Inverse Kinematics Based on High-Order Moments of Feature Points and Their Jacobian Matrices

Ko Ayusawa and Yoshihiko Nakamura

Abstract—In this paper, we propose the inverse kinematics method based on high-order moment features and their Jacobian matrices, which can use an arbitrary information source about the shape of the targeted kinematic chain as reference input. The method is especially useful to generate the motion of humanoid robots and human figures, and we can generate the whole body pose from a set of 3D markers or pixels of 2D images, without labeling each feature point. The moment features are computed from various types of quantities, for example, geometric points, mass density, pixel images, and the probability labeled as a specific link, the fact of which shows generality and versatility of the method. Some results of motion of a human figure from the label-less 3D markers and the 2D images are illustrated.

I. INTRODUCTION

Inverse kinematics (IK) is one of the important mathematical foundations in the field of robotics. The theory including non-linear and differential kinematics is widely applied to motion generation and control of many kinds of multi-body systems [1], [2], [3], [4], [5], [6]. The standard methods of IK calculate the generalized coordinates of a kinematic chain that achieve the desired position of the feature points located on some specific links and/or the orientation of the links. The IK of industrial manipulators usually requires the desired position and orientation only of the end effector. In the case of the large-scale kinematic chains such as humanoid robots, the desired position and orientation of several links, like the head, the hands and the feet, have to be determined. Since a humanoid robot is a structure-varying-system [7] and has no link fixed to the inertial frame, the 6 degree of freedom (DOF) of the floating base are among the generalized coordinates to be determined. The detailed model [5], [6] for human motion analysis requires more feature points placed close to some anatomical points. The more DOF the system possesses, the more feature points the IK needs to be defined and provided with their desired values.

The strategy to solve the generalized coordinates from the center of mass (COM) and its Jacobian matrix is often adopted for humanoid robots and human models [8], [9], [10], [11]. The COM is the first-order moments regarding mass property. In statistics, the second-order moments show the variance of stochastic variables, the third-order moments means the skewness, and the fourth-order moments imply the peakness, which are the high-order moment features of the geometric information. The high-order moment features are also used to recognize the shape of objects in the field of image recognition [12]. Some humanoid studies use not only the COM but also the total moment of inertia tensor [9] or the total angular momentum [11] for IK, which are closely related to the second-order moment features. Since only the COM and the second-order moment features are not enough to generate the whole body motion, the desired pose of several specific links are consequently required in those studies. Can we determine all the generalized coordinates of a kinematic chain only from the high-order moment features of whole geometry shape?

In this paper, we propose the IK method based on high-order moments and their Jacobian matrices. This method can use an arbitrary information source about the shape of the targeted kinematic chain as reference input. The moment features can be computed from various types of quantities, such as geometric points, mass density, pixel images, and the probability labeled as a specific link, which shows generality and versatility of the method. The method can readily be combined to the standard IK method setting simultaneous linear differential equations. Some results of the poses of a human model generated from label-free 3D markers and 2D images based on the high-order moment features are to be presented to show the performance of the proposed method.

II. INVERSE KINEMATICS BASED ON HIGH-ORDER MOMENTS

A. High-order moment feature and its Jacobian matrix

Let us consider the set of \( N \) feature points located on the kinematic chain with \( N_j \) DOF. And \( p_1, p_2, \ldots, p_N \in \mathbb{R}^3 \) are the position vectors of each feature point in the Cartesian space, and \( w_i \in \mathbb{R} \) \((1 \leq i \leq N)\) is the scalar parameter of the information of each feature point. If we consider only the information about geometric shape, \( w_i = 1 \). If we look at dynamics of the system, \( w_i \) means the mass of the feature point \( i \). In image processing, then \( w_i \) is a pixel of image data. If we have the probability that the point \( i \) is located on the link \( L_j \), \( w_i \) is the value of probability. \( w_i = 1 \) \((i \in L_j)\), \( w_i = 0 \) \((i \not\in L_j)\) are the special case when we have the knowledge that the point \( i \) is located on \( L_j \).

