Fuzzy Controller based Biped Robot Balance Control using 3D Image

Sangbum Park, Youngjoon Han, Hernsoo Hahn

Abstract

In this paper, we propose a balance control algorithm based on fuzzy controller using 3D (Dimensional) geometric information obtained from sequential images. The pose of a biped robot is acquired from image information instead of different sensors, and this pose information is used for fuzzy controller to keep biped robot in a stable balanced condition. Initially, the camera mounted on the head of robot obtains the pose information of the target object while walking. We acquired corresponding points between characteristic features on the target object and CAD (Computer Aided Design) information in DB(Data Base) using SURF (Speeded Up Robust Features) algorithm. Three arbitrary features from SURF algorithm are able to generate a virtual plane, which represent the geometric information of the target object's pose data. The camera pose information for the target object is calculated using pose information derived from the virtual planes. As the camera pose information is related to pose of a biped robot, we can easily get the pose information of a biped robot. With this the difference between current and desired pose of a biped robot is obtained. This difference is used as input data for the fuzzy controller. The fuzzy controller keeps the biped robot in stable pose using the recent pose and velocity of a biped robot. The efficiency of the proposed algorithm has been proven by the experiments performed on even floor using a mini humanoid robot.

Keywords: Biped robot, Balance control, Zero Moment Point, 3D Photometry, FLC(Fuzzy Logic Controller).

1. Introduction

Vision based robot control techniques use feedback information extracted from vision sensors to control the robot motion. This has been one of the major research issues for more than three decades. The recent technological development facilitated the advancement in the areas which has resulted in a number of successful systems. Such applications are mostly concentrated on visual servoing of manipulators which controls the pose of end-effector to approach a specific trajectory by measuring the target trajectory using the visual information [1, 2]. Since visual servoing is used to control a manipulator based on analysis of images being provided in real-time, Chesi [3] proposed a method of using 6 instead of 8 feature points which have been required for the traditional visual servoing techniques to extract the 3D pose of the reference object. Radu [4] proposed a method of aligning the pose of the first manipulator (where a camera is mounted on its end-effector) to the pose of the second manipulator using the visual information provided by the camera. The pose of the first manipulator is estimated from the image feature points of the second manipulator’s gripper.

There are also few other examples that have applied visual servoing techniques for balance control of biped robots. Different from the visual servoing techniques used for the manipulators whose base frame is fixed to the ground, but those which are used in biped robots, should consider the fact that the COM(Center Of Mass) or the ZMP (Zero Moment Point) of robot is not static while robot is walking. In order to apply the visual servoing technique used for manipulator having 6 DOF (Degrees Of Freedom) to the positioning control of a biped robot, Yamamura [5] fixed the geometrical relation between the camera and the robot frames to reduce the degree of freedom of each leg to 4. This approach of adapting a visual servoing technique for manipulator to the biped robot limits the robot’s locomotion significantly. With stereo vision, we can see where the objects are in relation to our own bodies with much greater precision, especially when those objects are moving toward or away from us in the depth dimension. Stereovision has been studied long, and common computational approaches include feature-based [6], area-based methods [7]. All these methods have their intrinsic problems, caused by the very assumptions inherent in these approaches. These methods, by using multiple cameras such as stereo vision are able to get more accurate and stable 3D information than single camera method. But, it is not a more economic way of finding the depth of all objects or a view.
To acquire 3D information from the single camera method, the algorithm to make relation between current and registered features is required. There are two promising approaches to detect salient regions in images. The SIFT (Scale Invariant Feature Transform) algorithm of Lowe includes both a keypoint detector and descriptor [8]. This is achieved by repeatedly convolving the image with Gaussian function. The SURF (Speeded Up Robust Features) has been recently proposed by Bay [9]. Like SIFT, the SURF approach describes a keypoint detector and descriptor. However, the SURF is performing better than SIFT and produce more correct matches per time interval. So, the quality and total number of the created keypoints and their descriptors are slightly better for SIFT [10]. In this paper, the better runtime performance outweighs the feature quality, so, we chose to using SURF instead of SIFT.

In order to deal with nonlinearity in real environment effectively, a new method that applies Fuzzy Logic into PID controller is presented in this paper. Classical PID controllers use a similar method to determine the next output of the system. PID controllers have some difficulties in tuning for the whole range of working points. Mostly tuning procedure is done for a specific working point. When the working point changes their performance deteriorate and the errors at the output signal or the controlled variable tend to increase. The humanoid robot in dynamic unstructured environments requires processing of large sensory cognitive tasks without any explicit computations. Humans use perceptions of time, shape, and other attributes of physical environment [11-13]. In this paper, a balance control method based on the fuzzy controller is proposed.

