Stable Whole-body Motion Generation for Humanoid robots to Imitate Human Motions

Seungsu Kim, ChangHwan Kim, Bumjae You and Sangrok Oh

Abstract—This work presents a methodology to generate dynamically stable whole-body motions for a humanoid robot, which are converted from human motion capture data. The methodology consists of the kinematic and dynamical mappings for human-likeness and stability, respectively. The kinematic mapping includes the scaling of human foot and Zero Moment Point (ZMP) trajectories considering the geometric differences between a humanoid robot and a human. It also provides the conversion of human upper body motions using the method in [1]. The dynamic mapping modifies the humanoid pelvis motion to ensure the movement stability of humanoid whole-body motions, which are converted from the kinematic mapping. In addition, we propose a simplified human model to obtain a human ZMP trajectory, which is used as a reference ZMP trajectory for the humanoid robot to imitate during the kinematic mapping. A human whole-body dancing motion is converted by the methodology and performed by a humanoid robot with online balancing controllers.

I. INTRODUCTION

A few humanoid robots have been developed and shown up to the public in the last decades, aiming to provide people with useful services. Most interaction among human beings happens through voice and gesture. Especially, using even simple gestures based on experiences, people can communicate with each other without serious confusions. When a humanoid robot serves a person, the interaction through gestures is essential and must be friendly to each other. Such gestures performed by the robot need to look as a human does, otherwise they may cause misunderstandings in the meaning of gestures. Due to the same reason, it is natural that a human-like motion of humanoid robot be comfortable to and acceptable by a human. Additionally, a humanoid robot may provide such entertainment as dancing or doing cute and joyful gestures. For this kind of purpose, the humanoid robot needs to be able to imitate the meaningful, artistic, and creative motions of human beings as close as possible without loosing the original meanings of motions.

It is one of the most fundamental functions for a humanoid robot to imitate an arm motions during the communication with a person through motions. The imitation of arm motions has been studied by several researchers. Pollard et al. in [2] developed a method to adapt captured human upper body motions to a humanoid robot. The method obtained the closest motions to the captured upper body motions of an actor, minimizing the posture differences between the humanoid robot and the actor. The limits of joint position and velocity were considered. Kim et al. in [1] also proposed a method to convert human arm motion capture data to the motions available to a humanoid robot using optimization. The position and orientation of hand and the orientation of upper arm of a human were imitated by a humanoid robot under bounded capacities of joint motors. Using this method human arm motions are converted, which will be discussed later in in the kinematic mapping process. Kim et al. in [3] suggested a mathematical representation to determine a human-like arm posture for a given position and orientation of human hand. The representation defined the Elbow Elevation Angle to characterize human-likeness in human arm motions. These studies above are all for the conversion of upper body motions.

Very few work on the conversion of human whole-body motions has been performed. Nakaoka et al. in [4] and [5] explored a procedure to let a humanoid robot (HRP-2) imitate a Japanese folk dance captured with a motion capture system. The entire dancing motion was first represented in terms of limited number of primitive motions. The position trajectories of joints were generated to imitate those primitive motions. These trajectories were then modified to satisfy mechanical constraints of the humanoid robot. Especially, for the dynamic stabilities the trajectory of pelvis was modified to be consistent with the desired ZMP trajectory. Their method however has the limitation of imitating very complicated ZMP trajectory of a human. Especially, for the case where the human stood on the ground with both feet, their method simply connected the ZMP trajectory from the center of one supporting foot to that of another supporting foot using third order polynomials. This may not imitate such complicated hip motions as oscillating motions (like going and coming back) between two feet. Those hip motions could be found quite easily in human motion capture data. Due to this reason, a new method to generate a ZMP trajectory for a humanoid robot to imitate a complex human motion directly, is necessary.

