Design and Control of the Twente Humanoid Head

L. C. Visser, R. Carloni and S. Stramigioli

***Abstract*— The Twente humanoid head features a four degree of freedom neck and two eyes that are implemented by using cameras. The cameras tilt on a common axis, but can rotate sideways independently, thus implementing another three degrees of freedom. A vision processing algorithm has been developed that selects interesting targets in the camera images. The image coordinates of the selected target are provided to a motion control algorithm, which controls the head to look at the target. The degrees of freedom and redundancy of the system are controlled such that natural human-like motions are obtained. The head is capable of showing expressions through mouth and eyebrows by means of light projection from the inside part of the exterior shell.**

I. INTRODUCTION

Over the last years, the workspace of robots has shifted from human-free factory environments to areas where the robots are explicitly expected to safely interact with humans. Robots that can communicate with humans in a human-like way are expected to considerably promote the acceptance of robots in society, and thus, as a result, the research interest in humanoid robotic heads is growing. Existing humanoid heads found in literature may be assigned to one of two categories, which differ in the number of degrees of freedom (DOFs) and motion speed. Maveric [5] and ASIMO [1] rely on three and two DOFs respectively and fast motions so to track objects effectively. On the other hand, QRIO [6] and iCub [2] are designed to interact with humans and thus have more DOFs as to mimic human-like motions. However, as a result, the speed of motion is limited.

This work presents the Twente humanoid head, shown in Fig. 1. The purpose of the system is to provide a research platform for human-machine interaction and, as such, it should be able to track targets but also exhibit human-like motions (e.g. observing an object, nodding, expressing surprise or curiosity).

II. MECHANICAL DESIGN

Human motions are characterized by range of motion, velocity and acceleration. In order to have the humanoid head move human-like, the system should meet these human characteristics as close as possible. The mechanical design of the Twente humanoid head has been presented in [3]. A trade-off had to be made between having a small number of DOFs, enabling fast motions close to human capabilities, and having a number of DOFs that enable complex human-like motions. It was found that four DOFs in the neck and three DOFs in the vision system are sufficient to achieve the desired behaviors. The DOFs in the neck enable the head to tilt around two axes and make pan and roll motions. The cameras in the vision system tilt on a common axis, but can rotate sideways freely.

{l.c.visser,r.carloni,s.stramigioli}@utwente.nl, IMPACT Institute, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 7500 AE Enschede, The Netherlands.

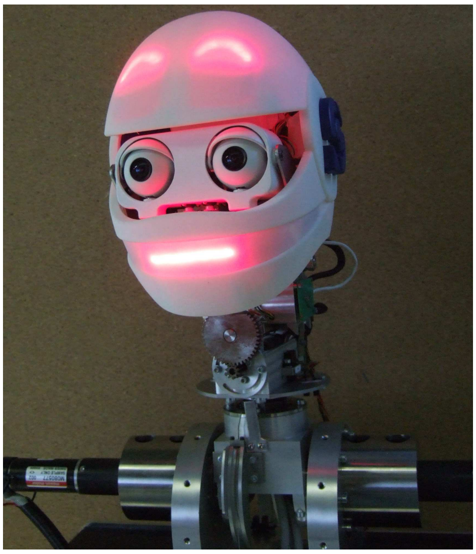

Fig. 1. The Twente humanoid head.

By combining one tilt motion and the pan motion of the neck in a differential drive, a compact realization is achieved. Because the differential drive itself is stationary, the total inertia of the system is kept low, enabling fast motions. To further improve performance, a gravity compensation mechanism has been designed for the lower tilt and the roll motion, so that motor torque and thus energy loss is reduced.

Table I provides a comparison between human biological data found in literature, and the performance characteristics of the humanoid head. In order to achieve the long range of tilt motion, the total range has been distributed equally over the two available axes.

III. VISION

The humanoid head perceives its environment by using two cameras. A biologically inspired vision processing algorithm has been designed that analyzes the camera images and selects the target of interest in the image plane. This section summarizes the work presented in [7].

The algorithm calculates a saliency map based on a given camera image. Saliency is a measure of interestingness of the different parts of the image and it can be calculated using different channels, e.g. color, intensity. The results from different channels are added so to realize a map in

TABLE I

BIOLOGICAL DATA OF A HUMAN HEAD AND MEASURED PERFORMANCE OF THE TWENTE HUMANOID HEAD

	Range (°)		Max. velocity (°/s)		Max. acceleration (°/s^2)	
	Human	Humanoid head	Human	Humanoid head	Human	Humanoid head
Tilt	-71 to $+103$	-35 to 41 (upper) ±36 (lower)	352	354	3300	3340
Roll	±63.5	±102	352	356	3300	3340
Pan	±100	±49	352	356	3300	3340

which the point with the highest value is selected as the most interesting point: the focus of attention (FOA). To prevent rapid switching between points with almost equal saliency, a positive bias is added to the current FOA. Additionally, over time a negative bias is added to the current and recent FOAs so that the whole field of view will receive focus at some time. The result is a natural target selection algorithm very similar to the way humans observe an unknown environment.