Now, we observe the feature points in the linear space \( S \) mapped from the Cartesian space, and \( s_i = f(p_i) \in S \subseteq \mathbb{R}^{N_s} \) is the position vector mapped from \( p_i \) in the Cartesian space. The examples of the function \( f \) are view coordinate transformation \((N_s = 3)\) and perspective transformation.
The mapping process of feature points from the Cartesian space to the Projective space (\(N_s = 2\)) is shown in Fig. 1. In the mapping process, we define the map \(p_i = f(p_i)\) of \(N \) feature points \(p_i\) with weight \(w_i\). The link \(L_j\) connects point \(s_j\) to \(p_i\). The mapped space \(R_s\) is shown in the figure with an example of mapped space \(R_s = R^{1 \times N_f}\). The desired value \(J_{s_1, x} = J_{s_{1, x}}\) of high-order moments at the position \(s_{1, x}\) is achieved by \(f(\mu)\). The inverse map \(J_m\) of \(J_s\) is \(J_m = J_{s, x}\) for \(x = \{x, y, z\}\). The Jacobian matrices \(J_{m, x}\) and \(J_{m, y}\) are Jacobian matrices of \(x, y, z\) with respect to \(q\).

\[
J_{s, x} = \begin{bmatrix} J_{s, x, 0} & J_{s, y, 0} & J_{s, z, 0} \\ J_{s, x, 1} & J_{s, y, 1} & J_{s, z, 1} \\ \vdots & \vdots & \vdots \\ J_{s, x, m} & J_{s, y, m} & J_{s, z, m} \end{bmatrix} \Rightarrow J_m = J_{s, x}^T T_{s, x}^T \Rightarrow J_m = J_{s, x}^T T_{s, x}^T
\]

where, \(m > 1\), \(N_m = m+2C_2\). If \(m = 1\), we choose \(J_{m, x} = J_{m, y} = J_{m, z} = 0\). As seen from Eq. (2) and Eq. (5), the moment \(m, x, y, z\) and the Jacobian matrix \(J_{m, x, y, z, m}\) of each feature point are computed in parallel. And \(J_{m, x, y, z, m}\) and \(J_{m, x, y, z, m}\) can be computed recursively as follow.

\[
J_{m, x, y, z, m} = \xi J_{m, x, y, z, m} + \eta J_{m, x, y, z, m} + \zeta J_{m, x, y, z, m} \Rightarrow J_m = J_{m, x, y, z, m}
\]

where, \(\xi, \eta, \zeta\) are the distances between point \(i\) and \(\xi, \eta, \zeta\) are the coordinates of \(x, y, z\) in \(R_s\) and \(\{x, y, z\}\) is an element of the index set.

Let us represent \(q \in R^{N_f}\) as the generalized coordinates of the kinematic chain, then we can define \(J_{m, x, y, z, m}\) on \(S\) with respect to \(q\) as follows.

\[
J_{m, x, y, z, m} = \frac{\partial J_{m, x, y, z, m}}{\partial q} = \frac{1}{W} \sum_{i=1}^{N} J_{i, m, x, y, z, m}
\]

where, \(J_{i, m, x, y, z, m} \in R^{1 \times N_f}\) is the high-order moment Jacobian matrix of \(\mu_i, m, x, y, z, m\) with respect to \(q\).
Let us define $e_m(q) \in \mathbb{R}^{N_m}$ as the error between $\mu_m(q) \in \mathbb{R}^{N_m}$ at $q$ and the desired moments $\mu_{m}^{ref} \in \mathbb{R}^{N_m}$ as below.

$$e_m(q) \triangleq \mu_{m}^{ref} - \mu_m(q)$$

(12)

Then, we compose $e \in \mathbb{R}^{N_\mu}$ of all $N_\mu(= \sum_{m=1}^{M} N_m)$ moment errors as follows.

$$e(q) \triangleq \begin{bmatrix} e_1(q) & \cdots & e_M(q) \end{bmatrix}^T$$

(13)

