbf{R}_j\mathbf{p} = \mathbf{q} \tag{2}$$

which is known as the *Paden-Kahan sub-problem 2*. For our problem the 3 rotation axes intersect at the same joint location. Consequently, since we are only interested in the orientations, we can translate the joint location to the origin $\mathbf{q}_j = O = (0,0,0)^T$. In this way, any point $\mathbf{p} = \lambda\omega_3$ with $\lambda \in \mathbb{R}$, $\lambda \neq 0$ is a valid choice for $\mathbf{p}$. Eq. (2) can decomposed in two subproblems

$$\exp(\theta_2\widehat{\omega}_2)\mathbf{p} = \mathbf{c} = \exp(-\theta_1\widehat{\omega}_1)\mathbf{q} \tag{3}$$

where $\mathbf{c}$ is the intersection point between the circles created by the rotating point $\mathbf{p}$ around axis ω_2 and the point

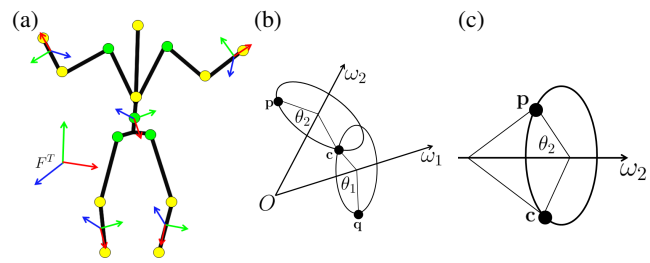

Figure 1: Inverse Kinematics: **(a)** decomposition into active (yellow) and passive (green) parameters. Paden-Kahan subproblem 2 **(b)** and sub-problem 1 **(c)**.

$\mathbf{q}$ rotating around axis ω_1 as shown in Fig. 1 (b). In order for Eq. (3) to have a solution, the points $\mathbf{p}$, $\mathbf{c}$ must lie in the same plane perpendicular to ω_2, and $\mathbf{q}$, $\mathbf{c}$ must lie in the same plane perpendicular to ω_1. This implies that the projection of the position vectors [1] $\mathbf{p}, \mathbf{c}, \mathbf{q}$ onto the span of ω_1, ω_2 respectively must be equal, see Fig. 2

$$\omega_2^T\mathbf{p} = \omega_2^T\mathbf{c} \quad \text{and} \quad \omega_1^T\mathbf{q} = \omega_1^T\mathbf{c} \tag{4}$$

Additionally, the norm of a vector is preserved after rotation and therefore

$$\|\mathbf{p}\| = \|\mathbf{c}\| = \|\mathbf{q}\| \tag{5}$$

Since ω_1 and ω_2 are not parallel, the vectors $\omega_1, \omega_2, \omega_1 \times \omega_2$ form a basis that span $\mathbb{R}^3$. Hence, we can write $\mathbf{c}$ in the new basis as

$$\mathbf{c} = \alpha\omega_1 + \beta\omega_2 + \gamma(\omega_1 \times \omega_2) \tag{6}$$

where α, β, γ are the new coordinates of $\mathbf{c}$. Now, using the fact that $\omega_2 \perp \omega_1 \times \omega_2$ and $\omega_1 \perp \omega_1 \times \omega_2$, we can substitute Eq. (6) into Eq. (4) to obtain a system of two equations with

[1] Since we translated the joint location to the origin we can consider the points as vectors with origin at the joint location $\mathbf{q}_j$.

two unknowns (α, β)

$$\begin{aligned} \omega_2^T \mathbf{p} &= \alpha \omega_2^T \omega_1 + \beta \\ \omega_1^T \mathbf{q} &= \alpha + \beta \omega_1^T \omega_2 \end{aligned} \tag{7}$$

from which we can isolate the first two coordinates of $\mathbf{c}$

$$\begin{aligned} \alpha &= \frac{(\omega_1^T \omega_2)\omega_2^T \mathbf{p} - \omega_1^T \mathbf{q}}{(\omega_1^T \omega_2)^2 - 1} \\ \beta &= \frac{(\omega_1^T \omega_2)\omega_1^T \mathbf{q} - \omega_2^T \mathbf{p}}{(\omega_1^T \omega_2)^2 - 1} \end{aligned} \tag{8}$$

