

Humanoid Robots That Behave, Speak, and Think Like Humans: A Reduction to Practice RRC-Humanoid Robot

Alan Rosen, David B. Rosen

Machine Consciousness Inc., Redondo Beach, CA, USA

Email: arozu1@gmail.com

Received 12 December 2014; accepted 9 January 2015; published 16 January 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A radical new approach is presented to programming human-like levels of Artificial Intelligence (AI) into a humanoid robot equipped with a verbal-phoneme sound generator. The system shares 3 important characteristics with human-like input data and processing: 1) The raw data and preliminary processing of the raw data are human-like. 2) All the data are subjective, that is related and correlated with a robotic self-identity coordinate frame. 3) All the data are programmed behaviorally into the system. A multi-tasking Relational Robotic Controller (RRC)-Humanoid Robot, described and published in the peer-reviewed literature, has been specifically designed to fulfill those 3 characteristics. A RRC-controlled system may be behaviorally programmed to achieve human-like high I.Q. levels of subjective AI for the visual signals and the declarative-verbal words and sentences heard by the robot. A proof of concept RRC-Humanoid Robot is under development and present status is presented at the end of the paper.

Keywords

Human-Like Artificial Intelligence, Humanoid Robots, Thinking Machines, Behavioral-Programming, Experiential Programming, Behavioral Speech Processing

1. Introduction

The design of “thinking computers” has been a goal of the discipline of Artificial Intelligence (AI) since the advent of digital computers. In 1950, Alan Turing, arguably, the founder of AI, posed the question “when is a machine thinking?” His approach to an answer was in terms of the behavior of the machine [1] [2]. He devised an I.Q. “Turing test” based on the conversational behavior of the machine; and deemed any machine that passed the I.Q.-test to be a thinking machine [2] [3].

We, in this paper, follow Alan Turing and describe a building path for a machine that can reach human-like, high I.Q. levels of AI, defined in terms of the behavior of the machine. But instead of programming the computer (robotic controller) with AI, we first program a “robotic self” into the system, that identifies the robotic system, and then program, experientially, all the AI that the robot gains with respect to, or into the robotic self coordinate frame of the system. So that it is the robotic self that develops a high IQ-level of intelligence, NOT the objective-mechanical digital computer system.

We have thereby designed a system, called a Relational Robotic Controller (RRC)-system that has a subjective identity and AI-knowledge associated with that identity. It is the “robotic self”, programmed into the computer that “thinks”, not the objective-mechanical digital computer.

An overview of a behavioral programming approach to the design-development of humanoid robots has been described in a previous paper [4]. The auditory RRC-Humanoid Robot is a human-like robotic system, controlled by a proprietary Relational Robotic Controller (RRC) [5]-[7].

1.1. But First, a Note about Human-Like Levels of AI

All programmable digital computers do not have a “self identity” as a human does, which could absorb and convert all data into subjective knowledge, knowledge absorbed relative to the “self” of the machine. Therefore, the computers do not have human-like intelligence. Computers have machine-like intelligence, *not* human-like intelligence.

Machine-like intelligence may refer to the objective knowledge programmed into all modern day computing devices. Human-like intelligence is obtained relative to the “self” of the machine. Human-like intelligence is called subjective knowledge. The following are six pre-requisites required to achieve human-like levels of AI:

1) The human brain relates, correlates, prioritizes and remembers sensory input data. Similarly, to achieve human-like intelligence, relating, correlating, prioritizing and remembering input patterns must be the essential analysis tool of the robotic controller. The RRC, a proprietary robotic controller of MCon Inc., was specifically designed to emulate the operation of the human brain. It also was designed to operate with a ‘self’ circuit that is the central hub of intelligence for the whole robotic system.

