INTERSPEECH 2017: Show & Tell Contribution
August 20–24, 2017, Stockholm, Sweden

Creating a Voice for *MiRo*, the World's First Commercial Biomimetic Robot

Roger K. Moore[1], *Ben Mitchinson*[2]

[1]Speech & Hearing Research Group, Dept. Computer Science, University of Sheffield, UK
[2]Dept. Psychology, University of Sheffield, UK

r.k.moore@sheffield.ac.uk, b.mitchinson@sheffield.ac.uk

Abstract

This paper introduces *MiRo* - the world's first commercial biomimetic robot - and describes how its vocal system was designed using a real-time parametric general-purpose mammalian vocal synthesiser tailored to the specific physical characteristics of the robot. *MiRo*'s capabilities will be demonstrated live during the hands-on interactive 'Show & Tell' session at INTERSPEECH-2017.

Index Terms: biomimetic robot, MiRo, mammalian vocalisation, vocal synthesis

1. Introduction

Recent years have seen growing interest in the use of animal-like robots for a range of applications such as assistance, companionship, pet therapy and edutainment. The majority of such devices are engineered to support specific use-cases, and often it is only the robot's physical appearance that mimics a particular animal. In contrast, *MiRo* (designed and built by Consequential Robotics Ltd. in collaboration with the University of Sheffield [1]) is the first commercial robot to be controlled by hardware/software modelled on the biological brain [2, 3].

MiRo is a highly featured, low-cost, programmable mobile developer platform, with a friendly animal-like appearance, six senses, a nodding and rotating head, moveable hearing-ears, large blinking seeing-eyes, and a responsive wagging tail. It has been designed to look like a cartoon hybrid of a generic mammal - see Fig. 1. *MiRo* is suited for developing 'companion' robots - future social robots that will share our personal space, interact with us and with each other and provide emotional engagement and entertainment. A unique brain-based biomimetic control system allows *MiRo* to behave in a life-like way: for example, listening for sounds and looking for movement, then approaching and responding to physical and verbal interactions. *MiRo*'s control architecture is outlined in Section 2.

Of particular interest here is *MiRo*'s ability to vocalise. This was achieved using a real-time parametric general-purpose mammalian vocal synthesiser [4] tailored to the physical and behavioural characteristics of the robot [5]. The design of *MiRo*'s voice is summarised in Section 3.

2. MiRo's Control Architecture

MiRo's control architecture operates across three embedded ARM processors that mimic aspects of spinal cord, brainstem and forebrain functionality - see Fig. 2. One important feature is that the control latency of loops through the lowest reprogrammable processor can be as low as a few milliseconds. If required, *MiRo* can be operated remotely through WiFi or Bluetooth, and can also be configured as a Robot Operating System (ROS) node.

Figure 1: *The MiRo biomimetic robot.*

2.1. Actuators

MiRo is constructed around a differential drive base and a neck with three 'degrees of freedom' (DoF). Additional DoFs include rotation for each ear, tail droop and wag, and eyelid open/close. All DoFs are equipped with proprioceptive sensors, and the platform also has an on-board loudspeaker.

2.2. Sensors

MiRo is equipped with stereo cameras in the eyes, stereo microphones in the ears and a sonar range-finder in the nose. Four light-level sensors are placed at each corner of the base, and two infrared 'cliff' sensors point down from its front. Eight capacitive sensors are arrayed along the inside of the body shell, over the top of the head and behind the ears. Internal sensors include twin accelerometers, a temperature sensor and battery-level monitoring.

2.3. Affect

MiRo represents its affective state (emotion, mood and temperament) as a point in a two-dimensional space covering 'valence' (unpleasantness-pleasantness) and 'arousal' (calm-excited) [6]. Events arising in *MiRo*'s sensorium are mapped into changes in affective state: e.g., stroking *MiRo* drives valence in a positive direction, whilst striking *MiRo* on the head drives valence in a negative direction. Baseline arousal is affected by general sound/light levels and the time of day. An individual event can cause an acute change: e.g., a very loud sound might raise arousal and decrease valence. *MiRo*'s movements are modulated by its affective state, and it expresses itself using a set of 'social pattern generators' that drive light displays, movement of the ears, tail, eyelids and - of particular importance here - *vocalisation*.

