Head-eyes system and gaze analysis of the humanoid robot Romeo

N. Pateromichelakis¹, A. Mazel¹, M. A. Hache¹, T. Koumpogiannis¹, R. Gelin¹, B. Maisonnier¹, and A. Berthoz²

Abstract— In this work we present the head assembly and the gaze shifting capabilities of the bipedal humanoid robot Romeo. The purpose of the head system is to provide a reliable hardware platform for Human-Robot Interaction (HRI) applications, stereo vision based navigation or gazing experiments, either as stand-alone version or in synergy with the body movements. Towards this purpose, the number of joints, the angular range, speed and acceleration of the eyes and neck rotations should be in agreement with human psychophysics, i.e. reproducing realistic kinematics. We present a mechatronic system of the head and neck; accompanied by experimental results of head-fixed eye movements, solely head rotations, as well as eye-head cooperative motion for large angles of gaze shifting. We compare the values achieved by the prototype, with those of an average adult human.

I. INTRODUCTION

The Romeo project's [2] objective is to develop a humanoid robot that can assist persons suffering from loss of autonomy, such as elderly or partial disability. The robot will be in close cooperation with humans, roaming around their personal environment and support their everyday needs. Due to this close relationship, it is of high importance that the robot is able to acquire a realistic perception of the surrounding environment and to interact with individuals in the most natural way. Since facial movements, voice and gestures are the principal means of communication between humans and between a human and a robot, special attention has been given on these subjects during the development of the robot. Based on the experience of Aldebaran Robotics [1], on the interaction of the humanoid robot NAO with users [3] or even with individuals with special needs [4] using voice and whole body movements, we believe that a smooth, human-like movement of the head-eyes system will enhance the effectiveness of the HRI.

Several robots, robotic heads or other devices have been developed focusing on social interactions with humans, but only few of them combine a head capable to perform social interactions using a moving head and eyes in synergy with a humanoid body. The iCub robot performs low speed movements with 3 degrees of freedom (DOF) on the neck and 3DOF for both the eyes [5, 6]. The KOBIAN robot has 3DOF on the eyes, and 4 on the neck [7], similarly with the ARMAR robot developed by [8] where the work is targeted on gaze control [9]. Apart from the full robots mentioned above, half a robot accompanied with a head that the authors are aware of, is the BARTHOC system [10] with 3 DOF for the eyes and 4 DOF for the head. A distinguished robotic head is the Flobi head, developed by [11] that has 3DOF for the eyes and 3DOF for the neck, with speeds close to human kinematics. Finally, several heads have been developed for HRI applications copying the human appearance, such as the heads made by Hiroshi Ishiguro [24], or the SAYA head [25], incorporating 2 or 3 DOF on the eyes, actuated at usually low speeds.

By contrast, the Romeo head employs 4 DOF in the eyes with independent control that allows stereo vision at different targets. The human-like speed achieved by the eyes and the neck allows gaze stabilization during walking and enhances the emotional expression of the robot. In addition, the head alone could serve as an experimental evaluator of control schemes suggested by various research groups or projects (e.g. recently the European project ROBOSOM) that conducted considerable work on human and robot perception and action.

Section II gives a short overview of the human gaze characteristics that defined the specifications of our system. Section III and Section IV present an overview of the robotic head and the eye mechanisms, respectively. Experimental results of independent movements that have been achieved by the eyes, the neck, as well as combined gaze are analyzed in Section V, and conclusions are drawn in Section VI.

Figure 1. A picture (left) and a 3D model (right) of the head and neck assembly of the humanoid robot Romeo.
II. HUMAN HEAD AND EYES MOVEMENTS

In this section we briefly summarize the physiology and the capabilities of the human gaze system (the sum of the head angle with respect to torso and eye angle with respect to the head), which constitutes the objective for the design of the robotic counterpart. In humans, reorientation of gaze is essential for visual exploration, scene perception, pursuit of a moving object, or anticipation and guidance of locomotor trajectories. Additionally, eye-head coordinated movements allow us to express basic emotions like “yes-no” behaviors or to achieve eye contact. Human joints, bodies, weights and range of motion are compared with the values achieved by the developed system (Table I).