In this work, we suggest a off-line methodology to convert whole-body motion capture data of a human into the dynamically stable motion for a humanoid robot. The robot is then able to imitate a complex human motion as close to the original motion as he can. The proposed method consists of two processes, kinematic mapping and dynamic mapping. The kinematic mapping deals with converting human upper body motions and kinematically scaling human feet, pelvis and ZMP trajectories under the geometric differences between the robot and the human. The dynamic mapping is to make sure all the motions converted from the kinematic mapping are necessarily stable to perform. For this the pelvis motion of humanoid robot is modified iteratively. Prior to doing the two processes above, it is needed to obtain a human ZMP trajectory, which will be
imitated by a humanoid robot. To do this, the work proposes a simplified human model to obtain the approximated human ZMP trajectory. The next coming three sections will explore those three processes in detail. An example of dancing motion will be converted and performed by the humanoid robot built in KIST to evaluate the proposed methodology.

II. SIMPLIFIED HUMAN MODEL

The acquisition and analysis of a human’s Zero Moment Point (ZMP) is essential for a humanoid robot to imitate human motions. The force plates are generally used to acquire the ZMP trajectory of human movements easily and accurately by measuring reaction forces and torques against the ground. It may difficult to measure the entire ZMP of an actor, when the actor moves all around the room in which only few force plates cover certain areas. In addition, it costs a lot to cover the all floors with the force plates. Due to this, a method to acquire a ZMP trajectory of a human without the force plates is needed. We introduce a simplified human model to obtain a ZMP trajectory of a human based on the marker trajectories of the motion capture system attached on an actor.

The human body is nonhomogeneous. In other words, each part of the human body has different densities. People have different percentages of muscles and fat in their body such that the total density of each person is also different. Zatsiorsky in [6] estimated the inertial characteristics of human body segments for 100 male subjects. He acquired the average center of mass (CoM), mass and inertia of human body segments. It is however just the average values of human body segments. Some researchers in [7], [8], [9] and [10] developed human models. To complete their human models, they measured a human body directly. Provic et al. in [11] also developed a human model to achieve the realistic human mass distribution. They used human morphological data from the literature and direct measures on the human test subject. However, the measuring process was a harassing work. And it couldn’t express the distribution of mass correctly, because of nonhomogeneous of human body. Therefore a method to get the distribution of mass automatically is needed.

Fig. 1. Motion capture data and simplified human model shape

We simplified a human body with cylinders, spheres or boxes as seen in Fig. 1. The Head and hands are simplified using spheres. The trunk, pelvis, neck, upper-arms, fore-arms, upper-legs and lower-legs are modeled using cylinders. The foot is modeled using a box. The radii of all the cylinders and spheres and the heights of boxes are the unknown parameters to be determined. The heights of cylinders and the widths and lengths of boxes can be obtained from the motion capture data. Each parts of the model are connected to each other by 3 Degrees of Freedom (DOF) joints. To determine the unknown parameters systematically as well as realistically, an optimization with the motion capture data is used. An actor shakes his/her arm, leg, hip, head and trunk uniformly and consistently, standing on the force plate as seen in Fig. 1. The motion capture system captures the motion and the reaction force and torque. The optimization problem is to find the unknown parameters, minimizing the error between the ZMP trajectory obtained from the captured motion and reaction data and the approximated ZMP trajectory of the simplified human model as follows,

\[
\min f(b) = \int_0^T \| \mathbf{P}_{\text{appr}}(\mathbf{M}, b, t) - \mathbf{P}_{\text{actor}}(t) \|^2 dt \tag{1}
\]

where \(\mathbf{M}\) denotes a motion capture data set that contains the position trajectories of all the markers attached on the actor. \(b\) denotes the vector of unknown parameters for the simplified human model. The components of \(b\) are in order of the radii of cylinders for calf, thigh, hip, trunk, neck, upper-arm and fore-arm, the radii of spheres for hand and head, and the heights of boxes for feet. \(T\) is the termination time of the motion performed by the actor. \(\mathbf{P}_{\text{appr}}(\mathbf{M}, b, t) = \begin{bmatrix} p_{\text{appr}}^x & p_{\text{appr}}^y & p_{\text{appr}}^z \end{bmatrix}^T\) is the approximated ZMP by the simplified human model with the motion capture data of the actor. This ZMP can be easily computed using the equations in [12]. \(\mathbf{P}_{\text{actor}}(t)\) is the actual ZMP of the actor computed from the measured reaction forces and torques. The details refer to [?]. During optimization process, the density for simplified model is considered as a constant: (weight of an actor) / (initial volume of the simplified model).