The rest of this paper is organized as follows. In Section 2, the structural characteristics and the walking patterns of the biped robot based on 3D-LIPM (Linear Inverted Pendulum Mode) are illustrated. In Section 3, the camera model is given, and details regarding the pose of the robot with respect to the object frame is measured from the target object image acquired by the camera is presented. The relationship between the pose and COM is also explained in Section 3. In Section 4, the balance control scheme is described, which is based on the proposed fuzzy system and implementation details on a compact size humanoid robot. Fuzzy logic method is able to represent human expert’s knowledge and does not require the mathematical model of the target system. The performance of the proposed balance control scheme evaluated by the experiments described in Section 5. Finally, conclusions are drawn in Section 6.

2. Structural Characteristics of the Biped Robot

In this Section, the structural characteristics of the biped robot used in this paper are illustrated. Afterwards, the walking pattern generation scheme is described which simplifies the structural model of the robot as an inverse pendulum using the 3D-LIPM scheme and generates the symmetric and periodic walking patterns.

A. Structure of the Biped Robot

Each leg of the biped robot has 6 degrees of freedom, so that the foot may generate any pose in 3D space as shown in Figure 1.(a). There are 3 rotational joints on the thigh, 1 rotational joint on the knee and 2 rotational joints on the ankle. For simplifying the kinematics analysis, the number of DH (Denavit-Hartenberg) parameters is reduced by aligning the ankle joint and the pelvis joint axes. The robot has no joint in the waist to align the upper body and the camera axis and each arm has 3 degrees of freedom as shown in Figure 1.(b).

B. Walking Pattern Generation using 3D-LIPM

Biped robot requires complex dynamic modeling and analysis processes to analyze simple symmetric and periodic walking motion of a leg having six degrees of freedom. Along with many other research, this paper also uses the 3D-LIPM scheme to approximate the motion of the biped robot as that of the inverse pendulum as shown Figure 2. The pendulum is the mass point which represents the most dominant motion of the biped robot [14, 15].

Figure 1. Distribution of joints in the biped robot.

Figure 2. Motion representation of the biped robot using the 3D-LIPM.
One important feature of human walking is that the height of pelvis changes little. This feature helps to simplify the complex dynamic equation by limiting the height of the motion of mass point in 3D-LIPM to a specific plane. By the analysis of the dynamics of the 3D-LIPM scheme, the relation between the COM and ZMP can be represented by the following equations [16].

\[
\begin{align*}
-\frac{x}{g} \ddot{x} + x &= P_x \\
-\frac{y}{g} \ddot{y} + y &= P_y
\end{align*}
\]  

(1)

In equation (1), \((x, y)\) and \((\dot{x}, \dot{y})\) are the differential and acceleration of the COM in a biped robot in X and Y-axes respectively, \(g\) is the gravity acceleration, \((P_x, P_y)\) are ZMP in \((X,Y)\) plane, and \(z_c\) is the height of the COM which does not change while the robot moves.

The principle of dynamic walking [17], guides that a biped robot may walk stable as long as the projection of the ZMP of the robot on the floor points to the contacted region in the feet. From this principle, the trajectory of ZMP which the robot should satisfy for stable walking can be constructed from the trajectory of the feet. Figure 3 shows an example of ZMP trajectory having a form of sine wave generated considering the symmetry and periodicity of the trajectory of the feet. As shown below, LSS (Left Single Side), DS (Double Side), RSS (Right Single Side) are included as the states of robot.

![Figure 3. ZMP trajectory of biped robot approximated as a sine wave form.](image)

To determine the \(y\) axis value of the COM first, the trajectory of \(P_y\) having the angular velocity \(\omega_y\) is represented as follows:

\[
P_y = A \sin \omega_y t
\]

(2)

By applying (2), (1) is represented as the canonical non-homogeneous 2\(^{nd}\) order differential equation as follows:

\[
\ddot{y} - \frac{g}{z_c} y = -\frac{g}{z_c} A \sin \omega_y t
\]

(3)

Since the force function \(P_y\) in the above equation is a sine wave, the steady state solution \(y\) will also be represented as a function of the sine waves whose angular velocity is \(\omega_y\).

\[
y = A_i \cos \omega_y t + B_i \sin \omega_y t
\]

(4)

By applying this function to (4) and solving the equation with respect to \(y\), (5) is obtained as the steady state solution.

\[
y = \frac{g}{z_c \omega_y^2 + g} P_y
\]

(5)

The \(x\) trajectory of ZMP is designed for the biped robot to have the lower speed in order to decrease the moment of inertia when the robot foot is touching on the floor. In addition, the biped robot has the highest speed at the instance when the ZMP of the biped robot points into the double support area. As shown in Figure 3, the force function \(P_x\) has the sine wave with the double angle velocity \((2\omega_y)\) of \(P_y\) and the average velocity of the biped robot according to \(x\) coordinate is the slope of the function respect to time coordinate. The trajectory of \(P_x\) having the angular velocity \(\omega_x\) is represented as follows:

\[
P_x = \theta_\theta \left[A \sin \omega_y t\right]
\]

(6)

where, the operator \(\theta_\theta\) expresses the rotation of \(\theta\) with respect to time coordinate. The \(x\) position of COM can be determined as follows:

\[
x = \theta_\theta \left[\frac{g}{z_c \omega_x^2 + g} P_x\right]
\]

(7)

Equation (5) and (7) shows that the coordinates of COM is linearly proportional to the coordinates of ZMP when the motion of COM of the biped robot is restricted to a plane where \(Z = z_c\) and the trajectory of ZMP is approximated as a sine wave.

