

Reduced Marker Layouts for Optical Motion Capture of Hands

Matthias Schröder¹ Jonathan Maycock² Mario Botsch¹

¹Computer Graphics & Geometry Processing Group ²Neuroinformatics Group

Bielefeld University, Germany

Abstract

We present a method for automatically generating reduced marker layouts for marker-based optical motion capture of human hand motions. Reducing the number of markers on the hand is important to ensure the generated motions are performed in a natural way and indeed a reduced marker set might be a technical requirement should simultaneous body motion capture also have to be carried out. The employed motion reconstruction method is based on subspace-constrained inverse kinematics, which allows for the recovery of realistic hand movements even from sparse input data. Our marker layout optimization is sensitive to the kinematic structure and the subspace representations of hand articulations utilized in the reconstruction method in order to generate sparse marker configurations that are optimal for solving the constrained inverse kinematics problem. We propose specific quality criteria for reduced marker sets that combine numerical stability with geometric feasibility of the resulting layout. These criteria are combined in an objective function that is minimized using a specialized surface-constrained particle swarm optimization scheme. Our method provides a principled way for determining reduced marker layouts based on subspace representations of hand articulations.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Hierarchy and geometric transformations I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Tracking

Keywords: Hand Tracking, Motion Capture, Marker Layout

1 Introduction

Marker-based optical motion capture, or *mocap*, is widely regarded as the standard method for acquiring motions of human performers in research, industry and indeed entertainment. Numerous commercial solutions [Vicon 2015; OptiTrack 2015; PhaseSpace 2015; Qualisys 2015] and considerable scientific literature exist on the topic. While there is a multitude of alternative solutions for motion tracking, such as marker-less methods [OrganicMotion 2015; Microsoft 2010] or systems using inertial sensors [Xsens 2015; Biosyn 2015], they are not as widely deployed due to the reliability of marker-based systems. Marker-based optical mocap systems track the 3D positions of markers attached to a performer, which can then be used to infer the articulation of a skeletal model of the tracked subject. Such systems typically consist of 4 to 32 cameras that capture at 30 to 2000 Hz and acquire the marker locations with very high accuracy [Kitagawa and Windsor 2008].

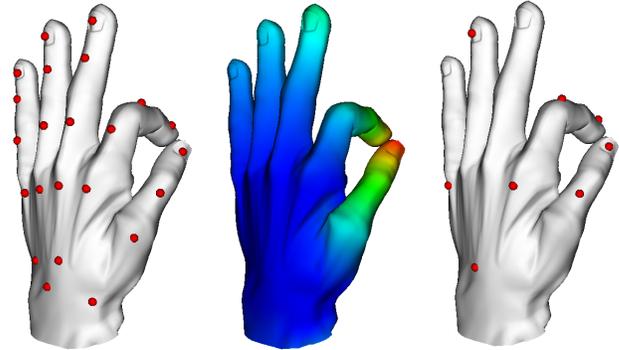


Figure 1: Our method generates reduced marker layouts for optical motion capture of hands by analyzing training data of hand movements. Left: full marker set covering all joints of the hand. Center: qualitative illustration of regions that are static (blue) and in motion (red) during the analyzed precision grasp movements. Right: reduced marker set that is sufficient to reconstruct the observed motions using our method.

However, despite the quality of marker-based mocap there are drawbacks and limitations to these systems. The captured data usually needs to be post-processed extensively, occlusions can cause gaps or mislabelings in the captured data, too many markers may inhibit natural movements, and any rotational information needs to be computed retrospectively. Some of these issues are amplified as the number of markers used for tracking increases. A common guideline for capturing articulated objects is to cover all major joints with markers [Guerra-Filho 2005; Kitagawa and Windsor 2008]. In addition to making the marker attachment process tedious and error-prone, a high number of markers causes problems when capturing multiple subjects or tracking body movements and hand articulations simultaneously. Capturing hand articulations in detail typically requires a dense marker set consisting of 18–23 markers in a small capture volume. In a large capture volume that also allows for full body mocap the resolution of the optical tracking system and the required size of the markers prohibit the usage of a full marker set. Instead, reduced marker sets have been employed in large capture volumes—however, this strongly limits the expressiveness of the captured hand motions. Therefore, body and hand movements are sometimes captured in isolated sessions and combined in post-processing [Wheatland et al. 2015].

In this work, we present a method to automatically determine reduced marker layouts for optical mocap based on inverse kinematics (IK). The motion reconstruction method is based on performing the IK optimization in a subspace learned from prior hand movements, which allows for realistic recovery of hand articulations even from sparse input data. Our method for reduced marker set optimization is sensitive to this reconstruction method, particularly the employed hand posture subspace, and thus produces layouts that are optimal for solving the subspace-constrained IK problem. We present an approach that minimizes an objective function, which jointly optimizes numerical stability of the IK problem and the geometric feasibility of the resulting layout. The optimization is done

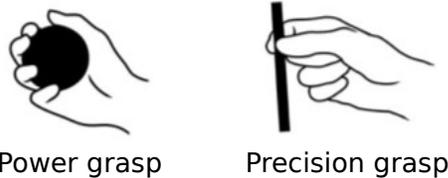


Figure 2: Examples of different grasping types in the taxonomy of Cutkosky [1989], which we use to evaluate our results. Power grasps usually involve the whole hand for interaction with large objects, whereas precision grasps involve only some fingers for handling smaller objects. Illustrations from [Zheng et al. 2011].

using a specialized surface-constrained particle swarm optimization (PSO) [Kennedy and Eberhart 1995; Kennedy and Eberhart 2001], which generates marker layouts bound to the surface of an animated 3D hand model (see Figure 1).

We show that, rather than specifying one marker per joint of the articulated object, it is sufficient to specify one marker per degree of freedom (DoF) of the parameter space that represents particular hand articulations. Reduced marker layouts can therefore be determined by reducing the parameter space of hand postures based on prior knowledge. Furthermore, we show the principles by which a reduced marker layout that best corresponds to the subspace DoFs can be determined. We demonstrate reduced marker layouts for various hand motions, in particular manual interaction movements based on the grasp taxonomy of Cutkosky [1989], which distinguishes between different types of power grasps and precision grasps (see Figure 2).

2 Related work

There is a substantial amount of literature on optical motion capture, therefore we focus on the related work that is most relevant to ours, which includes the topics of motion reconstruction based on motion subspace priors, as well as optimized or reduced marker configurations.

Employing subspace representations of human motions has been shown to be effective for motion reconstruction from sparse input. In [Chai and Hodgins 2005; Liu et al. 2006] local linear models were used to represent full-body motions and recover skeletal articulations from sparse marker sets. While these methods are completely data-driven and can therefore limit the space of recovered articulations, our approach uses data-driven subspaces as a prior but also allows for articulation refinements that lie outside of the ground truth database using a layered IK approach. Liu et al. [2006] also target the problem of determining reduced marker configurations by finding a subset of an initial input marker set that can produce accurate predictions of the remaining markers. In contrast, we present a bottom-up approach for generating optimal reduced marker layouts for hands based on the kinematic DoFs of an articulated hand model. While previous methods usually determine reduced marker sets by subsampling a specific initial marker set, our method more generally prescribes properties that candidate marker regions on the surface of a hand model should exhibit, and automatically computes the optimal marker placement within these regions.