In addition, a Viola-Jones face detection algorithm is implemented [8], so that the head can look at people and recognize them. This is necessary for human-robot interaction and gives the humanoid head a very human-like behavior.

IV. CONTROL

The design of the control algorithm is described in detail in [9]. The vision control algorithm provides target coordinates x in the camera image space. To make the head look at the target, the goal is to control the head such that the target is in the center of the image. Assuming that the origin of the image coordinates are in the center of the image, a proportional control law in the image space is given by

$$\dot{x}_d = -K_p x \tag{1}$$

where $K_p > 0$ is a proportional gain and $\dot{x}_d$ is the desired velocity of the target in the image space. The target velocity is given by the forward kinematic relation of the head

$$\dot{x} = F(q)\dot{q} \tag{2}$$

where q and $\dot{q}$ describe the joint configuration and velocities respectively. Because of the redundancy in the system, the inverse of (2) is given by

$$\dot{q} = F^{\sharp}\dot{x} + (I - F^{\sharp}F)z \tag{3}$$

where $F^{\sharp}$ denotes the weighted generalized inverse of F, I is the identity matrix, and z is an arbitrary vector which is projected onto the null-space of F. The first right-hand term of (3) is a minimum norm solution, where the norm is defined by a positive definite matrix M_q:

$$\|\dot{q}\| = \sqrt{\dot{q}^T M_q \dot{q}} \tag{4}$$

The matrix M_q defines the ratio between joint velocities and is chosen in accordance with biological studies [4], which show that humans use their eyes to quickly move their gaze, while the head follows more slowly. By this choice, the humanoid head intrinsically responds human-like to changing targets. When tracking targets, the vector z in

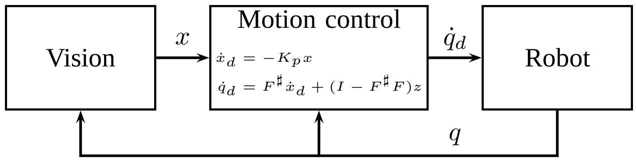

Fig. 2. Overview of the control scheme.

(3) is chosen to act as a proportional controller, so that the joints remain close to a desired neutral configuration q_0:

$$z = W(q_0 - q) \tag{5}$$

where W is a non-negative diagonal matrix. The vector z is also used to express emotions while the head is looking at a target by applying appropriate time varying functions to one or more of the joints. Examples are nodding in agreement, shaking in disagreement and moving the head forward and backward for curiosity and surprise. Fig. 2 shows an overview of the control scheme.

V. EXPRESSIONS

In designing the exterior shell, care was take to give the head a friendly appearance, and giving it human like features. To improve the expressions shown by the head, LEDs project eyebrows and a mouth from the inside of the exterior shell. The shapes of the eyebrows and mouth are changed to match the defined expression.

REFERENCES

[1] Y. Sagakami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki and K. Fujimura, "The intelligent ASIMO: system overview and integration", *IEEE Int. Conf. Intelligent Robots and Systems*, 2002.

[2] R. Beira, M. Lopes, M. Praca, J. Santos-Victor, A. Bernardino, G. Metta, F. Becchi and R. Saltaren, "Design of the robot-cub (iCub) head", *IEEE Int. Conf. Robotics and Automation*, 2006.

[3] D. M. Brouwer, J. Bennik, J. Leideman, H. M. J. R. Soemers and S. Stramigioli, "Mechatronic design of a fast and long range 4 degrees of freedom humanoid neck", *IEEE Int. Conf. Robotics and Automation*, 2009.

[4] H.H.L.M. Goossens and A.J. van Opstal, "Human eye-head coordination in two dimensions under different sensorimotor conditions", *Jour. Experimental Brain Research*, vol. 114, n. 3, pp 542-560, 1997.

[5] Maveric, http://www-humanoids.usc.edu/HH_summary.html.

[6] L. Geppert, "QRIO, the robot that could", *IEEE Spectrum*, vol. 41, n. 5, pp 34-37, 2004.

[7] R. Reilink, "Realtime Stereo Vision Processing for a Humanoid", *MSc-Report 019CE2008*, University of Twente, 2008.

[8] P. Viola and M. Jones, "Robust real-time face detection", *Int. Jour. Computer Vision*, 2001.

[9] L. C. Visser, R. Carloni and S. Stramigioli, "Vision Based Motion Control for a Humanoid Head", *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2009.