From Eq. (5) and Eq. (6) we can write

$$\|\mathbf{p}\|^2 = \|\mathbf{c}\|^2 = \alpha^2 + \beta^2 + 2\alpha\beta\omega_1^T\omega_2 + \gamma^2\|\omega_1 \times \omega_2\|^2 \tag{9}$$

and obtain the third coordinate γ as

$$\gamma^2 = \frac{\|\mathbf{p}\|^2 - \alpha^2 - \beta^2 - 2\alpha\beta\omega_1^T\omega_2}{\|\omega_1 \times \omega_2\|^2} \tag{10}$$

This last equation has no solution when the circles do not intersect, one solution when the circles are tangential and two solutions when the cirlces intersect at two points. For our choice of decomposition, the passive parameters correspond to $3DoF$ joints which are modelled as 3 concateneted revolute joints whose axis are mutually orthogonal. Therefore, there always exists a solution [3]. We note that the inverse kinematic solutions presented here are also valid for other decompositions, *e.g.* one could choose as passive paramters two rotation axes of the shoulder joint and one rotation axis of the elbow joints. However, the existence of solution should then be checked during the sampling process. Once we have the new coordinates (α, β, γ) we can obtain the intersection point $\mathbf{c}$ in the original coordinates using equation Eq. (6). Thereafter, Eq. (3) can be decomposed into two problems of the form

$$\begin{aligned} \exp(\theta_2 \widehat{\omega}_2)\mathbf{p} &= \mathbf{c} \\ \exp(-\theta_1 \widehat{\omega}_1)\mathbf{q} &= \mathbf{c} \end{aligned} \tag{11}$$

which simplifies to finding the rotation angle about a fixed axis that brings a point $\mathbf{p}$ to a second one $\mathbf{c}$, which is known as *Paden-Kahan sub-problem 1*

$$\exp(\theta_2 \widehat{\omega}_2)\mathbf{p} = \mathbf{c}. \tag{12}$$

This problem has a solution when the projections of the vectors $\mathbf{p}$ and $\mathbf{c}$ onto the orthogonal plane to ω_2 have equal lengths. Let $\mathbf{p}'$ and $\mathbf{c}'$ be the projections of $\mathbf{p}, \mathbf{c}$ onto the plane perpendicular to ω_2, see Fig. 2,

$$\mathbf{p}' = \mathbf{p} - \omega_2\omega_2^T\mathbf{p} \quad \text{and} \quad \mathbf{c}' = \mathbf{c} - \omega_2\omega_2^T\mathbf{c}. \tag{13}$$

If the projections have equal lengths $\|\mathbf{p}'\| = \|\mathbf{c}'\|$ then the problem is as simple as finding the angle between the two

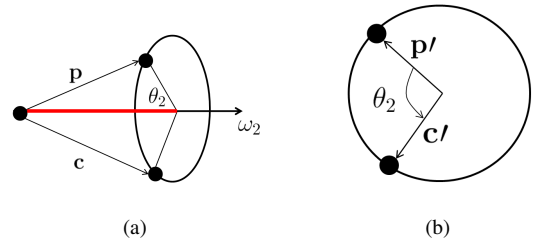

Figure 2: Paden-Kahan subproblem 1: **(a)** the projection length of $\mathbf{p}$ and $\mathbf{c}$ onto ω_2 must be equal, **(b)** the projection of the vectors $\mathbf{p}$ and $\mathbf{c}$ onto the orthogonal plane to the rotation axes ω_2

vectors

$$\begin{aligned} \omega_2^T(\mathbf{p}' \times \mathbf{c}') &= \sin\theta_2\|\mathbf{p}'\|\|\mathbf{c}'\| \\ \mathbf{p}' \cdot \mathbf{c}' &= \cos\theta_2\|\mathbf{p}'\|\|\mathbf{c}'\| \end{aligned} \tag{14}$$

By dividing the equations we finally obtain the rotation angle using the using the arc tangent

$$\theta_2 = \text{atan2}(\omega_2^T(\mathbf{p}' \times \mathbf{c}'), \mathbf{p}' \cdot \mathbf{c}'). \tag{15}$$