### 3. 3D Feature Extraction from 2D Image

#### A. 3D Feature Extraction based on SURF

SURF is based on sums of 2D Haar wavelet responses and makes an efficient use of integral images. The SURF detector is based on the Hessian matrix. Given a point \(x = [x, y]\) in an image \(I\), the Hessian matrix \(H(x, \sigma)\) in \(x\) at scale \(\sigma\) is defined as follows:
\[ H(x, \sigma) = \begin{bmatrix} L_{x\alpha}(x, \sigma) & L_{y\alpha}(x, \sigma) \\ L_{x\beta}(x, \sigma) & L_{y\beta}(x, \sigma) \end{bmatrix} \]  

(8)

where \( L_{x\alpha}(x, \sigma) \) is the convolution of the Gaussian second order derivative \( \frac{\partial^2}{\partial x^2} g(\sigma) \) with the image \( I \) in point \( x \), and similarly for \( L_{x\beta}(x, \sigma) \) and \( L_{y\beta}(x, \sigma) \). An example of one of this filters for the lowest scale analyzed as shown in Figure 4. Image convolutions with these box filters can be computed rapidly by using integral images [18].

![Figure 4. Gaussian second order derivative in xy-direction and corresponding box filter approximation.](image)

The location and scale of interest points are selected by relying on the determinant of the Hessian. Interest points are localized in scale and image space by applying non-maximum suppression in a 3×3 neighborhood. Finally, the local maxima found of the approximated Hessian matrix determinant are interpolated in scale and image space.

First of all, SURF constructs a circular region around the detected interest points in order to assign a unique orientation to the former and thus gain invariance to image rotations. The orientation is computed using Haar wavelet responses in both \( x \) and \( y \) directions. The dominant orientation is estimated and included in the interest point information. In a next step, SURF descriptors are constructed by extracting square regions around the interest points. The region is split up regularly into smaller 4×4 square sub-regions in order to retain some spatial information. The wavelet responses in horizontal and vertical directions are summed up over each sub-region and form a first set of entries to the feature vector. In order to bring in information about the polarity of the intensity changes, the underlying intensity pattern of each sub-region is described by a vector \[ V = \sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \]. This results in a descriptor vector for all 4×4 sub-regions of length 64, giving the standard SURF descriptor, SURF-64. On the other hand, the short descriptor with 3×3 sub-regions (SURF-32) that is still quite acceptable in comparison to other descriptors in the literature. Notice that the wavelet responses are invariant to a bias in illumination. Invariance to contrast is achieved by normalizing the descriptor vector to unit length.

An important feature of SURF is the fast extraction process, which takes profit of integral images and a fast non-maximum suppression algorithm. Also it is convenient, as the fast matching speed it permits, mainly achieved by a single step added to the indexing based on the sign of the Laplacian of the interest point. So, even though the quality and total number of the created keypoints and their descriptors are slightly better for SIFT, one has to pay for this with a disproportionate increase in computing time.

B. Homogeneous Matrix for the Virtual Planes

The proposed virtual plane method only use clearly detected features on the target object. These features can be constructed as virtual plane. We can consist a virtual plane from arbitrary 3 features in the feature set as shown Figure 5.

![Figure 5. Feature vectors for the object coordinates.](image)

The 3 feature vectors are based on the object coordinates. They are able to express on the virtual plane coordinates. We called the virtual plane which is consist of selected feature vectors \( \{F_0, F_1, F_2\} \). The feature vector \( F_0 \) is set as origin of the coordinates of virtual plane. The relationship between virtual plane \( \{P\} \) and object \( \{O\} \) coordinates is expressed as (9).

\[ P = H_O \cdot O \]  

(9)

The translation part of a homogeneous matrix is consisting elements of \( P_0 \), so we only concentrate to find rotational elements of the matrix.