<table>
<thead>
<tr>
<th>TABLE I. COMPARISON BETWEEN HUMAN AND ROBOTIC CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Diameter</td>
</tr>
<tr>
<td>Moving mass</td>
</tr>
<tr>
<td>FOV including movement, right eye [Right, Left, Up, Down]</td>
</tr>
<tr>
<td>Maximum movements of the eye [Right, Left, Up, Down]</td>
</tr>
<tr>
<td>Maximum non-controllable speed</td>
</tr>
<tr>
<td>Maximum controllable speed</td>
</tr>
<tr>
<td>Time to initiate</td>
</tr>
<tr>
<td>Weight</td>
</tr>
<tr>
<td>Max controlled speed for YAW</td>
</tr>
<tr>
<td>Range of motion [Right, Left, Up, Down, Roll]</td>
</tr>
<tr>
<td>Time to initiate</td>
</tr>
<tr>
<td>Binocular vision [Horizontal, vertical]</td>
</tr>
<tr>
<td>Total field of view [Horizontal, vertical]</td>
</tr>
</tbody>
</table>

The head of humans and some other primates can rotate around three axes and translate in three directions. The range of movements, speed, field of view (FOV) and other interesting characteristics are shown on table I.

The movement of the human eye is controlled by three pairs of muscles. It can reach angular speed of about 400°/sec and peak angular speed that depends on the type of movement: a) Stabilizing reflexes (vestibulo-ocular reflex, VOR, or optokinetic) which allows minimization of image slip on the retina during gaze shifting (up to at least 100°/sec) b) saccades are rapid eye movements which can be either reflex to an appearing target (a mechanism induced by object motion detection at the periphery of the visual field), or voluntary for visual search which is performed by sequence of saccades with a tight amplitude-duration relationship (angular eye velocity up to 500°/sec); in human and non-human primates, large gaze shifts are generally produced by a combination of saccades and rapid head movements c) smooth pursuit is slower. Its purpose is to maintain the image of a moving target on the fovea (0 to 60°/sec) [12, 13]. The sluggishness is compensated by the ability to anticipate the motion of the target in the case of repetitive motions d) vergence. These movements are convergent and allow the perception of objects that become close to the head in near action space.

Head motions are constrained in humans by the skeletal architecture of the neck, which provides separate control for yaw, pitch and roll movements. Each human eye is mechanically independent from the other, and it has been shown that the eyes and the head movements are controlled by independent closed loop driving circuits. However, these systems interact during the execution of the movement [14, 15]. It has been also shown that the gaze shift obeys the common physical principle of the “minimum effort rule” [16], and that all human eye rotations have their axis on a single frontal plane, which simplifies the control of 3D eye movements (Listings law). These principles may be the basis for an advanced control of the robotic gaze shift in the future.

III. THE ROBOTIC HEAD

In this section we briefly present the assembly of the head and the neck of the humanoid robot Romeo, focusing on the important for the application areas. Basic mechanical and electronic specifications will be given and joints, bodies, weights and range of movements will be presented.

Figure 2. Left: Overview of the joint axis and their relative position. Right: Representation of the joints.

The head measures about 200mm in length (from the nose to the back), 160mm in width (ear to ear) and 260mm in height. The neck is a cylinder approximately 80mm in diameter and 230mm in height. The head weighs 1.9Kg and
the neck assembly about 0.7Kg. The design of the outside shell of the face keeps the general characteristics of the human head, but respects the uncanny valley law [19], in order to make it attractive to users to interact with.

The robotic head is able to rotate around three axes, similarly to the human head. We consider the bottom of the neck as fixed to the body of the robot or to a base. Two of the DOF move the neck with respect to the fixed body in pitch (around Y axis) and yaw (around Z axis) and are named as “NeckPitch” and “NeckYaw”, respectively. Two more joints move the head in respect to the neck in pitch and roll (around X axis), and are named as “HeadPitch” and “HeadRoll” accordingly. The distance between the intersections of the “Head” and “Neck” axis is 95mm. Fig 2 shows the joint axis at their zero positions as well as their relative positions.

Each axis is actuated by a DC motor (Maxon, RE-max24, 11Watt) with gear ratios ranging from 130 to 201. The torque output of the motors is 36mNm for NeckYaw, 42mNm for the NeckPitch, 7mNm for the HeadPitch and 10.5mNm for the HeadRoll. Every axis is controlled by an individual PID controller on dedicated motor control boards, and incorporates two contactless rotary encoders (Austriamicrosystems AG, AS5045, Austria) with a resolution of 0.0879°. Both axis feedback and motor feedback are used to increase the reliability and the response of the feedback scheme.

The head incorporates 3 CPU boards based on the Intel Atom (Z500 series) processor, assigned as Audio, Video and AI, and are used for image processing, sound processing and high level reasoning, accordingly. High speed communication between the several processors is used, and special attention has been given to their synchronization. Head inclination is measured by an Inertial Measurement Unit (IMU) installed in the head. Finally, a wide selection of sensors and indicators is used, in order to enhance the HMI capabilities of the robot such as LEDs (mouth and tactile sensor), microphones, cameras and speakers. An optional depth sensor (ASUS Xtion) increases the performance of the robot navigation and perception.