2) Humans have a self-location and identification coordinate frame that is trained from infancy to give the human brain a proprioceptive self-knowledge capability. Even a baby, with a self-knowledge capability, instinctively knows the location of every surface point on its body, the location of its flailing limbs, and by extension, the location of every coordinate frame point in the near space defined by its flailing limbs. The fundamental design characteristic of the RRC-Humanoid Robot is a centralized hub of intelligence, a proprietary module that is the centralized “self location and identification” coordinate frame of the system. This module gives the RRC-Humanoid Robot a robotic form of proprioceptive knowledge, similar to human proprioceptive intelligence. In the RRC-Robot, the self-knowledge capability is the basis for all knowledge.

3) Human intelligence is experiential intelligence. Humans learn from, and remember their experiences throughout their lifetime. A behaviorally programmed RRC-Humanoid Robot emulates the experiential intelligence of a human.

4) In order to achieve contextual, or “self-knowledge” of visual data, auditory data, olfactory data, gustatory data, and vestibular data, all the data obtained from those human-like sensors must be related and correlated with the self-knowledge, self location and identification coordinate frame.

5) Human intelligence is gained only from the human-like sensors. In this paper we consider the external sensors: Tactile, visual, auditory, olfactory, gustatory, and vestibular sensors. These sensors provide for the sensations associated with human “feeling”, “seeing”, “hearing”, “smelling”, “tasting”, and “balancing”. The recording monitors of the RRC-Humanoid Robot emulate the external sensors of humans.

6) Human-like intelligence may be gained by a human-like RRC-Humanoid Robot. The mechanical robotic body and associated sensors simulate the human body and the human sensors. The robotic body must be bipedal, standing and walking upright with two arms hands and five fingers per hand free to manipulate objects in the environment. The 6 robotic sensors should be human-like sensors designed to gain the same information as is gained by the human sensors.

2. Behavioral Programming and the Development of Human-Like AI

Behavioral programming is achieved by training the humanoid robot to control its body, limbs, and verbal-pho-

neme sound generator on the basis of input data from the six external sensors. It is an experiential supervised programming technique analogous to the autonomous learning of a human. Behavioral programming techniques are employed for all the sensory input signals of the humanoid robot. The most important aspects of behaviorally programming the auditory sensors and the verbal phoneme sound generator is described in the following sections.

2.1. Programming a “Self Knowledge” Coordinate Frame

The tactile input signals are used to define the central hub of intelligence, the self-nodal map/coordinate frame, of the humanoid robot. The behavioral programming technique employed for the self location and identification “self-knowledge” coordinate frame is an itch-scratch methodology, wherein the robot is fully trained and remembers how to 1) reach and touch (scratch) all points located on the surface of the robotic body, and all points in the near space surrounding the robotic body; 2) to identify and locate all such points; and 3) to identify and locate all the “end joint” body parts (ends of fingers, elbow, knees, etc.) used to scratch all the itch points. When the level of training reaches the threshold of “self-knowledge”, the Self Nodal Map Module and associated Task Selector Module (TSM)-memory module will facilitate the robotic identification and recognition of all body parts, and the navigation of all movable parts of the robot towards any and every itch point located on the surface of the robotic body, and all points in the near space surrounding the robotic body.

The totality of the programmed “self location and identification” data, stored in a TSM-memory module, is the basis for the “self-knowledge” level of intelligence. The “self-knowledge” level of intelligence is analogous to the proprioceptive knowledge of a human. A RRC-Robot with a fully programmed “self-knowledge” capability “knows”, behaviorally, the location of every surface point of the robotic body, the location of flailing limbs, and by extension, the location of every coordinate frame point in the near space defined by flailing limbs.