![Figure 6. The rotational gap between virtual plane coordinates and x-y plane.](image)

(a) To parallel to the x-axis, and (b) To parallel to the y-axis

To parallel \( f_i \) into x-axis, we have to get different
angle on y and z-axes. The virtual plane coordinates are rotated by them. To parallel f_i into x-y plane, we only consider rotational gap (γ) on the x axis. This is shown as Figure 6. We can construct the (10) in the (9) using each rotational gap \( R = \{ \alpha, \beta, \gamma \} \).

\[
\rho \mathbf{H}_\alpha = \begin{bmatrix}
  Cac\beta & Cac\beta & 0 \\
  Sas\beta & Sas\beta & 0 \\
  -S\beta & S\beta & 0 \\
  0 & 0 & 1
\end{bmatrix}
\]

C. Calculation Camera Pose using Virtual Plane

Since the feature points are located in the x-y plane in the object frame, their z-coordinates are zero and they can be represented as \( \mathbf{P}_i = \{X_i, Y_i, Z_i, 1\} \). These points are represented as \( \{u_i, v_i\} \) in the computer image space and to derive inversely \( \mathbf{P}_i \) from \( \{u_i, v_i\} \) the following parallel (11), which is derived from (10).

\[
f \cdot x_i = \frac{R_{11}}{T_z} \cdot f \cdot y_i = \frac{R_{21}}{T_z} \cdot \frac{R_{31}}{T_z} \cdot x_i - \frac{x_i}{y_i} = \frac{R_{31}}{T_z} + \frac{f}{T_z} = x_i,
\]

\[
f \cdot x_i = \frac{R_{12}}{T_z} \cdot f \cdot y_i = \frac{R_{22}}{T_z} \cdot y_i - \frac{y_i}{x_i} = \frac{R_{22}}{T_z} + \frac{f}{T_z} = y_i.
\]

By inserting the coordinates of feature points in the object frame, the (11) is obtained which includes all the components of \( \mathbf{H}_\alpha \), and the goal matrix is further derived as (12).

\[
\mathbf{AK} = \mathbf{B}
\]

\[
\mathbf{A} = \begin{bmatrix}
  \beta_i \cdot X_i & \beta_i \cdot Y_i & 0 & 0 & -\left(u_i - u_0\right) \cdot X_i & -\left(u_i - u_0\right) \cdot Y_i & \beta_i \cdot X_i \\
  \beta_i \cdot X_i & \beta_i \cdot Y_i & 0 & 0 & -\left(u_i - u_0\right) \cdot X_i & -\left(u_i - u_0\right) \cdot Y_i & \beta_i \cdot Y_i \\
  0 & 0 & \beta_i \cdot X_i & \beta_i \cdot Y_i & -\left(v_i - v_0\right) \cdot X_i & -\left(v_i - v_0\right) \cdot Y_i & 0 \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  \beta_i \cdot X_i & \beta_i \cdot Y_i & 0 & 0 & -\left(u_i - u_0\right) \cdot X_i & -\left(u_i - u_0\right) \cdot Y_i & \beta_i \cdot X_i \\
  0 & 0 & \beta_i \cdot X_i & \beta_i \cdot Y_i & -\left(v_i - v_0\right) \cdot X_i & -\left(v_i - v_0\right) \cdot Y_i & 0 \\
\end{bmatrix}
\]

\[
\mathbf{K} = \begin{bmatrix}
  R_{11} & R_{12} & R_{13} & R_{14} \\
  R_{21} & R_{22} & R_{23} & R_{24} \\
  R_{31} & R_{32} & R_{33} & R_{34} \\
  R_{41} & R_{42} & R_{43} & R_{44}
\end{bmatrix}
\]

\[
\mathbf{B} = \begin{bmatrix}
  u_0 - u_0 \\
  v_0 - v_0 \\
  \vdots \\
  u_{i-1} - u_{i-1} \\
  v_{i-1} - v_{i-1}
\end{bmatrix}
\]

To calculate \( T_z \) in \( \mathbf{K} \) we use the property of the rotational matrix, that the column vectors are orthogonal matrix, that the column vectors are orthogonal and also condition that \( T_z \) is much larger than \( f \).

\[
T_z = f + \frac{1}{2} \left( \frac{1}{K_{11}^2 + K_{12}^2 + K_{13}^2} + \frac{1}{K_{21}^2 + K_{22}^2 + K_{23}^2} \right)
\]

In equation (13), \( K_i \) is the i-th elements of vector \( \mathbf{K} \) in the (12)). The solution of \( T_z \) is acquired from by averaging two calculated values, and the homogeneous transformation matrix \( \mathbf{H}_\alpha \) between camera frame and the object frame which can be calculated by it. We can easily obtain the geometrical relation between the camera and object coordinates as all features of the target object are in the same plane. However, most objects don’t have these features in this condition. So, we utilize the virtual plane method as described in section 3.B. The set of virtual planes include all the coordinates of target object. They can be expressed as shown in Figure 7.