From the optimal solutions of the problem, the unknown parameters are determined and the simplified human model is achieved. To evaluate the model, a simple whole-body motion was performed on force plates with the motion capture system. Figure 2 shows very good agreement between the approximated ZMP and the actual ZMP. In other words, the reasonably accurate ZMP trajectory of any motion of actor can be obtained from the marker trajectories without force plates, once the simplified human model for the actor is determined. Small errors exist in Fig. 2, which notices that the simplified human model could not reflect the delicate motion of actor perfectly. But the error is still small enough to neglect for this study.

III. KINEMATIC MAPPING

A humanoid robot needs to imitate an actor’s motions like dancing, bowing, sign language, etc. as close to the original motions as possible in the sense of keeping the meanings of the motions. For this motion similarity, this section discussed...
the kinematic mapping of actor’s motions to the humanoid robot in terms of upper-body and lower-body motions. In addition, the ZMP trajectory of actor contains his/her movement characteristics, especially for movement stability. A scaling of the ZMP trajectory to be consistent with the motions obtained from the kinematic mapping, is also explored. The simplified human model in the previous section is used to acquire the approximated ZMP trajectory of actor.

A. Mapping of upper-body motion

People use their arms, heads and waists to express their emotion and purpose more clearly during communication with others. The upper-body of actor usually provides more complicated motions to express the meaning of an action than the lower-body. To imitate human upper-body motions, the method developed by Kim et al. in [1] is used herein. In the method the geometric difference in the arm length is resolved by scaling the arm length of the humanoid with a constant based on the arm length ratio between the humanoid robot and the actor. The imitation of actor’s arm motion is then formulated as an optimization problem. The problem minimizes the error between the captured motion of actor’s arm and the approximated motion of humanoid robot’s arm subject to bounding the joint position and velocity. The optimal joint position trajectory is obtained discretely, since the optimization problem is solved at each data point. A cubic spline interpolation in [13] and [14] is used to smooth this discrete joint trajectory, providing the position and velocity values of joint as references during the real-time control.

B. Mapping of lower-body motion

Even if people fix their hands with a certain posture in the space, they may have the limited different poses of their arms, especially elbows. This is due to the redundancy that human’s arm has. Human legs however possess relatively less redundancy compared to arms. If the position and orientation of a foot is fixed, the pose of leg can be fixed mostly. Due to this, the inverse kinematics problem of determining the leg posture of a humanoid robot for the given position and orientation of foot is algebraically solved using the procedure in [7]. The mapping of lower-body motions mostly deals with the foot and pelvis motions like scaling the pose trajectory of foot and pelvis and detecting the contacting phase of foot against the ground.

First, as seen in Fig. 1 we acquire the position and orientation trajectories of actor’s pelvis and feet from the human motion capture data. We then determine the supporting phases of each foot according to its elevation from the ground. Once these supporting phases are obtained, the position and orientation of each foot are computed. It means that the foot prints of actor during the motion are determined. Herein we assume that the actor stand on the ground with at least one foot and that no slip occur between the foot and the ground during contact.