We can find θ_1 using the same procedure. Finally, θ_3 is obtained from Eq. (1) after substituting θ_1 and θ_2

$$\exp(\theta_3\widehat{\omega_3}) = \exp(\theta_1\widehat{\omega}_1)^T \exp(\theta_2\widehat{\omega}_2)^T\mathbf{R}_j = \mathbf{R} \tag{16}$$

where the rotation matrix $\mathbf{R}$ is known. The rotation angle θ_3 satisfies

$$2\cos\theta_3 = (\text{trace}(\mathbf{R}) - 1) \tag{17}$$

$$2\sin\theta_3 = \omega_3^T\mathbf{r} \tag{18}$$

where $\mathbf{r} = (\mathbf{R}_{32} - \mathbf{R}_{23}, \mathbf{R}_{13} - \mathbf{R}_{31}, \mathbf{R}_{21} - \mathbf{R}_{12})$ (page 584 of [1]). Finally, the rotation angle θ_3 can be computed from $\cos\theta_3$ and $\sin\theta_3$ using atan2.

By solving these sub-problems for every sensor, we are able to reconstruct the full state $\mathbf{x}$ using only a subset of the parameters $\mathbf{x}_a$ and the sensor measurements $\mathbf{z}^{sens}$.

2. Calibration and Synchronization

We explain here in more detail the calibration steps to align and synchronize IMU and camera coordinate systems. We recorded several motion sequences of subjects wearing 10 inertial sensors which we split in two groups of 5: the *tracking sensors* which we use for tracking and the *validation sensors* which we use for evaluation. The tracking sensors are placed in the back and the lower limbs and the validation sensors are placed on the chest and the upper limbs. An inertial sensor s measures the orientation of its local coordinate system F_s^S within a fixed global frame of reference

F^I. All sensors derive the same global frame of reference by merging information from a magnetic field sensor, an accelerometer and a rate gyro. The orientation data is given as a stream of rotation matrices $\mathbf{R}_s^{IS}(t)$ that define the coordinate transform from F_s^S to F^I. In the process of calibrating the camera, the global tracking coordinate system F^T is defined by a calibration cube placed into the recording volume. In order to bring F^I and F^T into correspondence, we carefully place the calibration cube such that the axes of F^T directly correspond to the axes of the known F^I using a compass. Like this, the orientation data $\mathbf{R}_s^{IS}(t)$ also directly maps from the local sensor coordinate system F_s^S to the global tracking coordinate system F^T and we note $\mathbf{R}^{TS} := \mathbf{R}^{IS}$. Therefore, in contrast to [6], no precalibration needs to be computed from a pre tracked calibration sequence. In this paper, we refer to the sensor orientations by $\mathbf{R}^{TS}$ and, where appropriate, by using the corresponding quaternion representation $\mathbf{q}^{TS}$. In our experiments, we use inertial sensors provided by Xsens [8].

To obtain multiview image sequences, we use four off-the-shelf consumer cameras that provide temporally unsynchronized videos. The videos are synchronized by cross correlating the audio signals with a known audio source position as proposed in [2]. Then, we synchronize the sensor measurements to the cameras using a clapping motion, which can be detected in the audio data as well as in the acceleration data measured by the inertial sensors.

References

[1] R. Hartley and A. Zisserman. *Multiple view geometry*, volume 642. Cambridge university press Cambridge, UK, 2003. 2

[2] N. Hasler, B. Rosenhahn, T. Thormählen, M. Wand, J. Gall, and H.-P. Seidel. Markerless motion capture with unsynchronized moving cameras. In *CVPR*, pages 224–231, 2009. 3

[3] R. Murray, Z. Li, and S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, Baton Rouge, 1994. 1, 2

[4] B. Paden. *Kinematics and control of robot manipulators*. PhD thesis, 1985. 1

[5] G. Pons-Moll, A. Baak, J. Gall, L. Leal-Taixe, M. Muller, H. Seidel, and B. Rosenhahn. Outdoor human motion capture using inverse kinematics and von mises-fisher sampling. In *Proc. IEEE Int Computer Vision (ICCV) Conf*, pages 1243–1250, 2011. 1

[6] G. Pons-Moll, A. Baak, T. Helten, M. Mueller, H.-P. Seidel, and B. Rosenhahn. Multisensor-fusion for 3d full-body human motion capture. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 663–670, 2010. 3

[7] G. Pons-Moll, L. Leal-Taixe, J. Gall, and B. Rosenhahn. Data-driven manifolds for outdoor motion capture. In *Theoretic Foundations of Computer Vision: Outdoor and Large-Scale Real-World Scene Analysis*, 2012. 1

[8] X. M. Technologies. http://www.xsens.com/. 3