Homogeneous matrices \( \mathbf{H}_\alpha \) in Figure 7 is the relation between world reference and object coordinate systems, and \( \mathbf{H}_\alpha \) is the relation between world reference and camera coordinates. The \( \mathbf{H}_\alpha \) is the known information, but \( \mathbf{H}_\alpha \) is doesn’t. To calculate this, we have to get relations between camera and every virtual plane.

The homogeneous matrix \( \rho \mathbf{H}_\alpha \) is about the relation between i-th virtual plane coordinates \( \{ \mathbf{P}_i \} \) and camera coordinates \( \{ \mathbf{C} \} \). From (12). The homogeneous matrix \( \mathbf{H}_\alpha \) mentioned earlier is the known information. This relation is obtained from CAD information for the target object. So, we can make a decision about the relationship between camera and object coordinates as \( \mathbf{H}_\alpha \).

We can generate (14), which is derived from (11).

\[
\begin{align*}
  X_{p_{ij}} &= f \cdot x_i \cdot X_{p_{ij}} + f \cdot y_i \cdot Y_{p_{ij}} + f \cdot z_i \cdot Z_{p_{ij}} + f \cdot \frac{T_i}{T_z} \\
  Y_{p_{ij}} &= f \cdot x_i \cdot Y_{p_{ij}} + f \cdot y_i \cdot Y_{p_{ij}} + f \cdot z_i \cdot Z_{p_{ij}} + f \cdot \frac{T_i}{T_z} \\
  Z_{p_{ij}} &= f \cdot x_i \cdot Z_{p_{ij}} + f \cdot y_i \cdot Z_{p_{ij}} + f \cdot z_i \cdot Z_{p_{ij}} + f \cdot \frac{T_i}{T_z} \\
\end{align*}
\]

\( \{X_{p_{ij}}, Y_{p_{ij}}, Z_{p_{ij}}\} \) and \( \{X_{p_{ij}}, Y_{p_{ij}}, Z_{p_{ij}}\} \) are about j-th virtual plane and i-th feature point in the target object from the current image. They also expressed as \( X_{p_{ij}}, Y_{p_{ij}}, Z_{p_{ij}} \) on the camera coordinate system. The i-th rotation elements of the homogeneous matrix between camera and target object coordinates is expressed as \( f \cdot \mathbf{R}_{ij} \). The translation on the x-axis is expressed \( T_x \). The rest translation data on the \( {x}, {y}, {z} \) is different.

\[
\begin{align*}
  \{X_{p_{ij}}, Y_{p_{ij}}, Z_{p_{ij}}\} &= f \cdot x_i \cdot X_{p_{ij}} + f \cdot y_i \cdot Y_{p_{ij}} + f \cdot z_i \cdot Z_{p_{ij}} + f \cdot \frac{T_i}{T_z} \\
  \{X_{p_{ij}}, Y_{p_{ij}}, Z_{p_{ij}}\} &= f \cdot x_i \cdot X_{p_{ij}} + f \cdot y_i \cdot Y_{p_{ij}} + f \cdot z_i \cdot Z_{p_{ij}} + f \cdot \frac{T_i}{T_z} \\
  \{X_{p_{ij}}, Y_{p_{ij}}, Z_{p_{ij}}\} &= f \cdot x_i \cdot X_{p_{ij}} + f \cdot y_i \cdot Y_{p_{ij}} + f \cdot z_i \cdot Z_{p_{ij}} + f \cdot \frac{T_i}{T_z}
\end{align*}
\]
unstable condition while walking in uneven surfaces or among the frames defined in Figure 6.

Therefore, the COM of the biped robot can be represented by (17) from the transformation relationship. 

\[
\mathbf{H}_m = \mathbf{H}_o \left( \mathbf{H}^T_o \mathbf{H}_o \right)^{-1} 
\]

(17)

The region of ZMP of the biped robot may result into an unstable condition while walking in uneven surfaces or because of external physical attacks. The normal method to keep a balance of the robot uses this distance to show how far from the center of the stable region. This distance and the velocity of ZMP are used as input data to the fuzzy controller. The fuzzy controller will generate the ZMP to recover a balance of the robot. This ZMP is converted to the COM of the robot using (7). These flow is shown in Figure 9.