2.2. Developing Self-Knowledge for the Visual and Auditory Sensors

In the visual and auditory RRC-Humanoid systems, experiential intelligence is obtained by performing behavioral programming on the processed raw data coming from the video-visual recording monitor and the auditory recording monitor. The raw data are processed in an Interface Circuit, inputted to the RRC and then behaviorally programmed to reach human-like levels of AI. Behavioral programming reaches the level of experiential human-like intelligence, when the RRC-Humanoid Robot demonstrates behaviorally that it has identified, recognized, visualized and comprehended in the same manner as does a human, the signals coming from the visual sensors, or the auditory sensors. The processing of the video-visual raw data was described in previous publications [4] [7]. The following sections will describe the processing of the auditory raw data in the Interface Circuit, and the behavioral programming of the processed data within the RRC-Humanoid Robot. On completion of behavioral programming, the RRC-Humanoid Robot demonstrates behaviorally human-like levels of AI for the identification, recognition, visualization or comprehension of the processed raw data. Note: Behavioral-programming of the Auditory RRC-Humanoid Robot generates an operational definition of the “identification”, “recognition” and “comprehension” levels of AI. Any human need merely verbally ask the RRC-Robot to identify, recognize, comprehend, or visualize any color or 3D-object in the FOV, in order to obtain a verbal response that is indistinguishable from the response of another human.

3. The Operation of the Auditory RRC-Humanoid Robot

3.1. Processing the Auditory Raw Data in the Interface Circuit

The auditory raw data consist of the output of the sound receiving microphones that are sensitive to the auditory frequency range of zero to 20,000 cps (simulating the human ear). When a talking sound is applied to the ear-like receiving microphones, they generate a sequence of frequencies and amplitudes as a function of time. These frequencies and amplitudes may be illustrated in an amplitude-frequency-time(a-f-t)-diagram shown at the top of **Figure 1**. These (a-f-t) data are inputted to the Interface Circuit.

The Interface Circuit stage performs the following functions: 1) It processes the a-f-t data into collective modalities that have been selected to be characteristic of 120 phoneme-sound combinations present in the English language. 2) With the aid of a spectrum analyzer, the collective modalities are tuned to the selected phoneme sound combinations in the English language, to musical sounds or to environmental noise. 3) It performs beha-

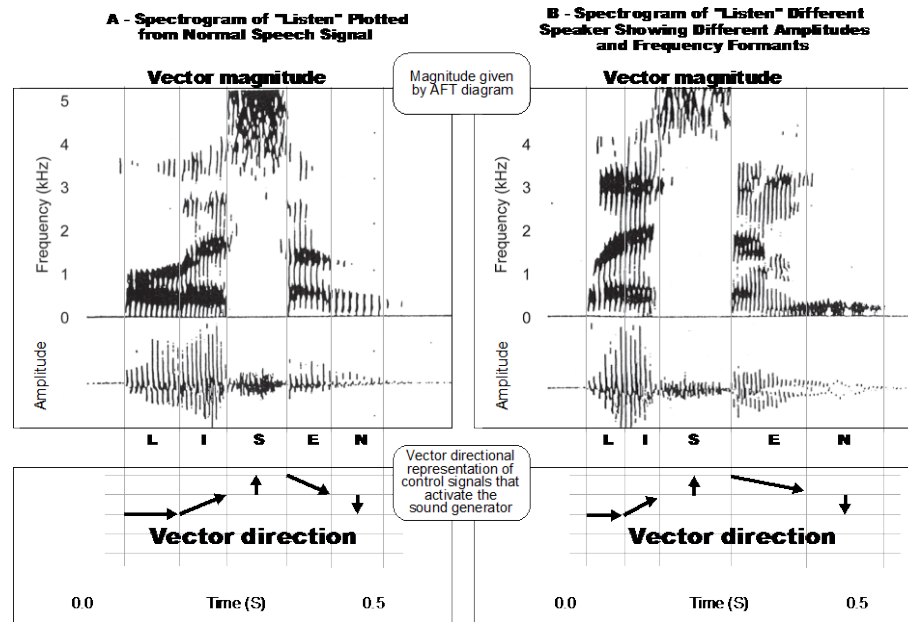


Figure 1. The magnitudes and directions of a sequence of multi-dimensional p-phoneme vectors representing the word “listen”. The vector direction is shown at the bottom, whereas the functional vector magnitude, the a-f-t-data, is shown at the top. Sections A and B show the differing amplitude and frequency formants for different speakers.