Other works deal with the optimal placement of markers, although not necessarily reduced marker layouts. Recently, Loper et al. [2014] demonstrated an approach that is able to capture fine details of soft tissue deformations in addition to full-body skeletal motions without having to rely on very dense marker sets. To improve the accuracy of their motion and shape capture, they extend their

initial sparse marker set in a greedy approach that iteratively adds the next best mesh vertex that minimizes an error metric. We show that, for the problem of finding good reduced hand marker layouts, such greedy approaches are outperformed by our PSO-based global search, as it is less prone to suboptimal local minima.

Le et al. [2013] explore the problem of determining optimal marker layouts for facial performance capture using an approach that minimizes the reconstruction error for ground truth sequences of high-resolution facial meshes. While their approach is based on surface deformations of facial meshes, we find reduced marker layouts by purposefully exploiting the kinematic structure and correlations within an articulated hand model.

While a common guideline for marker placement on hands is to use one marker per joint [Guerra-Filho 2005; Kitagawa and Windsor 2008], reduced marker layouts for hands have been frequently discussed. In [Kitagawa and Windsor 2008] an example for a reduced “mitten” layout was given, where only one marker was placed at the tip of a single finger. Given an estimation for the global location and orientation of the hand, the relative movement of this marker can be interpreted as the simultaneous bending of all fingers. Our work examines this concept more closely by considering how correlations and redundancies in hand articulations affect marker placement.

Regarding the degree of realism of finger motions with reduced marker sets, Hoyet et al. [2012] found that humans are not particularly sensitive to the subtle details of finger animations and that the perceived quality of motions is not significantly affected by reduced marker sets. While they manually selected reduced marker configurations, we present an automatic approach based on subspace-constrained inverse kinematics. In contrast, Chang et al. [2007] determine the most important markers in a reduced marker set for the purpose of grasp motion recognition by using supervised feature selection based on the prediction accuracy of grasp classifiers.

In [Kang et al. 2012; Wheatland et al. 2013] a data-driven approach for hand motion reconstruction from sparse marker sets was used, where motions are synthesized by finding database postures that most resemble the low-dimensional input. Wheatland et al. [2013] compute a subset of an initial full marker set by performing principal component analysis (PCA) on the marker trajectories and selecting the most influential ones. Our method differs from theirs in two significant aspects: First, our IK-based approach allows for the recovery of hand articulations that are not present in the prior database, and second, we determine reduced marker layouts in a bottom-up way based on the PCA of joint angles, which explicitly captures the correlations and redundancies present within hand kinematics, unlike positional marker trajectories.

Using PCA or other dimension reduction techniques for hand kinematics has found widespread success in hand tracking, animation, and automation [Bernstein 1967; Wu et al. 2001; Kato et al. 2006; Mulatto et al. 2013; Schröder et al. 2014; Tagliasacchi et al. 2015]. To reconstruct the kinematic parameters of an articulated hand model from positional marker data, we follow our previous subspace-constrained IK approach from [Schröder et al. 2014]. In that work we showed that using subspace constraints the hand posture estimations remain realistic even when input data is missing. Here, we reverse the problem and seek to find the minimal amount of marker input data necessary to reconstruct postures accurately using subspace priors. As in previous works on reduced marker sets for hand mocap [Chang et al. 2007; Kang et al. 2012; Wheatland et al. 2013; Hoyet et al. 2012], our marker layouts describe only the articulation of the hand, whereas the global position is given by markers placed on the forearm near the wrist.

In the following we describe the employed motion reconstruction method (Section 3), before discussing the specific quality criteria

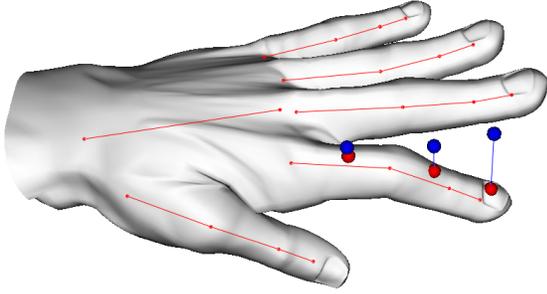


Figure 3: Hand model and its underlying skeleton. Also shown are three exemplary markers on the hand (red), which should move towards their target positions (blue) using inverse kinematics.

for reduced marker layouts (Section 4) and presenting our layout optimization scheme (Section 5). Finally we show and discuss some results of our marker layout optimization and motion reconstruction approach (Sections 6 and 7).

3 Motion reconstruction

Given a set of target marker positions from an optical mocap system, our motion reconstruction algorithm estimates the hand posture from which the observed positions originate by fitting an articulated hand model to the data. The hand model consists of 16 joints, which are driven by 26 kinematic parameters $\theta = (\theta_1, \dots, \theta_{26})^T$. Of those parameters, 6 describe the global pose of the hand: 3 for translation and 3 for rotation. The remaining 20 parameters describe the posture of the fingers, where each finger defines 4 joint angle parameters. The hand geometry is represented by a triangle mesh, which is animated using linear blend skinning [Jacobson et al. 2014]. On the surface of this model effector positions are defined, which correspond to the marker target positions in the input data. The associations between the target and effector positions can be obtained by either manually labeling the observed data or computing the labels automatically [Meyer et al. 2014; Maycock et al. 2015]. Figure 3 shows the hand model with its underlying skeleton and some exemplary markers. The problem of finding the hand model parameters that move the effector positions to their corresponding targets is solved using inverse kinematics. We apply the subspace-constrained IK of Schröder et al. [2014] to the marker-based mocap problem.

3.1 Inverse kinematics

The positions of the k effectors on the surface of the hand model are represented as a stacked vector $\mathbf{x} \in \mathbb{R}^{3k}$ and move relative to the model articulation. They can therefore be expressed as a function of the kinematic parameters: $\mathbf{x} = \mathbf{x}(\theta)$. These effector positions are subject to move to their corresponding target positions $\mathbf{t} \in \mathbb{R}^{3k}$. The IK problem $\mathbf{t} = \mathbf{x}(\theta)$ is solved by finding an update to the kinematic parameter vector θ that minimizes the objective function

$$E_{\text{IK}}(\delta\theta) = \frac{1}{2} \|\mathbf{x}(\theta + \delta\theta) - \mathbf{t}\|^2 + \frac{1}{2} \|\mathbf{D} \delta\theta\|^2. \quad (1)$$

In this objective function, the first term models the least squares error between the positions of the effector points \mathbf{x}_i and the positions of their corresponding target points \mathbf{t}_i . The second term is a selective damping for the parameter update $\delta\theta$ through a diagonal matrix \mathbf{D} , which stabilizes the solution and is used for joint limit avoidance [Schröder et al. 2014].

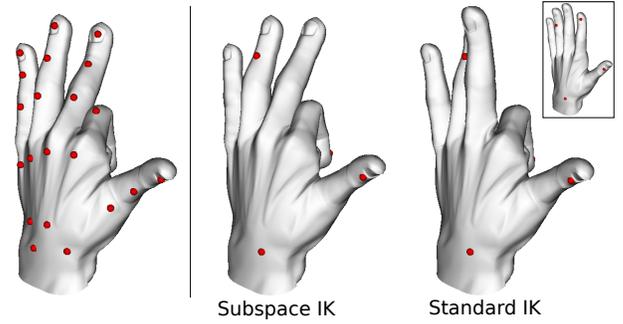


Figure 4: Hand postures reconstructed from full marker set (left) and from reduced marker set using standard IK (center) and subspace-IK (right). While standard IK cannot articulate the marker-less fingers, subspace IK captures the correlations between fingers and articulates them using the reduced marker set.