Beginning from certain $q_0$, we compute $q$ sequentially to minimize the evaluation function Eq.(14). $q_k$ and Eq.(14) are the solution and the evaluation function in the $k$-th process of optimization.

$$E_k = \frac{1}{2} r_k^T W_e r_k + \frac{1}{2} (q_k - q_{k-1})^T W_q (q_k - q_{k-1})$$

(14)

where, $r_k \triangleq \lambda e_{k-1} - J_{\mu,k-1}(q_k - q_{k-1})$, $\lambda_e \in \mathbb{R}$ is a positive scalar, $W_e \in \mathbb{R}^{N_e \times N_e}$ and $W_q \in \mathbb{R}^{N_q \times N_q}$ are the weight matrices and symmetric positive definite, and $e_k \triangleq e(q_k)$, $J_{\mu,k} \triangleq J_{\mu}(q_k)$ are the moments and Jacobian matrices in the $k$-th process. In the computation of $J_{\mu,k}$, the basic Jacobian matrix [14] can be used instead of $J_{x,i}, J_{y,i}, J_{z,i}$. As $q_{k-1}$ is the solution in the previous process and is constant in the $k$-th step, then we can compute $q_k$ as follows.

$$q_k = q_{k-1} + \lambda_e J_{\mu,k-1}^* e_{k-1}$$

(15)

where, $J_{\mu,k}^*$ is the singularity-robust inverse matrix weighted by $W_e$ and $W_w$, and can be defined as follows.

$$J_{\mu,k}^* \triangleq W_q^{-1} W_e J_{\mu,k} W_q^{-1} J_{\mu,k}^T + W_e^{-1}$$

(16)

Using Eq.(15) sequentially, we can obtain the convergent solution $q^*$. The conceptual diagram of the method is presented in Fig.2.

With the DOF of the system less than the number of moment features (i.e. $N_j < N_m$), no solution satisfies $\mu_m(q) = \mu_{m}^{ref}$. In this case, the solution of Eq.(15) does not satisfy $\mu_m(q) = \mu_{m}^{ref}$, but minimizes $\frac{1}{2} e_m^T W_m e_m$. For singular points or over-constraints, solving the non-linear equations $\mu_m(q) = \mu_{m}^{ref}$ is formulated by non-linear optimization. Especially in the cases of the large-DOF systems such as humanoid systems, non-linear optimization is often adopted for computational stability and practical safety [15], [16]. The IK based on the high-order moment features applies to such non-linear optimization problems.

It should also be noted that the moment feature can be computed from various types of quantities $w$ as mentioned above. And they are easily combined by solving the simultaneous linear differential equations. For example, if we use $L$ different features, we only solve the following equations.

$$\begin{bmatrix} \mu_{w^1} \\ \vdots \\ \mu_{w^L} \end{bmatrix} = \begin{bmatrix} J_{w^1} \\ \vdots \\ J_{w^L} \end{bmatrix} \dot{q}$$

(17)

where, the high-order moment based on $l$-th type of quantity $w^l$ is defined as $\mu_{w^l}$ and the Jacobian matrix as $J_{w^l}$. The use of various types of information is important to avoid the local minimum in the process of optimization. As seen from above, the method can also be combined to the standard IK method setting simultaneous linear differential equations, which is also important for the practical application.

III. EXAMPLES OF GENERATED POSE OF HUMAN MODEL

A. Results of inverse kinematics using label-less markers

This sub-section shows some results of motion of a human model generated by the proposed IK method. The human model is made of 34 DOF and 15 rigid links, and is shown in the left of Fig.3. The joint configuration and type are as follows: upper torso, lower torso, head, upper arms, lower arms, hands, thighs, shanks, and feet. The waist, the neck, the shoulders, the wrists, the hip joints and the ankles are modeled with spherical joints, and the elbows and the knees are modeled with rotational joints. It represents the most important DOF that are used in daily activities such as locomotion. In addition to the 34 DOF of its kinematics, 6 DOF are used to define the generalized coordinates of the
First, a walking motion data of the human model was prepared in advance by motion capturing. The motion data was sampled at 200Hz. The markers trajectories \( p_{M_i}(1 \leq i \leq 35) \) were computed from the motion data with the model and used as reference input of IK. It was checked whether the proposed IK could restore the original motion data from the reference trajectories. The validation, the restored trajectories \( \hat{p}_{M_i} \) were compared with the original references \( p_{M_i} \). In this validation process, no noise was added to the reference trajectories. It should be noted that the usual IK methods need the labels of markers: one-to-one correspondence between the reference trajectory and the marker of the model. Meanwhile the proposed IK method does not require the labels.