\[
\begin{bmatrix}
\mathbf{f}_s x_{p_1} \\
\mathbf{f}_s y_{p_1} \\
\mathbf{f}_s z_{p_1} \\
\vdots \\
\mathbf{f}_s x_{p_i} \\
\mathbf{f}_s y_{p_i} \\
\mathbf{f}_s z_{p_i} \\
\vdots \\
\mathbf{f}_s x_{p_d} \\
\mathbf{f}_s y_{p_d} \\
\mathbf{f}_s z_{p_d} \\
\end{bmatrix} = 
\begin{bmatrix}
\mathbf{f}_s x_{Y_{p_1}} \\
\mathbf{f}_s y_{Y_{p_1}} \\
\mathbf{f}_s z_{Y_{p_1}} \\
\vdots \\
\mathbf{f}_s x_{Y_{p_i}} \\
\mathbf{f}_s y_{Y_{p_i}} \\
\mathbf{f}_s z_{Y_{p_i}} \\
\vdots \\
\mathbf{f}_s x_{Y_{p_d}} \\
\mathbf{f}_s y_{Y_{p_d}} \\
\mathbf{f}_s z_{Y_{p_d}} \\
\end{bmatrix} = 
\begin{bmatrix}
\mathbf{f}_s x_{Z_{p_1}} \\
\mathbf{f}_s y_{Z_{p_1}} \\
\mathbf{f}_s z_{Z_{p_1}} \\
\vdots \\
\mathbf{f}_s x_{Z_{p_i}} \\
\mathbf{f}_s y_{Z_{p_i}} \\
\mathbf{f}_s z_{Z_{p_i}} \\
\vdots \\
\mathbf{f}_s x_{Z_{p_d}} \\
\mathbf{f}_s y_{Z_{p_d}} \\
\mathbf{f}_s z_{Z_{p_d}} \\
\end{bmatrix} = 
\begin{bmatrix}
\mathbf{f}_s x_{p_1} - u_{p_1} - u_{y_{p_1}} \mathbf{x}_{p_1} - (u_{p_1} - u_{y_{p_1}}) \mathbf{y}_{p_1} - (u_{p_1} - u_{z_{p_1}}) \mathbf{z}_{p_1} \\
\mathbf{f}_s y_{p_1} - v_{p_1} - v_{y_{p_1}} \mathbf{x}_{p_1} - (v_{p_1} - v_{y_{p_1}}) \mathbf{y}_{p_1} - (v_{p_1} - v_{z_{p_1}}) \mathbf{z}_{p_1} \\
\mathbf{f}_s z_{p_1} - w_{p_1} - w_{y_{p_1}} \mathbf{x}_{p_1} - (w_{p_1} - w_{y_{p_1}}) \mathbf{y}_{p_1} - (w_{p_1} - w_{z_{p_1}}) \mathbf{z}_{p_1} \\
\vdots \\
\mathbf{f}_s x_{p_i} - u_{p_i} - u_{y_{p_i}} \mathbf{x}_{p_i} - (u_{p_i} - u_{y_{p_i}}) \mathbf{y}_{p_i} - (u_{p_i} - u_{z_{p_i}}) \mathbf{z}_{p_i} \\
\mathbf{f}_s y_{p_i} - v_{p_i} - v_{y_{p_i}} \mathbf{x}_{p_i} - (v_{p_i} - v_{y_{p_i}}) \mathbf{y}_{p_i} - (v_{p_i} - v_{z_{p_i}}) \mathbf{z}_{p_i} \\
\mathbf{f}_s z_{p_i} - w_{p_i} - w_{y_{p_i}} \mathbf{x}_{p_i} - (w_{p_i} - w_{y_{p_i}}) \mathbf{y}_{p_i} - (w_{p_i} - w_{z_{p_i}}) \mathbf{z}_{p_i} \\
\vdots \\
\mathbf{f}_s x_{p_d} - u_{p_d} - u_{y_{p_d}} \mathbf{x}_{p_d} - (u_{p_d} - u_{y_{p_d}}) \mathbf{y}_{p_d} - (u_{p_d} - u_{z_{p_d}}) \mathbf{z}_{p_d} \\
\mathbf{f}_s y_{p_d} - v_{p_d} - v_{y_{p_d}} \mathbf{x}_{p_d} - (v_{p_d} - v_{y_{p_d}}) \mathbf{y}_{p_d} - (v_{p_d} - v_{z_{p_d}}) \mathbf{z}_{p_d} \\
\mathbf{f}_s z_{p_d} - w_{p_d} - w_{y_{p_d}} \mathbf{x}_{p_d} - (w_{p_d} - w_{y_{p_d}}) \mathbf{y}_{p_d} - (w_{p_d} - w_{z_{p_d}}) \mathbf{z}_{p_d} \\
\end{bmatrix}
\]

Equation (16) can be expressed after deriving from (12).

4. Balance Control using Fuzzy Logic Controller

As shown in Figure 8, the camera frame is fixed on the top of the robot frame, which can be automatically determined once the pose of the camera is measured. Thus, the balance control of the biped robot becomes a problem while determining the stable pose of the robot while walking.