To find the parameter update $\delta\theta$, the objective function (1) is minimized using a Gauss-Newton approach, in which a linear system of the following form is solved in each iteration:

$$(\mathbf{J}^T \mathbf{J} + \mathbf{D}) \delta\theta = \mathbf{J}^T (\mathbf{t} - \mathbf{x}(\theta)), \quad (2)$$

where $\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \theta}$ is the $(3k \times 26)$ Jacobian matrix of the effector positions [Buss 2004]. After solving the linear system, the resulting update $\delta\theta$ is scaled using a line search in order to guarantee convergence. The process of solving the linear system (2) and updating the effector positions is iterated 5–10 times.

The result of this process is an update to the kinematic parameter vector θ that moves the effector positions on the model to the marker target positions in the input data. Given a full marker set that specifies the articulation of every joint this produces accurate reconstructions of the input motion. However, when using reduced marker sets the input data is sparse and the motions of joints that are not constrained by marker positions cannot be recovered. For this reason, a subspace prior that captures the correlations of joint movements is employed in the inverse kinematics scheme.

3.2 Subspace prior

To obtain a subspace representation of hand articulations, the publicly available database¹ of Schröder et al. [2014], which contains a high variety of human hand motions, was processed using PCA. Performing PCA on this database of 20-dimensional hand postures yields a set of eigenvectors and eigenvalues, which can be used to construct a $26 \times (6 + l)$ matrix of principal components \mathbf{M} , which maps between the full 20-dimensional posture space and a reduced l -dimensional subspace. The additional 6 dimensions encode the global pose of the hand, which is not captured in the PCA model. The number of subspace dimensions, l , determines the amount of variance in the input data covered by the subspace and can be seen as a control variable for the eventual number of markers, k , employed in a reduced marker layout. It was shown in [Schröder et al. 2014] that in order to represent 90% of given hand movements, 3–6 subspace dimensions are sufficient.

Given the PCA matrix \mathbf{M} , the full parameter vector $\theta \in \mathbb{R}^{26}$ can be computed from the reduced subspace parameters $\alpha \in \mathbb{R}^{6+l}$ as

$$\theta = \mathbf{M}\alpha + \mu, \quad (3)$$

¹<http://graphics.uni-bielefeld.de/publications/icra14/>

where $\boldsymbol{\mu} \in \mathbb{R}^{26}$ is the mean of the database postures. This makes it possible to represent the forward kinematics of the effector points \mathbf{x} subject to the subspace parameters: $\mathbf{x} = \mathbf{x}(\boldsymbol{\alpha}) = \mathbf{x}(\boldsymbol{\theta}(\boldsymbol{\alpha}))$.

Based on this representation, the IK problem can be expressed in terms of the subspace parameters. Optimizing for the subspace parameters in (1) and (2) is possible using the subspace Jacobian

$$\mathbf{J}_{\text{PC}} := \frac{\partial \mathbf{x}}{\partial \boldsymbol{\alpha}} = \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \cdot \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\alpha}} = \mathbf{J} \cdot \mathbf{M}. \quad (4)$$

Substituting \mathbf{J}_{PC} for \mathbf{J} in the linear system (2) and analogously changing the damping matrix \mathbf{D} yields the IK solution for the subspace parameters. This solution naturally constrains the reconstructed hand postures to linear combinations of the principal components of the posture database and allows joints to move in correlation to others even when they are not constrained by markers.

However, as there can be variations between the movements contained in the database and the ones observed in the mocap data, we only use this subspace estimate as an initialization for a subsequent refinement of the full posture parameters. By removing the subspace constraints after the initialization of the subspace parameters $\boldsymbol{\alpha}$ and refining the estimate by solving the IK problem again for the full parameter vector $\boldsymbol{\theta}$, the joints with markers are allowed to move more closely to the observed marker positions. This layered IK scheme makes it possible to obtain hand motion reconstructions that are both realistic, due to the subspace prior, and accurate, due to the full kinematic refinement. Figure 4 shows a comparison of standard IK with the subspace approach we employ.

4 Reduced marker layouts

Subspace-constrained inverse kinematics makes it possible to fully articulate a hand model based on a sparse set of marker points. However, the choice of marker placement is not arbitrary, and finding the optimal marker layout requires a method for assessing the quality of a given layout in relation to others. In the following, we discuss the general considerations taken into account and the specific quality metrics employed in our marker layout optimization.

For a given hand motion trajectory, the most straightforward way to evaluate the quality of a given marker set is to compare a large marker set as ground truth data with one reconstructed using a reduced marker set. The specific metric we consider here is the positional reconstruction error, which measures the deviation of the reconstructed trajectories of the model vertices \mathcal{V} from the ground truth trajectories. While this is an intuitive measurement for the deviations in the results of the motion reconstruction (see, e.g., Figure 5), it is not convenient as a metric for choosing an optimal marker layout. Its computation is prohibitively inefficient and it does not generalize beyond the specific input trajectory. Instead, we use metrics that effectively incorporate the IK problem setup, the subspace DoFs, and generic geometric considerations.

A reduced marker set must be configured in such a way that the subspace IK can produce the most accurate results. Additionally, the layout must be designed such that it is well suited for practical use, which means that it should be unobtrusive, easy to apply, and should obviate occlusions and self-contact. In the following, we break these requirements down into two categories: numerical stability and geometric feasibility.

4.1 Numerical stability

Our IK hand motion reconstruction is based on solving the linear system (2). The numerical stability of the IK problem is measured by the invertibility of the left hand side matrix $\mathbf{J}^T \mathbf{J} + \mathbf{D}$,

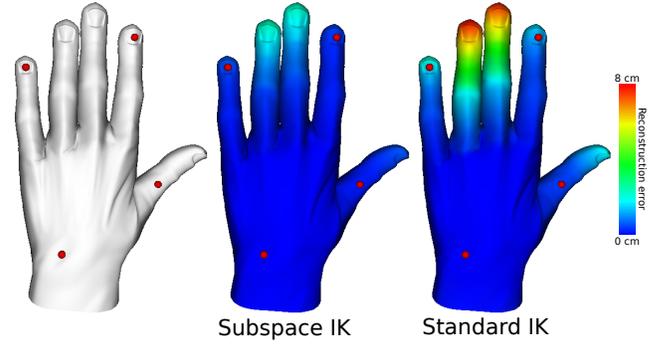


Figure 5: Visualization of the reconstruction error of a reduced marker layout for grasping motions. In the analyzed motion data, all fingers bend forward simultaneously and the reduced layout only specifies markers on three out of the five fingers. Using standard IK this causes a large error in the marker-less fingers, but using subspace IK the movement of these fingers is correlated to the outer markers, which lowers the error.

the key component of which is the Jacobian \mathbf{J} (or \mathbf{J}_{PC}), which is the derivative of the marker positions with respect to the kinematic (or subspace) parameters. Different marker layouts define different Jacobians, each marker defines three rows in the Jacobian matrix. Therefore we denote the Jacobian matrix produced by a specific marker layout \mathcal{M} as $\mathbf{J}_{\mathcal{M}}$. Each kinematic (or subspace) DoF corresponds to a column in the Jacobian. As we are only interested in the minimal layout necessary for accurate posture estimation (joint angles), we omit the three columns in the Jacobian that correspond to translational DoFs, which means that $\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}}$ is a 23×23 matrix for the full parameter space and a $(3 + l) \times (3 + l)$ matrix for the reduced parameter space.