In the validation, the first to fifth order moments of feature points were chosen to be used in the proposed IK. In this case, the number of moment features is \( N_m = 55(= \sum_{i=1}^{35} i+2C_2) \). Although the number of moment features is larger than the DOF of the system (i.e. \( N_m > N_j \)), there always exists the solution to satisfy \( \mathbf{\mu}_m(q) = \mathbf{\mu}_{m}^{\text{ref}} \). It is because no noise was added to the original marker trajectories. The numerical solution of Eq.(15) is thus expected to converge \( \mathbf{\mu}_{m}^{\text{ref}} \), without thinking about numerical errors, singular points, and local minimums. In the validation, the parameters used in Eq.(15) were chosen as follows: \( \lambda_e = 1 \), \( W_q = 10^{-3} E \), and \( W_e \) was chosen as the diagonal matrix whose diagonal elements corresponding with \( m \)-th order moments are equal to \( \sigma^m \), where \( \sigma = ||\mathbf{\mu}_2|| \).

Fig.4 shows the mean errors of each distance between the referenced marker and the marker restored from the two IK methods (i.e. \( \sum_{i=1}^{35} ||p_{M_i} - \hat{p}_{M_i}|| / 35 \)). The black thin line shows the error of the proposed method using only moment features, and the gray thick line means the error of the usual IK using the labels of markers. The error of standard IK was less than 1mm and those of the proposed IK was less than 5mm, even when in the neighborhood of the singular points such as knee extension. Though the accuracy and convergence speed of the proposed method was inferior to the standard IK, the proposed method restored the original walking motion without labeling.

Secondly, another walking motion of a human subject was measured by a motion capture system. The positions of 35 reflective optical markers pasted on the body were captured by the system. It was checked whether the proposed method could recover the walking motion from the measured marker trajectories with noise. In this case, the solution may not satisfy \( \mathbf{\mu}_m(q) = \mathbf{\mu}_{m}^{\text{ref}} \), but minimizes \( \frac{1}{2} e_m^T W_m e_m \). The IK computation started with the two different initial generalized coordinates. In this validation, the parameters of Eq.(15) were changed for the computational stability as follows: \( \lambda_e = 0.2 \), \( W_q = 10^{-1} E \).

The generated motion is shown in Fig.5. As in the top of Fig.5, the walking motion was successfully recovered, when the initial pose of the human model was close to the first frame of the real motion. In the bottom of Fig.5, the initial pose of the model was back to front and looking at the opposite side, and so was his walking motion. Although the whole body motion of the 34 DOF human model was generated using only high-order moments of geometry, with only one type of moment feature the solution may converge to a local minimum, depending on the initial condition. Since the higher-order moments tend to contain the larger numerical errors, it is recommended to increase types of moment feature \( w_i \) or combine the standard IK method in practice, rather than choosing their higher moments.

\( B. \) Results of inverse kinematics using 2D perspective projection planes

In this sub-section, the IK based on the moment features on the 2D perspective projection planes was verified with the same human model, the same marker set, and the same reference trajectories \( p_{M_i} \) as in the previous section. Some camera viewpoints were located to compute the moment features of the markers on the projection plane of each camera. In this validation, occlusion of markers was not considered, and the parameters of Eq.(15) are the same as
in the previous sub-section. If we consider first to fifth order moments, $20(= 6 + 5 + 4 + 3 + 2)$ moments can be obtained from the 2D plane of one camera viewpoint. As the human model has 40 DOF, at least two cameras are required to determine all the DOF of the model. In this sub-section, at most three camera viewpoints were used to generate the pose of the model.