IV. THE EYE MECHANISM

In this section we present the mechanism for eye control of the humanoid robot Romeo. First, we show the adaptation of an existing mechanism into the current assembly. Then, the electrical design is presented, followed by an overview of the control strategy that has been followed.

A. Mechanical design

The mechanism of the eyes of the humanoid robot ROMEO is designed to satisfy the specifications of the human eye in speed, acceleration and field of view. Each eye is a spherical mechanism with two actuated DOF (Pitch and Yaw) and one passive (Roll), based on the simplified 2-DOF version of the Agile Eye that has been developed by [17]. The two eye mechanisms are identical and mirrored with respect to the vertical plane in the middle of the head. Dedicated control for each eye with high precision feedback is followed, in order to magnify the research capabilities of the platform, thus, improving the accuracy of the stereo vision and gaze shifting.

![Figure 3. Back, Left and isometric views of the eye mechanism.](image)

Similarly to the rest of the robot, extensive use of 3D printing and metal sintering technologies allowed us to achieve complex structures with low weight and inertia, which is especially important for moving assemblies. The actuating motors has been attentively arranged to minimize the volume of the complete mechanism, by maintaining the right angle (90°) per pair of motors which is essential for the mechanism and providing an angle of 20.75° between the motors on the side plane. This angle reduces the workspace of the mechanism; however, this extra workspace is not used due to the limited rotation angle of the eyes. The two axis of rotation and the axis of the line of sight intersect at the center of the mechanism (Fig.4, middle and right), whereas the center of mass of the eye parts 5,6,7 is placed as close to that as possible.

Fig.4 shows the components that constitute the mechanism. A brushed DC motor (Maxon DCX10L) with a 4:1 reduction gearhead is used on every input axis. A contactless magnetic rotary encoder is used as feedback sensor (Austriamicrosystems AG, AS5045, Austria) with a resolution of 0.0879°. The proximal parts and the two parts that form the rotational joint in the middle of the mechanism are made by high-strength aluminum using laser sintering, and are machined afterwards to achieve the required tolerances. Ceramic ball bearings have been used to decrease
the inertia of the mechanism. The part between the camera board and the rotational joint is printed in a very light and durable plastic (PA-UD) and finally the front eye shell is molded using light polypropylene (PUX-220), to achieve the desired visual result from the outside. The total weight of the moving parts (parts 4 to 7 in Fig.4) of each eye is 11.2 grams, with very low inertia, due to the spherical mechanism.

B. Electrical design

From the electrical point of view, the system consists of the face-board, one pair of moving cameras in the eyes, the processor responsible for the video acquisition and processing and finally the dedicated IMU.

The central unit is the face-board is about 100x48x16mm in size and it is placed above the eyes mechanism. It receives input from the 2 cameras in the eyes (Aptina Imaging, MT9M114, resolution 1.3Mp, 1280x960 active pixels at 30fps) and it transmits the data to the processing unit that is responsible for video analysis through a PCIe bus. The face-board controls the eyes mechanism by governing the 4 DC motors with 4 dedicated PID control loops.

The IMU is a board 32x23x3mm (LxWxH) in size made by Aldebaran Robotics, that incorporates a 3-axis digital accelerometer with 16-bit data output (ST Microelectronics, LIS331DLH) and a 3-axis digital gyroscope (ST Microelectronics, L3G4200D). The board is connected to the front-face board through I2C protocol that provides a high-speed response of the eye mechanism to inclination changes of the head. This is essential for applications such as gaze stabilization or reduction of blur on the frames, during head movements or walking.

C. Control design

The forward and inverse kinematics of the mechanism will be not analyzed in this work, since similar mechanisms have been well documented by other works [17, 18]. Fig. 6 shows the logical diagram of the control loop that is implemented. The user or the high level software sends the desired set points for the eyes to the Hardware Abstraction Layer (HAL). The HAL is a software package installed in every processor inside the robot, and it is responsible for the communication between software and hardware modules. The HAL converts the desired eye positions to motor positions, according to the given inverse kinematics. This information goes to the controller (Microchip, dsPIC33EP) on the face-board, which implements an individual PID controller for every motor. A dual full bridge PWM motor driver (Allegro A3995) is used to drive the motors. The time loop of the processor is 3ms and of the face-board is 1ms, allowing quick control of the movements of the eyes. The refresh time of the data to and from the user is 10ms.

Figure 5. The mechatronic system of the eyes consists mainly of the eyes mechanism, the faceboard and the IMU.

Figure 6. An overview of the control loop that is used in the eyes. The setpoint of the eye movement which is given by the user is converted into the system’s coordinations within the HAL, and forwarded to the faceboard which is responsible for the closed loop control of the eyes.