A criterion for the invertibility of a matrix is its condition number, which is low when the problem is well-conditioned and high when it is ill-conditioned. As we are interested in the numerically most stable marker layout, we omit the damping matrix \mathbf{D} , which is not impacted by the markers, and only regard the condition number of the matrix $\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}}$. We compute the condition number of the matrix $\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}}$ using its singular values as

$$\kappa(\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}}) = \left| \frac{\sigma_{\max}(\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}})}{\sigma_{\min}(\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}})} \right|, \quad (5)$$

where $\sigma_{\max}(\mathbf{A})$ and $\sigma_{\min}(\mathbf{A})$ denote the maximum and minimum singular values of matrix \mathbf{A} , respectively.

Optimizing the marker layout \mathcal{M} for the condition number $\kappa(\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}})$ produces marker layouts whose IK solutions are numerically stable by covering the kinematic DoFs of the hand. Taking into account the subspace prior in the IK system by using the subspace Jacobian \mathbf{J}_{PC} from (4), the marker positions tend toward positions that optimally cover the subspace DoFs. Figure 6 illustrates this concept. Note that the number of markers needed to specify the IK problem is determined by the number of DoFs representing the posture. The full posture space therefore cannot be used to produce sparse marker sets of less than 8 markers, since the IK problem would be under-specified. Employing a subspace representation facilitates reduced marker sets.

4.2 Geometric feasibility

Optimizing only for the condition number of the system matrix produces numerically stable and kinematically meaningful marker layouts, however they can be unsuitable for practical use by placing

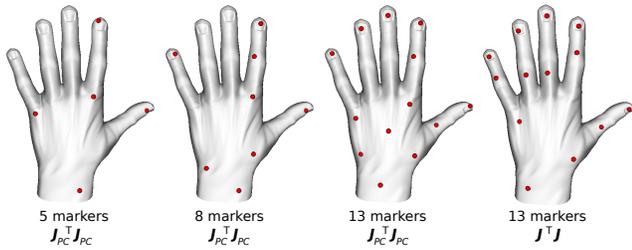


Figure 6: Marker layouts of different sizes for a precision grasp movement involving the index finger and thumb. The rightmost layout with 13 markers was computed using the full Jacobian \mathbf{J} for the condition number metric, whereas the others were computed using the reduced Jacobian \mathbf{J}_{PC} .

markers at positions that are obstructive for the mocap performer or are sensitive to occlusions and self-contact.

Therefore, we consider geometric feasibility in addition to numerical stability in order to produce well-conditioned marker layouts that are also good in practice. We do this in part by limiting the areas where markers can be placed. While this could be done by manually predefining allowed regions, this would cause the need for user intervention. Instead, we define some generic properties that the model vertices should exhibit to select feasible ones automatically. Additionally, we need to model geometric properties that cannot be accounted for by preselecting vertices, as they change during hand motions (e.g., self-contact).

The first set of geometric feasibility properties is the potential areas for positioning the markers on the surface of the hand model. As the hand naturally bends inwards and can come in contact with objects in the front, markers should generally not be placed on the front side, but rather on the back. Similarly, the markers should be prevented from touching the other fingers during motion and therefore markers should not be placed towards the sides of the fingers. We therefore define feasible regions on the surface of the hand model based on the vertex normals. Only vertex positions $\mathbf{p}_i \in \mathcal{V}$ whose normals \mathbf{n}_i satisfy the condition $\mathbf{n}_i \cdot \mathbf{h} > 0.9$, where \mathbf{h} is the hand model’s back-facing vector, are eligible as marker positions.

The second set of geometric feasibility properties taken into account is marker movement. In practice, markers placed near the joint pivot can move non-rigidly along with the joint rotation due to stretching and sliding of the skin. To prevent this, we identify regions on the skinned mesh that move rigidly relative to joints by considering the hand model vertices’ convex skinning weights [Jacobson et al. 2014] and only using vertices with weight 1 for one single joint. Another movement-related issue is when markers can come in contact with each other during motions, which is especially important even with reduced marker layouts when using large markers. To prevent marker contact from occurring, we maximize the minimum distance between markers across multiple key-frames in the input trajectory. For a single frame, the minimum distance between two markers in a marker set \mathcal{M} is

$$\text{md}(\mathcal{M}) = \min_{\mathbf{a} \in \mathcal{M}} \left\{ \min_{\mathbf{b} \in \mathcal{M} \setminus \{\mathbf{a}\}} \{\|\mathbf{a} - \mathbf{b}\|^2\} \right\}. \quad (6)$$

Maximizing this objective function over all frames causes markers to spatially disperse as far from each other as possible, particularly when finger movements cause otherwise spatially distant markers to approach each other more closely.

The combination of these criteria serve as a geometric regularization to the kinematic constraints imposed on the marker set, causing the markers to be placed in geometrically feasible hand regions.

The layouts shown in Figure 6 combine the numerical and geometric criteria.

5 Layout optimization

We now combine the quality measures for reduced marker layouts in an energy minimization scheme, in which the marker set \mathcal{M} that minimizes an objective function $E(\mathcal{M})$ is found using stochastic optimization. To this end, we employ a specialized surface-constrained particle swarm optimization (PSO) scheme, which confines the solution domain to the vertices \mathcal{V} of an animated hand model. In addition to the vertices, the input to this optimization includes the vertex normals and skinning weights, as well as a training set of example hand motions. The marker set quality properties are evaluated on the model’s vertex positions. A distinction can be made between static properties, which are invariant to hand motion and relative marker placements, and dynamic properties, which vary with different motions and marker layouts.

Static aspects of marker layout quality are those that prevent negative effects of skin sliding (using vertices’ skinning weights) and obstructiveness (using vertices’ normals). These properties can easily be incorporated by preselecting only the vertices that satisfy them. This yields a set of preselected vertices $\mathcal{V}' \subset \mathcal{V}$ on the hand model surface that are eligible as potential marker positions. Ultimately, the optimized marker layout will be a subset $\mathcal{M} \subset \mathcal{V}'$ of this preselection.

In contrast, dynamic aspects of marker layout quality cannot be evaluated as isolated vertex properties, as they vary with changes in hand articulation and placement of the remaining markers within the layout. These include the numerical stability measured by the condition number of the IK system matrix, $\kappa(\mathbf{J}_{\mathcal{M}}^T \mathbf{J}_{\mathcal{M}})$, and the minimum marker distance $\text{md}(\mathcal{M})$. To account for these changes with respect to different hand articulations, we evaluate and accumulate these metrics over a set \mathcal{F} of representative key-frames of a given input hand motion trajectory, which can be automatically computed using farthest point optimization [Schlömer et al. 2011] in the hand posture domain. These dynamic properties of the marker set \mathcal{M} are modeled in the objective function $E(\mathcal{M})$, whose definition and optimization are discussed in the following.

5.1 Objective function

The objective function to be minimized during marker layout optimization is a weighted combination of energy terms with respect to marker set \mathcal{M}

$$E(\mathcal{M}) = w_1 \cdot E_{\text{cond}}(\mathcal{M}) + w_2 \cdot E_{\text{dist}}(\mathcal{M}), \quad (7)$$

where $E_{\text{cond}}(\mathcal{M})$ penalizes the condition number of the IK system matrix induced by the marker layout Jacobian, and $E_{\text{dist}}(\mathcal{M})$ penalizes the minimum distance between any two marker positions in the layout. Both terms are evaluated over a set \mathcal{F} of frames from a hand motion trajectory that are representative of the movements that should be captured in the reduced marker set. We denote the marker configuration of layout \mathcal{M} in frame $f \in \mathcal{F}$ as $\mathcal{M}_{(f)}$.