For the supporting phases by a single foot, we scale the human’s foot trajectory considering the geometric differences in the leg length and the hip joint distance between the actor and the humanoid robot. It should be noticed that the orientations of actor’s pelvis and swing foot are not scaled but kept as they are. only the positions of swing foot and pelvis are scaled by the equations as follows,

\[ s^\text{robot} = K s^\text{human} \]  \hspace{1cm} (2)
\[ r_x^\text{robot} = k_x r_x^\text{human} \]  \hspace{1cm} (3)
\[ r_y^\text{robot} = k_y r_y^\text{human} \]  \hspace{1cm} (4)
\[ r_z^\text{robot} = k_z (r_z^\text{human} - r_z^\text{init}) + r_z^\text{init} \]  \hspace{1cm} (5)

where \( s^\text{robot} \) and \( s^\text{human} \) are the position vectors of swing foot with respect to the coordinate at the supporting foot for the actor and the robot, respectively. \( K \) is a scaling diagonal matrix: \( K = \text{diag}[k_x, k_y, k_z] \). In this study, we set \( k_x = k_y = k_z = l^{\text{robot}} / l^{\text{human}} \). However, they can be redefined by a user. \( r = [r_x, r_y, r_z]^T \) is the pelvis position vector with respect to the same coordinate as for the swing foot, which is defined for the actor and the robot separately with the superscripts of `human` and `robot`. \( r_{\text{init}}^{\text{robot}} \) is the initial position value of z axis of robot pelvis, \( r_{\text{init}}^{\text{human}} \) is the maximum position value of z axis of actor’s pelvis trajectory, \( l \) is the leg length. The constant used for scaling the trajectories of swing foot and pelvis is defined as the ratio of robot’s leg length and actor’s. Different constants may however be used by users for other cases. When double supporting phase, foot trajectories are fixed as the first values of the double supporting phase. Because we assumed that slip is not occurred.

C. Mapping of ZMP trajectory

It has been well known by Vukobratovic in [15] that the ZMP is necessarily in some margin of the supporting area, unless a human or a humanoid robot may fall down. The ZMP trajectory of humanoid during the imitation of actor’s locomotion must be within the supporting area. Through the mapping of lower-body motion, the stable zone for the ZMP of humanoid robot needs to change, since the foot trajectory is scaled. Therefore, it is needed to modify and scale the ZMP trajectory of actor suitable for the humanoid foot trajectory scaled in the early subsection.

Nakaoka et al. in [5] had proposed the way to generate the ZMP trajectory for a humanoid robot to imitate a whole-body motion of human. Once they converted the foot prints for the humanoid robot from actor’s, they connected simply the centers of the feet with a third order polynomial during a double supporting phase. Due to simple such connection, their method may not be sufficient to imitate the complex ZMP.
trajectory of an actor. To keep the dynamically complicated characteristics of the actor, we use the projection of actor’s ZMP trajectory. To obtain the most reliable stability zone for the humanoid’s locomotion, the ZMP trajectory is projected to lines or points in x-y plane. During a double supporting phase, the ZMP trajectory of actor is projected to the line that connects the centers of both supporting feet. For a single supporting phase, the ZMP trajectory of actor is laid on the center of the supporting foot.

The ZMP trajectory of actor is projected to the line connecting the centers of both supporting feet. And it is also scaled simultaneously, since the relative position of both supporting feet is already scaled earlier. This projection and scaling is done using Eq. 6 as follows,

$$\mathbf{p}^\text{human} \cdot \mathbf{s}^\text{human} \cdot \mathbf{s}^\text{robot} = \mathbf{p}^\text{robot}$$  \hspace{1cm} (6)

where \(\mathbf{p}^\text{human}\) is the ZMP vector of actor with respect to the supporting foot during a double supporting phase. \(\mathbf{p}^\text{robot}\) is the ZMP vector for the humanoid robot projected and scaled from \(\mathbf{p}^\text{human}\).

Figure 3 shows an example of the mapping of a ZMP trajectory during a double supporting phase. The ZMP trajectory of actor in blue in the upper figure is scaled and projected to the blue line of the lower figure for the ZMP trajectory of humanoid robot. For example, some points on the human ZMP trajectory (the points a, b and c point on the upper figure of Fig. 3) are projected to the points a’, b’ and c’ in the lower figure for the humanoid robot using Eq. (6). For the case of switching a single supporting phase to double or reversely, the ZMP trajectory of humanoid robot obtained from the mapping may not be continuous. It may cause a discontinuous lower-body motion. For that case the trajectory is connected by third order polynomial equation with the time interval of 0.2sec, which can be defined by a user.