Fig.6 shows the mean errors of each distance between the referenced marker and the marker restored by the IK with the 2D moment features on the 2D planes. The gray long dashed thin line shows the error using 2D moments with three camera viewpoints, the black dotted log dashed thin line means the error with two viewpoints, and the black dotted thick line is with one viewpoint. Although the moments from one viewpoint, of course, could not reconstruct the walking motion, the moments from more than one viewpoint could restore the walking motion. The error with three viewpoints was in the same range of the error with 3D markers (i.e. the black thin line in Fig.4). The redundant camera viewpoints obtained the same performance as that with 3D moment features.

C. Results of inverse kinematics using 2D images

This sub-section shows some poses of the human model generated by the proposed IK methods using the moments of 2D input images. In this case, 14 feature points located on each joint center as shown in the right of Fig.3 were used instead of the 35 markers. Fig.7 shows the five input images and the poses of the human model computed from the proposed IK using the input images. In the process of computing the moments of those images, the images were binarized; $w_i = 1$ in the black or gray area of the images and $w_i = 0$ otherwise. In Fig.7 the whole poses of the human model fit the lines or the characters of the images. As seen above, the proposed IK can generate a whole body pose with 2D images that are free from the structure of the kinematic chain.

The proposed IK method was also combined with the standard IK method using the pin-and-drag interface [5]. For the implementation, the paint interface to draw 2D input images was added to the pin-and-drag system. In Fig.8, the position and orientation of both feet were pinned by the pin-and-drag interface, and drew the two 2D input images from the front and side views. For the IK computation, the position and rotation of both feet and the moment features of the two images were used as reference input. They were combined by solving the simultaneous linear differential equations. The whole pose followed the lines in the two images, with the position and orientation of both feet fixed.

In all the examples we saw in this section, the feature points were arbitrarily chosen. If the CG model of the system is available, we can use the vertices of a polygon mesh for the feature points. Noted that the results depend on the density of the feature points. For example, if the feature points are mostly located on the torso link, the movement of joints has little influence on the moment features. For using vertices of polygons as the feature points, we would need to control the weight $w_i$ to equalize the density of feature points; for example, $w_i = V_k/n_k$, where $V_k$ is the volume of polygon object $k$ where vertex $i$ exists, and $n_k$ is the number of vertices of link $k$. Fig.9 shows the generated pose when all the vertices of polygon meshes were used as the feature points with $w_i = V_k/n_k$, where the number of vertices of the model is 1464. The whole pose of the model fit the lines of the image. More detail analysis of the sensitivity of the IK to the feature point location is a future work.

IV. Conclusion

In this paper, we proposed the IK method based on high-order moment features and their Jacobian matrices. It can use an arbitrary information source about the shape of the
Fig. 7. Poses generated by the proposed method using high-order moment features of the geometric feature points. First, a walking motion of the human figure was prepared in advance, and the reference trajectories of the 3D markers pasted on the figure were computed without labeling and noise. It was checked whether the proposed IK could restore the original motion data from the reference trajectories. For the validation, the restored trajectories were compared with the original references, and the proposed method restored precisely the original motion. The walking motion could also be restored with the moment features of markers projected on the 2D perspective planes of several camera views. The performance of the proposed method was tested by generating the motion of a human figure only from the high-order moment features of the geometric feature points. The small figures show the input 2D images, and the large ones show the results of the proposed method.

We note that other geometric features can be also implemented in the same way using their Jacobian matrices. The merits using moment features are ease of implementation and low computational cost, which are useful for real-time computation of IK. Moreover, if the rigid body is approximated by a finite distribution of feature points, moment features obtained from them are closely related with inertial parameters [17], and have an important role when studying dynamics of humanoid systems. The proposed method has the potential to be a technology that can generate the pose of CG characters by simple 2D images or a technology of marker-less motion capture using video pictures.

**REFERENCES**