![Figure 8. The relationships among the frames as defined in the working environment.](image)

For generating the stable robot walking pattern, this paper uses the ZMP trajectory as explained in Section 2, which calculates the COM trajectory. As the biped robot is used for our task, it is structured so as to have the COM coincided with the center of the pelvis, the biped robot has a small magnitude of error in the position difference between them which may be ignored [16]. Therefore, the COM of the biped robot can be represented by (17) from the transformation relationship among the frames defined in Figure 6.

\[
\begin{align*}
\mathbf{H}_m &= \mathbf{H}_o \left( \mathbf{H}_o^T \mathbf{H}_o \right)^{-1} \\
&= \mathbf{H}_o^T \mathbf{H}_o \mathbf{H}_m
\end{align*}
\]

(17)

The region of ZMP of the biped robot may result into an unstable condition while walking in uneven surfaces or

![Figure 9. Intelligent fuzzy compensation system.](image)

Initially, we obtain the ZMP from the camera image. The difference between measured ZMP, desired ZMP, and the velocity of measured ZMP are inputs to the fuzzy controller. This distance is in the Euclidian geometry. The recovered ZMP from FLC is converted to the COM using the converter from ZMP domain to the COM domain in the inverse kinematics module. In addition, the inverse kinematics module calculates each joint angle gaps and they are transferred to the robot control module. The robot control module is converting from each joint angle to each joint variable. The FLC input/output membership functions are shown in

![Figure 8. The relationships among the frames as defined in the working environment.](image)

![Figure 9. Intelligent fuzzy compensation system.](image)
Due to the size of the robot sole which is 80mm × 48mm, the size of stable region is fixed to 50mm × 28mm.

Table 1. Fuzzy Rules.

<table>
<thead>
<tr>
<th>ZMP</th>
<th>NB</th>
<th>NM</th>
<th>NS</th>
<th>ZO</th>
<th>PS</th>
<th>PM</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After the defuzzification process a compensated ZMP is added to the original ZMP. The membership functions NB (negative big), NM (negative middle), NS (negative small), ZO (zero), PS (positive small), PM (positive middle), PB (positive big) can be depicted as shown in Table 1. In this case, the most commonly used COA (Center Of Area) method that generates the center of gravity of the possibility distribution of a fuzzy set is considered.

5. Experimental Evaluation

A. Measurement of the Camera Pose from a Target Object

To acquire the statistical characteristics of the measurement error, the distance and orientation of the target object are measured repetitively using the camera mounted on the head of the robot and the average and variance of the measurement errors are calculated. Initially the camera frame was located at (0mm, 0mm, 350mm) and rotated by (-90deg, 90deg, 0deg) with reference to the world frame. The orientation was measured by rotating the lens axis about the individual axes by 15deg from -30deg to +30deg, and the distance was measured by moving 15mm each step from a distance of 850mm toward the target object along the x-axis as shown in Figure 12.

Table 2 summarizes the average and variance error measurements included in the orientation and distance measurements.
Table 2. Error characteristics include in the orientation and distance measurements.

(a) Error included in the orientation measurement.

<table>
<thead>
<tr>
<th>Rotation axis and angle [deg]</th>
<th>Average[deg]</th>
<th>Variance[deg²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+30</td>
<td>0.3839</td>
<td>1.716</td>
</tr>
<tr>
<td>+15</td>
<td>0.1645</td>
<td>1.478</td>
</tr>
<tr>
<td>-15</td>
<td>0.1537</td>
<td>1.796</td>
</tr>
<tr>
<td>-30</td>
<td>0.2483</td>
<td>1.608</td>
</tr>
<tr>
<td>Pitching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+30</td>
<td>0.4505</td>
<td>1.664</td>
</tr>
<tr>
<td>+15</td>
<td>0.1959</td>
<td>1.416</td>
</tr>
<tr>
<td>-15</td>
<td>0.1256</td>
<td>1.030</td>
</tr>
<tr>
<td>-30</td>
<td>0.4130</td>
<td>1.144</td>
</tr>
<tr>
<td>Yawing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+30</td>
<td>0.2499</td>
<td>1.574</td>
</tr>
<tr>
<td>+15</td>
<td>0.0904</td>
<td>1.819</td>
</tr>
<tr>
<td>-15</td>
<td>0.1705</td>
<td>1.788</td>
</tr>
<tr>
<td>-30</td>
<td>0.2520</td>
<td>1.104</td>
</tr>
<tr>
<td>No rotation</td>
<td>0.1237</td>
<td>1.149</td>
</tr>
<tr>
<td>Average</td>
<td>0.2415</td>
<td>1.511</td>
</tr>
</tbody>
</table>

(b) Error included in the distance measurement.