V. EXPERIMENTAL EVALUATION

This section presents the experimental results acquired by the system. We present results as the eyes motor and the neck yaw response, to quick and slow activations. Then, co-actuated motion of the eyes and the head is presented, followed by a validation of these results in respect to the values achieved by the average adult human [12, 14]. Although gaze shifting can be directed towards any direction within the mechanical limits of the mechanism, during this study we will only focus on the yaw movement of the neck and the eyes, primarily because this movement along the horizontal plane is extensively used by humans. In addition, it has been experimentally observed that neck and eyes movements of the robot in other directions can achieve similar speeds and accelerations due to their similar mechanical characteristics.

A. Head-fixed saccadic movements

To provide a basic evaluation of the response of the motors that control the eyes, a command for a high-speed rotation of about 450°/sec, and a low speed rotation of about 131°/sec, has been sent to a motor (Fig.7).
The hysteresis between the setpoint and the recorded value is about 10ms for both speeds, since this is constrained by the refreshing time on the high level software. It has been experimentally observed and theoretically supported, that the feedback rate limits the speed where the motor can be adequately controlled to 500°/sec. This comes in agreement with the corresponding values of human eyes, where voluntary eye movements have not been observed to exceed 400°/sec; any value above that shows ballistic motor behaviour. Additionally, a passive rotation about 2° around the line of sight (roll axis) of the mechanism has been observed, similarly to human eye torsion (2° to 4°).

### B. Head only movements

Further investigation includes the response of the yaw joint of the neck, which is rotated at 45° around the z-axis (yaw), at low and high speed (67°/sec and 200°/sec respectively) with a trajectory configured by the user prior to the movement. The maximum speed that the NeckYaw can perform in a controlled way is 250°/sec. The hysteresis of about 30ms that is observed in Fig. 8 is due to the time delay between the motor board and the processor that controls all the joints of the robot (only the eyes are controlled by dedicated board with high communication).

### C. Head and eyes cooperative motion

In order to compare the gaze shift of the robot with that of humans, an experiment that simulates the human gaze has been performed. The gaze that is presented here is using the characteristics of human gaze shifting as described in [14], and does not uses any feedback from the cameras to specify the trajectory, it solely serves the purpose of showing the correlation to the human gazing.

The head and the eyes are initially at zero position. At an arbitrary time of 60ms a command to move the neck towards a point at 45° and the eyes at 13° on the horizontal is given. The eyes respond to the command after 10ms and the neck after 50ms. Since the eyes can move much faster they achieve their maximum allowed angle before the neck starts to rotate. While the neck rotation is progressing, the eyes start to counter-rotate, in order to compensate for the ongoing movement. This, in humans, is provided by the vestibulo-ocular reflex but in this experiment is pre-programmed so that eyes and neck are reaching their final positions after about 700ms. Fig. 9 shows the joint measurements from the eyes yaw and the neck yaw. Finally, Fig. 10 shows the frames from the key time points during the corresponding experiment illustrated in Fig. 9.

---

**Figure 7.** Example of an eye-motor response on a high speed (up) and a low speed (bottom) trajectory.

**Figure 8.** Example of the setpoints and the joint response acquired during neck yaw rotation for 45°, for low speed and high speed (67°/sec and 200°/sec respectively).

**Figure 9.** Gaze measurements from the platform, with characteristics inspired by human gaze.
VI. CONCLUSION

The humanoid robot Romeo developed by Aldebaran robotics is designed to endeavor the living quality of elderly or impaired persons. In this perspective, gazing is one of the most significant abilities that robotic human assistant has to fulfill. In this paper, the design of the neck, head and eyes mechanism is presented. Experiments on solely eyes or neck movements and finally combined gazing have been performed, and the results have been presented and correlated to human values. Eye rotations with speeds up to 400°/sec and neck of up to 250°/sec have been achieved in a controlled manner. Nonetheless, higher speeds could be realized in a non-controlled way, similarly to the saccadic movement of the human eyes or the neck reflexes.

Future experiments using an external video recorder with high frame-rate and definition could lead to more solid results. Improvements on the mechanics, electronics and control will enhance the reliability of the setup during long endurance tests or experiments. Finally, the motion of head together with gestures and body movements of the whole robot and the evaluation by experts will emerge the importance of this work.

ACKNOWLEDGMENT

This work has been supported by BPI France in the framework of the ROMEO2 Project [2]. The authors would like to acknowledge Y. Dupraz and P. Leboucher from College de France for suggesting the eyes mechanism, and the ROMEO team for their support and valuable remarks. Sample of the gaze experiments can be seen in the video accompanying the paper.

REFERENCES