Based on (5), the energy term penalizing the condition numbers of the induced system matrices is defined as

$$E_{\text{cond}}(\mathcal{M}) = \frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \kappa(\mathbf{J}_{(f)}^T \mathbf{J}_{(f)}), \quad (8)$$

where $\mathbf{J}_{(f)}$ denotes the Jacobian of marker configuration $\mathcal{M}_{(f)}$. This term minimizes the average condition number across all frames $f \in \mathcal{F}$. Since the considered marker layout is a subset of the

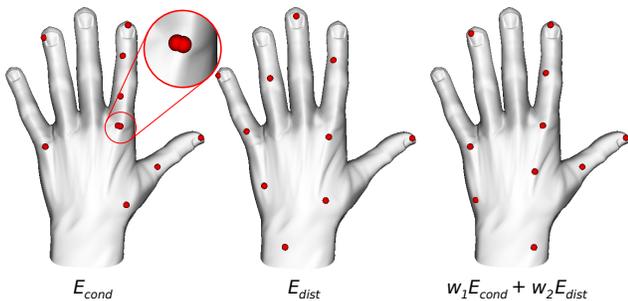


Figure 7: Example 10-marker layouts for the individual objective function terms. The input data is a precision grasp, where mostly the index finger and thumb are in motion. Left: When optimizing only for the numerical stability E_{cond} , markers can be placed in close proximity, which is geometrically impractical. Center: Optimizing for the geometric distance E_{dist} results in spatially distant markers, but the layout does not capture the analyzed hand articulations. Right: A weighted combination of the two terms results in a layout that is both numerically stable and geometrically feasible.

preselected vertices, $\mathcal{M} \subset \mathcal{V}'$, we can precompute the vertex Jacobian $\mathbf{J}_{\mathcal{V}'}$ for all frames in \mathcal{F} and construct the respective marker Jacobians by selecting the corresponding rows in this matrix.

Based on (6), the energy term penalizing the minimum distance between two marker positions across all key-frames is defined as

$$E_{dist}(\mathcal{M}) = -\frac{1}{L} \min_{f \in \mathcal{F}} \{md(\mathcal{M}_{(f)})\}, \quad (9)$$

where L is the length of the hand model, making the term scale invariant. As we want to maximize the minimum distance between two markers, this term aims to minimize the negative of the overall minimum distance over all frames \mathcal{F} .

Combining these two energy terms integrates the desired numerical stability and geometric feasibility properties of the marker layout in a single objective function. The results of minimizing the two energy terms and their weighted sum is illustrated in Figure 7. In this particular example, the condition energy places two markers close to each other, because the linear system for the subspace parameters is over-specified by the number of markers, which means that close-by markers do not corrupt the matrix conditioning. Combining the two energies improves the resulting layout. We use weights $w_1 = 0.1$ and $w_2 = 100$ in all our experiments. In the following, the optimization of the objective function (7) is detailed.

5.2 Marker PSO

We find reduced marker layouts by minimizing the objective function (7) using particle swarm optimization (PSO). PSO is a stochastic meta-heuristic for finding global optima of arbitrary objective functions without the need for prior knowledge or assumptions about the optimized problem. The method has recently found widespread application and success in the context of visual hand tracking [Oikonomidis et al. 2011; Qian et al. 2014; Sharp et al. 2015]. Our use of PSO for marker placement optimization aims to overcome the issues of suboptimal local minima often associated with non-global or greedy approaches.

In the PSO method, an optimal solution to a given problem is found by iteratively updating and evaluating candidate solutions, or solution hypotheses. A large set of such hypotheses is managed as a swarm or population of particles, each of which has an associated position \mathbf{x}_t and velocity \mathbf{v}_t in the solution domain of the objective

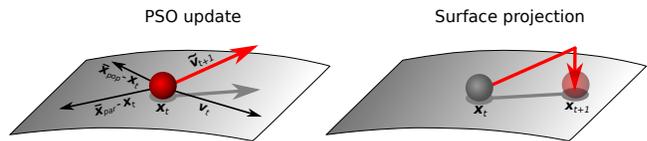


Figure 8: Illustration of a PSO update for one marker position. First, the new velocity $\tilde{\mathbf{v}}_{t+1}$ of the marker is computed as a weighted linear combination of the vectors towards the particle's local optimum $\bar{\mathbf{x}}_{par} - \mathbf{x}_t$, the population's global optimum $\bar{\mathbf{x}}_{pop} - \mathbf{x}_t$ and the particle's current velocity vector \mathbf{v}_t . This update can send the marker off the surface of the hand model due to the curvature of the model surface. Therefore, in a second step, the new position \mathbf{x}_{t+1} is computed by projecting back onto the surface. The velocity \mathbf{v}_{t+1} is then recomputed accordingly as $\mathbf{x}_{t+1} - \mathbf{x}_t$.

function at iteration t . Each particle keeps track of its local previous best position $\bar{\mathbf{x}}_{par}$ in the solution domain, and the population keeps track of the global optimum $\bar{\mathbf{x}}_{pop}$ across all particles.

In each iteration of the PSO process, the velocity of every particle is updated such that the particle is attracted to both the local and global optimum, while still moving according to its inertia. The local and global optima are updated after each particle movement by evaluating the objective function at the new particle position. Finally, the solution of the PSO process is the global optimum achieved after a given number of iterations or after convergence of the optimum value.

In our application, the solution domain of the objective function is the domain of marker layouts \mathcal{M} . To map this to the PSO scheme, we define a particle at iteration t as the stacked vector of k marker positions $\mathbf{x}_t \in \mathbb{R}^{3k}$ of the candidate solution. We further modify the generic PSO scheme such that the 3D positions within each particle are constrained to the surface of the hand model. Specifically, after every particle update we project each marker position in \mathbf{x}_t onto its spatially closest vertex in the set \mathcal{V}' of preselected feasible positions on the hand model.

The new position \mathbf{x}_{t+1} of a particle is determined by computing its new velocity \mathbf{v}_{t+1} and translating along this vector. To this end, we first compute the standard PSO velocity update as

$$\tilde{\mathbf{v}}_{t+1} = w \cdot (\mathbf{v}_t + c_1 \cdot r_1 \cdot (\bar{\mathbf{x}}_{par} - \mathbf{x}_t) + c_2 \cdot r_2 \cdot (\bar{\mathbf{x}}_{pop} - \mathbf{x}_t)), \quad (10)$$

where w is a weight determining the overall step length of the update, c_1 and c_2 are importance weights for the local and global attractors, respectively, and r_1 and r_2 are uniformly distributed random numbers in $[0, 1]$. Due to the curvature of the hand model surface, applying this linear update to the current particle position can cause the markers to stray from the surface. To counteract this, we project the updated marker positions back onto the permissible regions defined by vertices \mathcal{V}' , which we denote by a projection operator $\Pi_{\mathcal{V}'}$. The final particle position update is therefore

$$\mathbf{x}_{t+1} = \Pi_{\mathcal{V}'}(\mathbf{x}_t + \tilde{\mathbf{v}}_{t+1}). \quad (11)$$

After this, the new particle velocity is computed as $\mathbf{v}_{t+1} = \mathbf{x}_{t+1} - \mathbf{x}_t$. Figure 8 illustrates the surface-constrained PSO update.