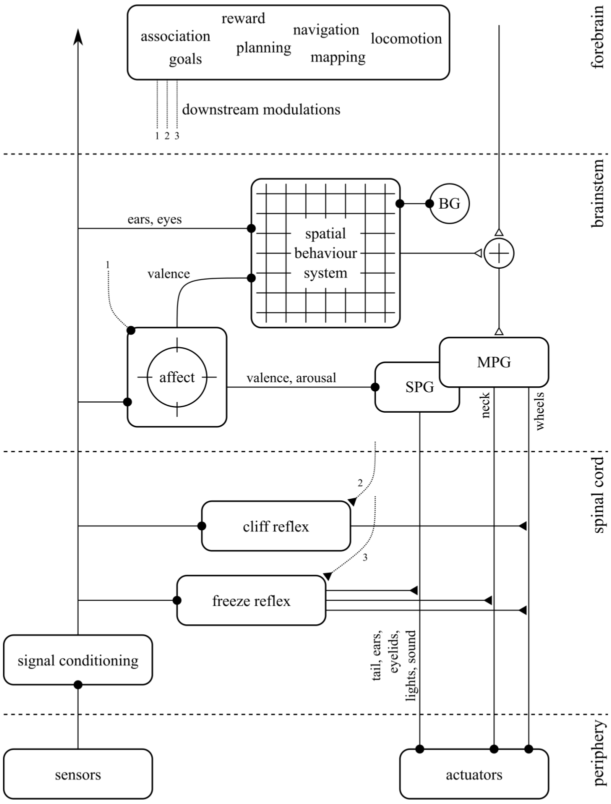

Figure 2: *Illustration of MiRo's control architecture loosely mapped onto brain regions (spinal cord, brainstem, forebrain). Signal pathways are excitatory (open triangles), inhibitory (closed triangles), or complex (closed circles). BG is the Basal Ganglia. SPG/MPG are the Social/Motor Pattern Generators.*

3. MiRo's Voice

A special requirement of *MiRo*'s design was that its voice should be 'appropriate' to the biomimetic nature of the robot. Hence, rather than using a fixed inventory of pre-recorded animal sounds, *MiRo*'s ability to vocalise was developed using a general-purpose mammalian vocal synthesiser [4] tailored to the specific physical characteristics of the robot [5]. The approach taken was to decompose the overall sound production system into the relevant anatomical components (body, lungs, vocal cords, tongue and mouth), and to model each component within a real-time parametrically-controlled sound generator. The net result is that every vocalisation is both appropriate and unique.

3.1. Synthesis Software

The overall structure of the synthesis software is based on a simulation of the flow of energy through a generic mammalian vocal apparatus with body mass corresponding to an animal of *MiRo*'s size. The key modules correspond to the lungs, larynx and vocal tract. The robot has a breathing rhythm, the frequency of which is linked to arousal (see Section 2.3), and vocalisation is initiated stochastically during the exhalation phase. Breathing is simulated as cyclic airflow into and out of the lungs with an amplitude and duration that is calculated from the flow rate, lung capacity and body mass. When vocalising, the larynx modulates the airflow using the simulated action of a set of vocal folds vibrating at a fundamental frequency that is also determined by the body mass. The vocal tract then simulates three acoustic resonances (formants) using band-pass filters whose frequencies are determined by the vocal tract length and the degree of mouth opening. A syllabic rate parameter controls the opening and closing of the mouth, and a vibrating uvula adds a 'cute' robotic timbre to the voice.

3.2. Vocal Emotion

In order to allow the injection of emotion into the vocalisations, key parameters were linked to *MiRo*'s two-dimensional affect map (as discussed in Section 2.3). Arousal modulates the airflow rate and, thereby, the amplitude and tempo of the vocalisations; high arousal leads to high airflow and short vocalisations (and *vice versa*). Valence influences the variance of the fundamental frequency and the voice quality; high valence leads to expressive vocalisation whereas low valence produces more monotonic utterances. For example, stroking *MiRo*'s head increases valence, which leads to 'happier' vocalisations (and a wagging tail).

4. Conclusion

This paper has introduced *MiRo* - the world's first commercial biomimetic robot - and described how its vocal system was designed using a real-time parametric general-purpose mammalian vocal synthesiser tailored to the specific physical characteristics of the robot. *MiRo*'s capabilities will be demonstrated live during the hands-on interactive 'Show & Tell' session at INTERSPEECH-2017.

5. Acknowledgements

This work was partially supported by the European Commission [grant numbers EU-FP6-507422, EU-FP6-034434, EU-FP7-231868 and EU-FP7-611971], and the UK Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/I013512/1].

6. References

[1] MiRo: The Biomimetic Robot. [Online]. Available: http://consequentialrobotics.com/miro/

[2] B. Mitchinson and T. J. Prescott, "MiRo: A robot mammal with a biomimetic brain-based control system," in *Biomimetic and Biohybrid Systems. Living Machines 2016. Lecture Notes in Computer Science*, N. Lepora, A. Mura, M. Mangan, P. Verschure, M. Desmulliez, and T. Prescott, Eds. Springer, 2016, vol. 9793, pp. 179–191.

[3] E. C. Collins, T. J. Prescott, B. Mitchinson, and S. Conran, "MiRo: a versatile biomimetic edutainment robot," in *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology - ACE '15*. Iskandar, Malaysia: ACM Press, 2015, pp. 1–4.

[4] R. K. Moore, "A real-time parametric general-purpose mammalian vocal synthesiser," in *INTERSPEECH*, San Francisco, CA, 2016, pp. 2636–2640.

[5] R. K. Moore and B. Mitchinson, "A biomimetic vocalisation system for MiRo," in *Living Machines 2017*, Stanford, CA, 2017. [Online]. Available: http://arxiv.org/abs/1705.05472

[6] E. C. Collins, T. J. Prescott, and B. Mitchinson, "Saying it with light: a pilot study of affective communication using the MiRo robot," in *Proceedings of the 4th International Conference on Biomimetic and Biohybrid Systems - Volume 9222*. Barcelona, Spain: Springer-Verlag New York, Inc., 2015, pp. 243–255.