In this section, we introduced the kinematic mapping process to imitate human whole-body motion in the motion similarity point of view. The only use of the kinematic mapping is however not good enough to guarantee whether the converted motion is dynamically stable or not. A method to modify partially the kinematically mapped motion to ensure dynamic stability will be discussed in the following section.

IV. DYNAMIC MAPPING

The desired ZMP trajectory for the humanoid robot obtained in Sec. III-C may be inconsistent with the lower- and upper-body motions from the kinematic mapping, since no constraint on the dynamics of the humanoid robot is considered. This section then explore a method to achieve the dynamic stability for the humanoid robot. The motions modified through the procedure will correspond to the desired ZMP trajectory in the humanoid robot dynamics point of view.

The method aims to modify the x- and y-axis values of the pelvis position to satisfy the desired ZMP trajectory in in Sec. III-C, when the upper-body motion, the foot trajectory, the pelvis orientation and the z-axis value of pelvis position are obtained through the kinematic mapping process. It is then necessary to obtain the pelvis trajectory that satisfies the desired ZMP trajectory and the kinematically mapped motions mentioned above. For this, we first calculate the trajectory of center of mass (CoM) of the humanoid robot from the desired ZMP trajectory using an inverted pendulum model for simplicity. After that, we determine the x- and y-axis values of the pelvis position that satisfy the COM trajectory and kinematic mapped motion.

A. CoM trajectory generation from given ZMP trajectory

The x- and y-axis CoM trajectories can be calculated from a given ZMP trajectory, initial and final conditions of CoM and z-axis trajectory of CoM using an inverted pendulum model. This idea is similar to the fast motion pattern generation for a desired ZMP pattern proposed in [16] and [17]. We will introduce an easy and simple implement skill of that idea herein.

The relationship between CoM and ZMP of the inverted pendulum is

$$\mathbf{p}^\text{human} \cdot \mathbf{s}^\text{human} \cdot \mathbf{s}^\text{robot} = \mathbf{p}^\text{robot}$$  \hspace{1cm} (6)

Figure 3. ZMP conversion

Through the two equations above the CoM at \(t_{i+1}\) can be expressed in terms of the CoMs at \(t_i\) and \(t_{i-1}\).

$$c(t_{i+1}) = A(t_i) p(t_i) + B(t_i) c(t_i) - c(t_{i-1})$$  \hspace{1cm} (9)

where

$$A(t_i) = -\frac{\ddot{z}(t_i) + g}{\dot{z}(t_i)} \cdot \Delta t^2; \quad B(t_i) = -A(t_i) + 2$$  \hspace{1cm} (10)

By solving Eq.(9) iteratively, the CoM at \(t_i\), \(c(t_i)\), is obtained in terms of the CoMs at \(t_{i-1}\) and \(t_n\), as follows,

$$c(t_i) = \frac{c(t_n) - \sum_{k=1}^{n-i} (m(t_k) A(t_k) p(t_{n-k})) + m(t_{n-1}) c(t_{i-1})}{m(t_{n-i+1})}$$  \hspace{1cm} (11)
If the initial and final CoM, z-axis CoM trajectory and the entire ZMP trajectory are given, x- and y-axis CoM trajectory can be obtained. It is noticed that the first and final status of a motion is assumed to be steady. For this, we added 3 second of steady motion at the first and last. The time can be redefined by a user. Therefore the CoM and ZMP will be same at the first and final time frame from Eq.7.

## B. Modifying of pelvis trajectory for desired ZMP

Choi et al.[18] proposed the kinematic resolution method based on a CoM jacobian with an embedded motion. The method makes it possible to generate lower-body joint positions satisfying a desired CoM trajectory, an upper-body motion and a foot position and orientation in real-time.