Similar to Oikonomidis et al. [2011], we perturb one randomly chosen marker position in 50% of the particles once in every third iteration, and use the weights $c_1 = 2.8$, $c_2 = 1.3$ and $w = 2 / \left| 2 - \psi - \sqrt{\psi^2 - 4\psi} \right|$ with $\psi = c_1 + c_2$. We use a total of 1000 particles, perform 100 PSO iterations and use between 3 and 10 keyframes depending on the input hand motion trajectory. Using this method, we can find reduced marker layouts that optimize

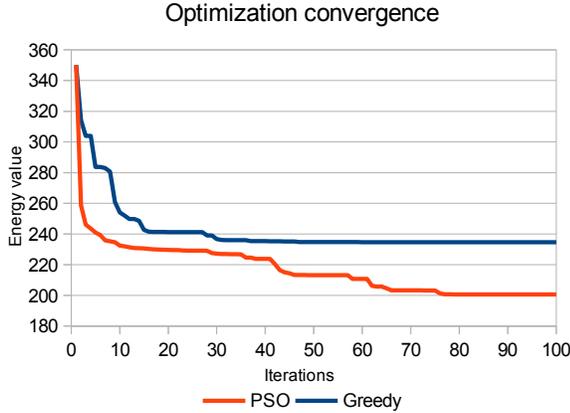


Figure 9: Comparison of the convergence behaviors our PSO energy minimization and a greedy approach with identical initialization. While the greedy approach converges to a suboptimal local minimum after about 50 iterations, our stochastic optimization minimizes the energy faster and achieves a better result.

the objective function (7) and as a result are numerically stable and geometrically feasible.

6 Results

We evaluated the convergence properties of our marker PSO scheme and the motion reconstruction accuracy of the marker layouts generated using our method in a varied set of evaluation trials. The hand movements involved in these trials included a variety of grasping and other manual interaction movements, as well as generic finger movements and gestures. In the performed trials, we measured run-time statistics and average per-vertex errors of the reconstructed hand motions compared to the ground truth input. For proper evaluation of the accuracy of our approach, the input motions being reconstructed were not contained in the database used to generate the subspace model. As our reduced marker sets are optimized to represent only rotational DoFs of the hand articulation, an initial estimate for the global position of the hand is given by a fixed anchor marker on the forearm.

We evaluate the convergence properties of our PSO-based marker layout optimization by comparing it to a more straightforward greedy approach. For this, we adapted the farthest point optimization scheme of [Schlömer et al. 2011] to find the marker subset of the initial vertex set \mathcal{V}' that minimizes the objective function (7). Briefly stated, this method first iteratively selects the next best vertex as a marker position that reduces the objective value until the desired number of markers has been placed. Then, this greedy process is repeated such that each selected marker position is replaced by the next better remaining vertex position, until no more substitutions can be done to improve the objective value. This is already a more sophisticated approach than the greedy methods for constraint selection used in [Loper et al. 2014; Thiery et al. 2012] and can therefore serve as an upper bound for the effectiveness of such methods. Figure 9 compares this greedy approach with our PSO-based one with identical initialization and shows that our method converges faster and achieves better objective values. The runtime for our PSO method varies between 5–10 seconds for 100 iterations, depending on the number of selected key-frames \mathcal{F} (typically up to 10). For the same problem setup, the greedy approach takes between 45 seconds and 3 minutes to converge.

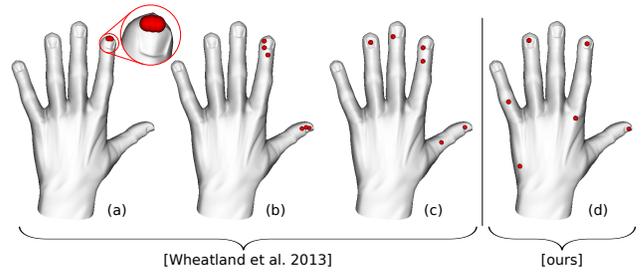


Figure 10: Comparison of 6-marker layouts generated for precision grasps using Wheatland et al. [2013] and our approach. In (a) the complete set of preselected vertices \mathcal{V}' is used as the base marker set, which causes the selected markers to cluster at the index fingertip, as it exhibits the most movement. In (b) a random subset with 5% of \mathcal{V}' is used as the base marker set, which leaves 3 candidate positions per joint. In this case the markers cluster around the index and thumb tips. In (c) 1% of \mathcal{V}' is used, which leaves one candidate position per joint. The resulting marker set is distributed among the most active joints in the input motion. In (d) our approach generates a marker layout from the complete set \mathcal{V}' based on the DoFs of our subspace model.

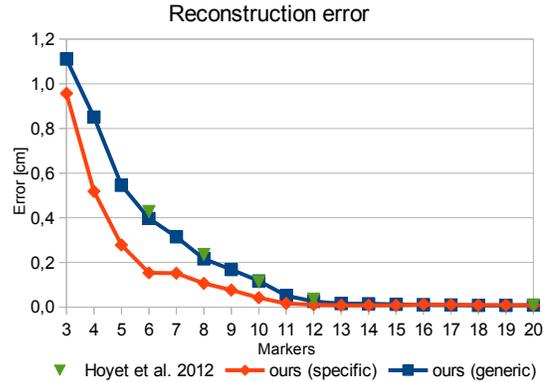


Figure 11: Average reconstruction errors for grasping motions, using marker sets of varying sizes. Marker layouts specifically generated from grasping training motions lead to lower reconstruction error than layouts based on generic training data. The manually selected marker layouts of Hoyet et al. [2012] produce similar results to our automatically generated generic layouts.

A comparison of our marker layout optimization with the marker subset selection approach of Wheatland et al. [2013] is shown in Figure 10. A crucial aspect to note regarding this comparison is that the two methods are based on different marker layout generation paradigms. While Wheatland et al. [2013] select the most influential markers in an initial base marker set, our method generates marker layouts more freely within the dense set of preselected vertices \mathcal{V}' (see Section 5). Figure 10 shows that the results of the subset selection method are strongly influenced by the choice of the base marker layout. As the method of Wheatland et al. [2013] is based on computing an importance ranking for the base markers according to their positional trajectories, the selected marker layouts are clustered around the areas of the hand that move the most in the considered hand motion. In contrast, our method is sensitive to the hand kinematics and the subspace model employed in our approach, which produces layouts that are well-suited for subspace-constrained IK.