<table>
<thead>
<tr>
<th>Rotation axis and angle [mm]</th>
<th>Average[mm]</th>
<th>Variance[mm²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>700</td>
<td>2.799</td>
<td>3.117</td>
</tr>
<tr>
<td>750</td>
<td>1.540</td>
<td>4.046</td>
</tr>
<tr>
<td>800</td>
<td>0.924</td>
<td>2.128</td>
</tr>
<tr>
<td>850</td>
<td>0.984</td>
<td>2.185</td>
</tr>
<tr>
<td>900</td>
<td>1.299</td>
<td>3.242</td>
</tr>
<tr>
<td>950</td>
<td>2.098</td>
<td>3.889</td>
</tr>
<tr>
<td>Average</td>
<td>1.607</td>
<td>3.101</td>
</tr>
</tbody>
</table>

The average and variance of the rotation measurement errors are 0.2415° and 1.511°² respectively. In Table 1.(a), the error occurring in case of rolling appears smaller than those occurring in the other cases, since the rotation is about the optical axis and thus there is no distortion included. The average and variance of the distance measurement errors are 1.607 mm and 3.101 mm² respectively. These statistical error characteristics are used in the fuzzy controller to estimate the ZMP of the robot in the experiments described in sub section 5.2.

B. Robot Walking on a Flat Floor without Obstacle

The robot initially located at (0 mm, 0 mm, 230 mm) in the world frame and walks towards the reference object along x-axis of the robot frame which coincides with the z-axis of the reference object frame. The performance of the balance control algorithm is tested by measuring how accurately the robot follows the desired ZMP trajectory. Figure 13 shows four trajectories, acquired when the robot walk with an optimal speed of 20 mm/sec.

The first one is the ZMP trajectory without using fuzzy controller, the second one is obtained by fuzzy controller, and the third one is the error signal between the desired ZMP trajectory and the actual ZMP trajectory shown in Figure 13.(b).

The performance of the proposed fuzzy controller is shown in Figure 13.(d). The experimental results have shown that the balance control using only camera is successful since the robot follows the desired ZMP trajectory within a reasonably small error boundary. The average error was 2.24 mm without using fuzzy controller and it was 0.74 mm while using the fuzzy controller. Since the accuracy of balance control can be affected by the periodic time of commands along with the walking speed, the algorithm was tested while robot walking with the maximum speed of 40 mm/sec. This can be achieved by adjusting the periodic time of commands, which is represented as the sum of image processing and servo control times as given in (20).

Command period (50 ms) = Image processing(40 ms) + Servo control(10 ms) (20)

The maximum periodic time of commands (cycle) is 50 ms (20 Hz) was empirically determined on the condition that the frame rate of the camera is 30 fps. At the maximum periodic time of command, the walking speed of the robot becomes 40 mm/sec. Experimental results acquired when the robot walks with the maximum speed are shown in Figure 14. As expected, the measurement errors increased to twice of those acquired when the robot walks at the optimal speed.

The average errors are displayed as shown in Figure 14. The ZMP error of the proposed fuzzy control with respect to y-axis was 1.65 mm, but the general control had the ZMP error of 3.28 mm.

6. Conclusion

In general, biped robot demands several sensors for effective balance control. This paper proposes fuzzy controller based biped robot balance control using camera image. The pose of biped robot is thus determined without using any dynamic sensor. The 3D pose for the biped robot can be acquired using the proposed virtual plane method even in cases of limited feature information or image having noise. The pose of the robot at each sampling time is determined using fuzzy controller, so that the current ZMP of the robot may be located at the desired ZMP positioned inside the stable region of the robot. The experimental results have shown that the ZMP trajectory generated by the fuzzy controller, where the current ZMP is determined only using the visual information follows the desired trajectory very closely. The proposed method provides a cost-effective solution to maintain the biped robot balance.
Figure 13. Acquired ZMP trajectories using the proposed algorithm are compared with the reference trajectory when the robot walks with the speed of 20 [mm/sec].

Figure 14. Acquired ZMP trajectories using the proposed algorithm are compared with the reference trajectory when the robot walks with the speed of 40 [mm/sec].
Acknowledgment

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)"(NIPA-2009-(C1090-0902-0007)), and this work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD).

References


Sangbum Park received the B.S. and M.S. degrees from Electronic Engineering department of Soongsil university, Seoul, South Korea, in 2004 and 2006 respectively. He has been with School of Electronic engineering, Soongsil university since 2006, where he is currently pursuing Ph.D. His current research interests include biped walking robot, robotics vision.
Youngjoon Han received his B.S., M.S. and Ph.D. degrees in Electronic Engineering from Soongsil university, Seoul, Korea, in 1996, 1998, and 2003, respectively. He is currently a assistant professor in the School of Electronic Engineering at Soongsil university. His research interests include robot vision system, and visual servo control.

Hernsoo Hahn received the B.S. and M.S. degrees in Electronic Engineering at Soongsil university and Younsei university, Korea in 1982 and 1983 respectively. He received Ph.D degree in Computer Engineering from University of Southern California in 1991, and became an assistant professor at the School Electronic Engineering in Soongsil university in 1992. Currently, he is a professor. His research interests include application of vision sensors to mobile robots and measurement systems.