**V. ONLINE BALANCING CONTROL**

The motion generated so far from the proposed methodology may not be performed by the humanoid robot as stably and exactly as we expect, since there are many disturbances in the robot like electric motors, harmonic drives, mechanical links mechanisms and so on. The robot may not act as the desired motion in real implementation. Due to this, a real-time balancing control algorithm is also required. We use three controllers, pelvis-ZMP controller, pelvis orientation controller and foot-landing controller.

The Pelvis-ZMP Controller is designed as a main balancing controller to chase the desired pelvis and ZMP trajectory. This controller is based on the CoM-ZMP walking controller suggested in [18]. The controller herein uses a variation of pelvis position instead of that of CoM position as follows,

\[
 u_i = q_i^d - k_{p,i} e_{p,i} + k_{q,i} e_{q,i}
\]

(13)

where \( u \) is the control input for the humanoid robot, \( q \) is the pelvis position and \( p \) is the ZMP position. The gains \( k_p \) and \( k_q \) obey the stable gain conditions given in [18]. The subscript, \( i \) for \( i = x \) or \( y \), denotes the x- and y-axis values of variables above.

Due to the existing flexibility of links, joints and harmonic drives, position and orientation errors in the swing foot and pelvis may occur. These errors could have serious influence on balancing control of the humanoid robot. We attached an IMU sensor at the pelvis center and designed a PI controller to maintain a desired pelvis orientation (both of roll and pitch) as

\[
 \varphi_{ref} = \varphi_d + k_{\varphi,p}(\varphi_d - \varphi) + k_{\varphi,l}\int(\varphi_d - \varphi)dt
\]

(14)

where \( \varphi \) can be the roll or pitch angle in the pelvis orientation measured from the IMU sensor, \( \varphi_d \) is the desired value of \( \varphi \). \( k_{\varphi,p} \) and \( k_{\varphi,l} \) are gains, which are experimentally determined.

In addition, the impedance controller in [12] is employed to observe the landing impact of a foot and it makes the humanoid robot land softly.

**VI. EXPERIMENT**

The motion of an actor captured with the motion capture system was generated using the proposed methodology. Figure 5 shows the ZMP trajectory and foot print of the actor and the humanoid robot, respectively, when the actor performs a simple dance motion. Figure 6 shows the ZMP and CoM trajectory of the motion for the humanoid robot. The snapshots of dancing motion performed by the KIST humanoid robot, *Mahru*, are given in Fig. 8, showing good agreements in the motion between the two characters. *Mahru* weighs 63Kg and is 150cm tall. Figure 7 shows the pelvis trajectories of the actor and the robot during first 9 second of the motion, as seen in the first 3 figures of Fig. 8. The pelvis trajectory of actor was scaled using Eq.4. The pelvis trajectory of humanoid robot...
was obtained using the two methods, the proposed method and the method by Nakaoka et al. in [4], for comparison. It is observed that the pelvis trajectory done by the proposed method is much closer to the scaled pelvis trajectory of actor than that from the Nakaoka’s method.

kinematic mapping an upper-body motion, a leg motion, a pelvis motion and the ZMP trajectory of a human was properly converted to the humanoid robot without loosing motion similarity. The dynamic mapping modified the pelvis trajectory to make sure the whole motion obtained from the kinematic mapping were dynamically stable. From these three processes we could generate a dynamically stable and kinematically similar motion to imitate a human motion. To resolve the disturbances in the robot system the three online controllers for balancing and soft stepping were used as well. As an example, a dancing motion was imitated by the KIST humanoid robot with good agreement.

Fig. 5. ZMP trajectory and Foot print of a human and humanoid robot

Fig. 6. CoM and ZMP trajectory of a Humanoid robot

VII. CONCLUSION

In this paper, we proposed a methodology for a humanoid robot to imitate a human whole-body motion. We introduced a simplified human model to obtain the ZMP trajectory of a human. This model enables the humanoid robot to imitate a complex human ZMP trajectory with ease. Through the

REFERENCES

Fig. 8. A performance done by MAHRU using the developed method.