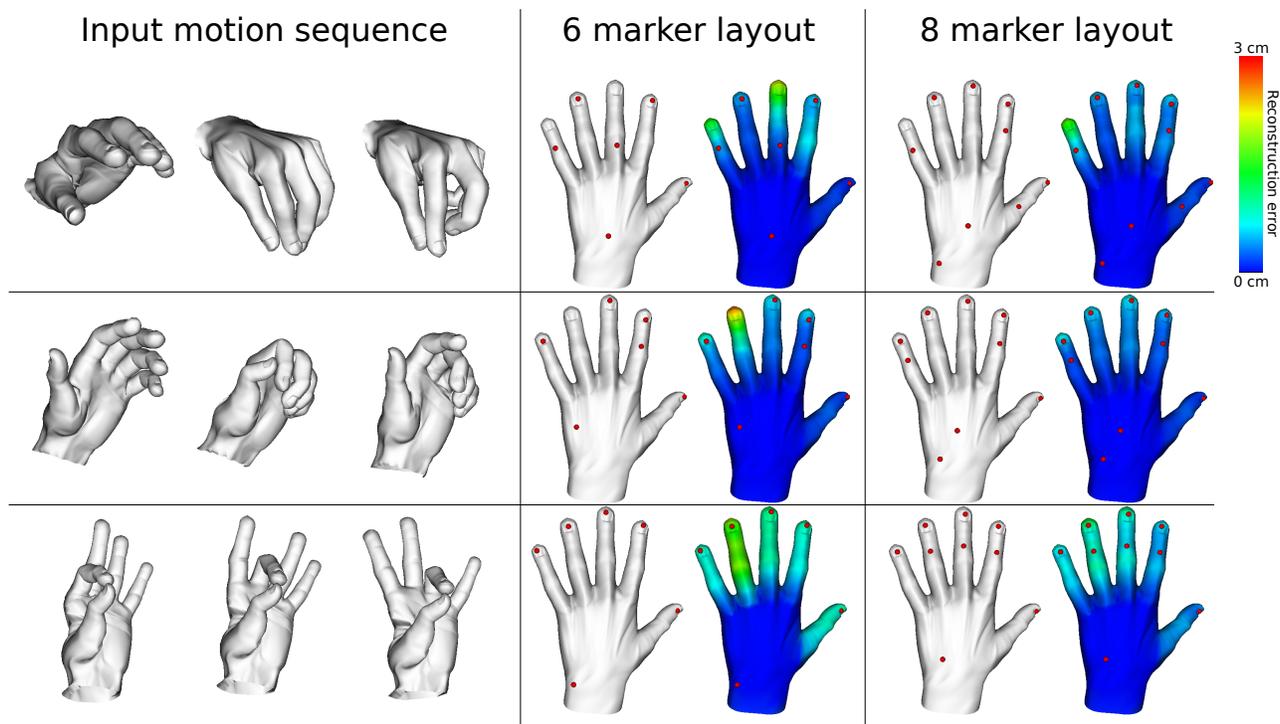


Figure 12: Reduced marker layouts for some example motions. First row: precision grasp motion involving multiple fingers. Second row: power grasp motion of a small object. Third row: sequence in which the thumb touches all the other fingers. The marker layouts are optimized to allow for accurate reconstruction of the input motion. Marker-less fingers tend to have slightly larger reconstruction errors, however they still move in correlation to the marker-constrained fingers due to the subspace approach.

To assess the suitability of our marker layouts for motion reconstruction, we compare the reconstruction error of different marker layouts in Figure 11. The testbed of this evaluation is a set of grasping motions based on the grasp taxonomy of [Cutkosky 1989]. We generated two different types of marker sets with varying sizes—a specific type based on grasping input motions and a generic type based on general gestures and hand articulations. The reconstruction error is lower when using the grasp-specific marker layouts. In particular, to achieve a reconstruction error below 2 mm, a specific layout generated by our method only requires 6 markers, whereas generic layouts require 9 markers or more. Additionally, Figure 11 also compares our automatically generated marker layouts with the empirically selected layouts of Hoyet et al. [2012], who also performed motion reconstruction based on constrained IK. Their manually selected layouts (see their Figures 4 and 8) produce similar results to our automatically generated *generic* layouts.

Figure 12 shows some examples for reduced marker layouts computed by our approach for various different movements. The results show that markers are preferentially placed in areas that have the most involvement in the considered hand motion. If the motions contain more varied articulations for specific fingers over others, these fingers will receive more markers, as the low-frequency details of the remaining markers are not influenced by as many subspace DoFs. In the third row of Figure 12, the input motion involves all fingers and the reduced marker layout accordingly distributes markers across all of them. This last example demonstrates that it is possible to combine multiple types of hand motions (e.g., each finger touching the thumb) and then solving for the corresponding marker layout. To this end, the PCA matrix \mathbf{M} in (3) and the subspace Jacobian \mathbf{J}_{PC} in (4) are simply computed from the union of the different sets of training data.

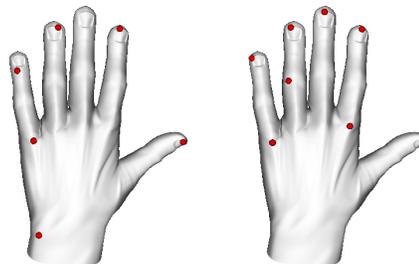


Figure 13: Generic 6-marker and 8-marker layouts generated from our full set of available training data.

Please see the accompanying video for comparisons of the full and reduced marker layouts using standard IK and subspace IK for reconstructing the hand motion sequences shown in Figure 12.

Although our method is primarily designed for generating *specific* marker layouts to be used for tracking specific types of hand movements (e.g., power grasp, precision grasp), we also generated two *generic* marker layouts by using all available training data from [Schröder et al. 2014] for constructing the posture subspace, which then includes a wide range of strongly varying hand motion trajectories. The resulting 6-marker and 8-marker layouts are shown in Figure 13. Since in the training data the middle finger moved mostly in correlation with other fingers, its movements can be inferred from the subspace, such that the 6-marker layout does not place a marker on it. The 8-marker layout interestingly is quite similar (up to the marker on the ring finger) to the empirically chosen 6F-2T layout of Hoyet et al. [2012] (see their Figure 8b).

Method	Average error	Maximum error
Standard IK	1.79 cm	7.9 cm
Subspace IK	0.89 cm	2.1 cm

Table 1: Reconstruction errors for subspace-IK and standard IK with a 4-marker layout generated for a variety of manual interaction motions. While the standard method can deviate by almost 8 cm, the subspace method achieves adequate results consistently.

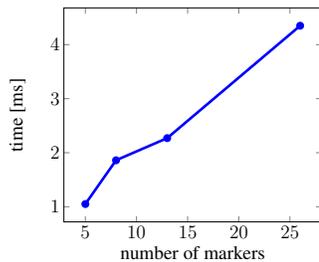


Figure 14: Per-frame computational time for tracking a typical (time-coherent) hand motion sequence, measured on an 8-core Intel Xeon(R) E5-1620 CPU at 3.60 GHz.

We verify the accuracy of our motion reconstruction by comparing the average reconstruction errors of subspace-IK to standard IK for a reduced marker layout produced by our method. Table 1 shows the average and maximum errors for a variety of manual interaction motions using standard IK and subspace-IK and the 4-marker layout shown in Figure 5. The improvement of the subspace method over the standard method ranges from 9 mm to almost 6 cm.

Although an increase in computational performance is not the primary motivation for employing sparse marker layouts, Figure 14 shows that the tracking time scales approximately linearly with the number of markers. Note, however, that even for an over-specified 26-marker layout the IK-based hand tracking requires less than 5 ms without any GPU-acceleration [Schröder et al. 2014].

7 Discussion

We have presented a method that automatically computes reduced marker layouts for optical motion capture by employing a subspace-constrained IK motion reconstruction. Our layout optimization minimizes an objective function that measures the numerical stability and geometric feasibility of the reduced marker configuration. This objective function is minimized using a specialized surface-constrained particle swarm optimization. The resulting marker layouts are suitable for solving the subspace-IK motion reconstruction, and are specific to the type of hand motions that should be expressed with and recovered from the sparse marker data.

Our method makes it possible to generate marker layouts that are fine-tuned to the parameters of a given mocap setup. If there is a limitation to the number of markers that can be used in the mocap setup, our method computes the optimal placements for the given number of markers that allows for realistic expressive motion reconstruction. An insight provided by our work is that it is sufficient for high quality motion reconstruction to place individual markers that correspond to low-dimensional control parameters of hand articulations. For instance, to track grasping motions it is sufficient to only place one marker on the thumb, index finger, pinky finger, and wrist. The subspace based reconstruction will plausibly interpolate the movements of joints that are not immediately constrained by markers.

Limitations of our approach include the stochastic nature of the particle swarm optimization and the need for parameter tweaking. Another drawback of our subspace-oriented method is that while it produces good results for specific hand movements, it does not necessarily provide a general-purpose marker layout result that can be used for all types of motions and produce high-quality results. The marker placement as well as the motion reconstruction are limited by the subspace priors employed. However, given prior knowledge of the motions intended to be tracked, our method produces accurate and robust results. Beyond marker placement, our approach could be used generally to identify salient regions in articulated bodies, which could be of interest for different avenues of motion detection and reconstruction.

Acknowledgments

This research was supported by the Cluster of Excellence Cognitive Interaction Technology CITEC (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

References

- BERNSTEIN, N. 1967. *The Co-ordination and Regulation of Movements*. Pergamon Press Ltd.
- BIOSYN, 2015. <http://www.biosynsystems.net/>.
- BUSS, S. R. 2004. Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods. Tech. rep., University of California, San Diego.
- CHAI, J., AND HODGINS, J. K. 2005. Performance animation from low-dimensional control signals. *ACM Transactions on Graphics* 24, 3.
- CHANG, L., POLLARD, N., MITCHELL, T., AND XING, E. 2007. Feature selection for grasp recognition from optical markers. In *Proc. of International Conference on Intelligent Robots and Systems (IROS)*.
- CUTKOSKY, M. 1989. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation* 5, 3.
- GUERRA-FILHO, G. B. 2005. Optical motion capture: Theory and implementation. *Journal of Theoretical and Applied Informatics* 12, 2.
- HOYET, L., RYALL, K., MCDONNELL, R., AND O’SULLIVAN, C. 2012. Sleight of hand: Perception of finger motion from reduced marker sets. In *Proc. of Symposium on Interactive 3D Graphics and Games*.
- JACOBSON, A., DENG, Z., KAVAN, L., AND LEWIS, J. 2014. Skinning: Real-time shape deformation. In *ACM SIGGRAPH Course Notes*.
- KANG, C., WHEATLAND, N., NEFF, M., AND ZORDAN, V. B. 2012. Automatic hand-over animation for free-hand motions from low resolution input. In *Proc. of ACM Motion in Games*.
- KATO, M., CHEN, Y.-W., AND XU, G. 2006. Articulated hand motion tracking using ICA-based motion analysis and particle filtering. *Journal of Multimedia* 1, 3.
- KENNEDY, J., AND EBERHART, R. 1995. Particle swarm optimization. In *Proc. of IEEE International Conference on Neural Networks*.

- KENNEDY, J., AND EBERHART, R. C. 2001. *Swarm Intelligence*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- KITAGAWA, M., AND WINDSOR, B. 2008. *MoCap for Artists: Workflow and Techniques for Motion Capture*. Focal Press.
- LE, B. H., ZHU, M., AND DENG, Z. 2013. Marker optimization for facial motion acquisition and deformation. *IEEE Transaction on Visualization and Computer Graphics* 19, 11.
- LIU, G., ZHANG, J., WANG, W., AND MCMILLAN, L. 2006. Human motion estimation from a reduced marker set. In *Proc. of Symposium on Interactive 3D Graphics and Games*.
- LOPER, M. M., MAHMOOD, N., AND BLACK, M. J. 2014. MoSh: Motion and shape capture from sparse markers. *ACM Transactions on Graphics* 33, 6.
- MAYCOCK, J., RÖHLIG, T., SCHRÖDER, M., BOTSCH, M., AND RITTER, H. 2015. Fully automatic optical motion tracking using an inverse kinematics approach. In *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*.
- MEYER, J., KUDERER, M., MÜLLER, J., AND BURGARD, W. 2014. Online marker labeling for fully automatic skeleton tracking in optical motion capture. In *Proc. of International Conference on Robotics and Automation*.
- MICROSOFT, 2010. Kinect. <https://dev.windows.com/en-us/kinect/hardware>.
- MULATTO, S., FORMAGLIO, A., MALVEZZI, M., AND PRATICCHIZZO, D. 2013. Using postural synergies to animate a low-dimensional hand avatar in haptic simulation. *IEEE Transactions on Haptics* 6, 5.
- OIKONOMIDIS, I., KYRIAZIS, N., AND ARGYROS, A. A. 2011. Efficient model-based 3D tracking of hand articulation using Kinect. In *British Machine Vision Conference*.
- OPTITRACK, 2015. <http://www.optitrack.com/>.
- ORGANICMOTION, 2015. <http://www.organicmotion.com/>.
- PHASESPACE, 2015. <http://www.phasespace.com/>.
- QIAN, C., SUN, X., WEI, Y., TANG, X., AND SUN, J. 2014. Real-time and robust hand tracking from depth. In *Proc. of Computer Vision and Pattern Recognition*.
- QUALISYS, 2015. <http://www.qualisys.com/>.
- SCHLÖMER, T., HECK, D., AND DEUSSEN, O. 2011. Farthest-point optimized point sets with maximized minimum distance. In *Proc. of ACM Symposium on High Performance Graphics*.
- SCHRÖDER, M., MAYCOCK, J., RITTER, H., AND BOTSCH, M. 2014. Real-time hand tracking using synergistic inverse kinematics. In *Proc. of International Conference on Robotics and Automation*.
- SHARP, T., KESKIN, C., ROBERTSON, D., TAYLOR, J., SHOTTON, J., KIM, D., RHEMANN, C., LEICHTER, I., VINNIKOV, A., WEI, Y., FREEDMAN, D., KOHLI, P., KRUPKA, E., FITZGIBBON, A., AND IZADI, S. 2015. Accurate, robust, and flexible real-time hand tracking. In *Proc. of ACM Conference on Human Factors in Computing Systems (CHI)*.
- TAGLIASACCHI, A., SCHRÖDER, M., TKACH, A., BOUAZIZ, S., BOTSCH, M., AND PAULY, M. 2015. Robust articulated-icp for real-time hand tracking. *Computer Graphics Forum* 34, 5.
- THIERY, J.-M., TIERNY, J., AND BOUBEKEUR, T. 2012. CageR: Cage-based reverse engineering of animated 3d shapes. *Computer Graphics Forum* 31, 8.
- VICON, 2015. <http://www.vicon.com/>.
- WHEATLAND, N., JÖRG, S., AND ZORDAN, V. 2013. Automatic hand-over animation using principle component analysis. In *Proc. of ACM Motion in Games*.
- WHEATLAND, N., WANG, Y., SONG, H., NEFF, M., ZORDAN, V., AND JÖRG, S. 2015. State of the art in hand and finger modeling and animation. In *Eurographics State of the Art Reports*.
- WU, Y., LIN, J. Y., AND HUANG, T. S. 2001. Capturing natural hand articulation. In *Proc. of International Conference on Computer Vision*.
- XSENS, 2015. <https://www.xsens.com/>.
- ZHENG, J., DE LA ROSA, S., AND DOLLAR, A. 2011. An investigation of grasp type and frequency in daily household and machine shop tasks. In *Proc. of International Conference on Robotics and Automation*.