vioral speech processing on sequential sets of phoneme sounds and identifies and recognizes these sequential sets as “words” or “sentences” formed by the person speaking to the microphones of the auditory recording monitor. (Note: behavioral speech processing starts in the Interface Circuit stage by training the Robot to “repeat the heard words.”) The speech processing described in Section 4 includes a capability to recognize, identify and correct, incorrect grammatical structures. 4) Finally, the (a-f-t) data must be formatted within the Interface Circuit so that it is compatible with the RRC-Phoneme Sound Generator and the input to the RRC-Multi-dimensional Nodal Map Module [8].

The final output of the Interface Circuit stage, the “words and sentences”, music or environmental noise, are inputted to their respective multi-dimensional Nodal Map Modules of the RRC-Humanoid Robot, for behavioral programming.

The operation of the RRC-circuits is described in Sections 3.2. and 3.3. The operation of the memory system of the RRC robot is described in Section 3.4. And Section 5 describes the behavioral programming of an RRC-robot that “knows” and understands what it “hears”, and demonstrates that knowledge verbally.

3.2. The Operation of the RRC-Circuits That Perform Identification, Recognition, and Comprehension of Verbal-Phoneme Words and Sentences

3.2.1. Introduction to the RRC-Robot

The RRC-controller is made up of the array of Relational Correlation Sequencer (RCS) [5]-[7]. Control, computation and data storage are performed by relating and correlating each and every input data signal (tactile, visual, or auditory signals) with each other; and relating and correlating each and every input data signal with each and every output-control signal generated by the RRC-controller.

When the input sensory signals are related to the self-location and identification Nodal Map Module, the sensory signals are said to be “internalized” to the robotic “self” of the system. When the data from the other sensors is internalized the system gains a contextual form of self-location and identification knowledge for those other sensors. Multiple relations and correlations are required in order to achieve higher levels of AI-identification, AI-recognition, AI-visualization and AI-comprehension of all input data signals. It is the internalization of all the input data that allow the RRC-robot to identify, recognize, visualize and comprehend the input signals and patterns as does a human.

3.2.2. The Operation of the RRC

The auditory search engine is used to search the multi-dimensional auditory Nodal Map Module of the RRC-robot for a-f-t-sound patterns that will be recognized by the RRC as a Task Initiating Triggers (TIT). The TIT-patterns are used to activate any of the tasks listed in the Hierarchical Task Diagram (HTD) shown in **Figure 2**. Each of the prime level tasks, shown in **Figure 2**, has a programmed Task Selector Module (TSM) associated with it. The dominant electronic component of each TSM is a pattern recognition circuit that is programmed to recognize and prioritize the TIT-pattern detected by each of the TSMs as they operate on each of the input Nodal Map Modules. The total collective of TSMs shown in **Figure 2**, form the declarative memory system of the auditory RRC-Humanoid Robot. The programming/training of the auditory RRC-robot is a process of training the pattern recognition circuits of each TSM associated with each task, to recognize, identify and prioritize input-signal TIT patterns that initiate the lower level tasks shown in the figure.

3.3. The Search Engine Mode of Operation

The operation of a multi-tasking RRC-robot has been described in the literature [5]-[7]. The following presents some aspects of the operation of an RRC-Robot that are applicable to the behavioral programming of the system. An RRC-Humanoid Robot operates by searching the environment for tactile signals or patterns, visual signals or patterns or auditory signals or patterns. The declarative auditory search engine, shown in the Hierarchical Task Diagram (HTD) of **Figure 2**, operates concurrently to guide the robot in the performance of the verbal tasks listed in the figure. The declarative auditory search engine, shown at the top of the figure, searches the sound environment for phoneme-based sounds consisting of words and sentences described by the a-f-t signals shown

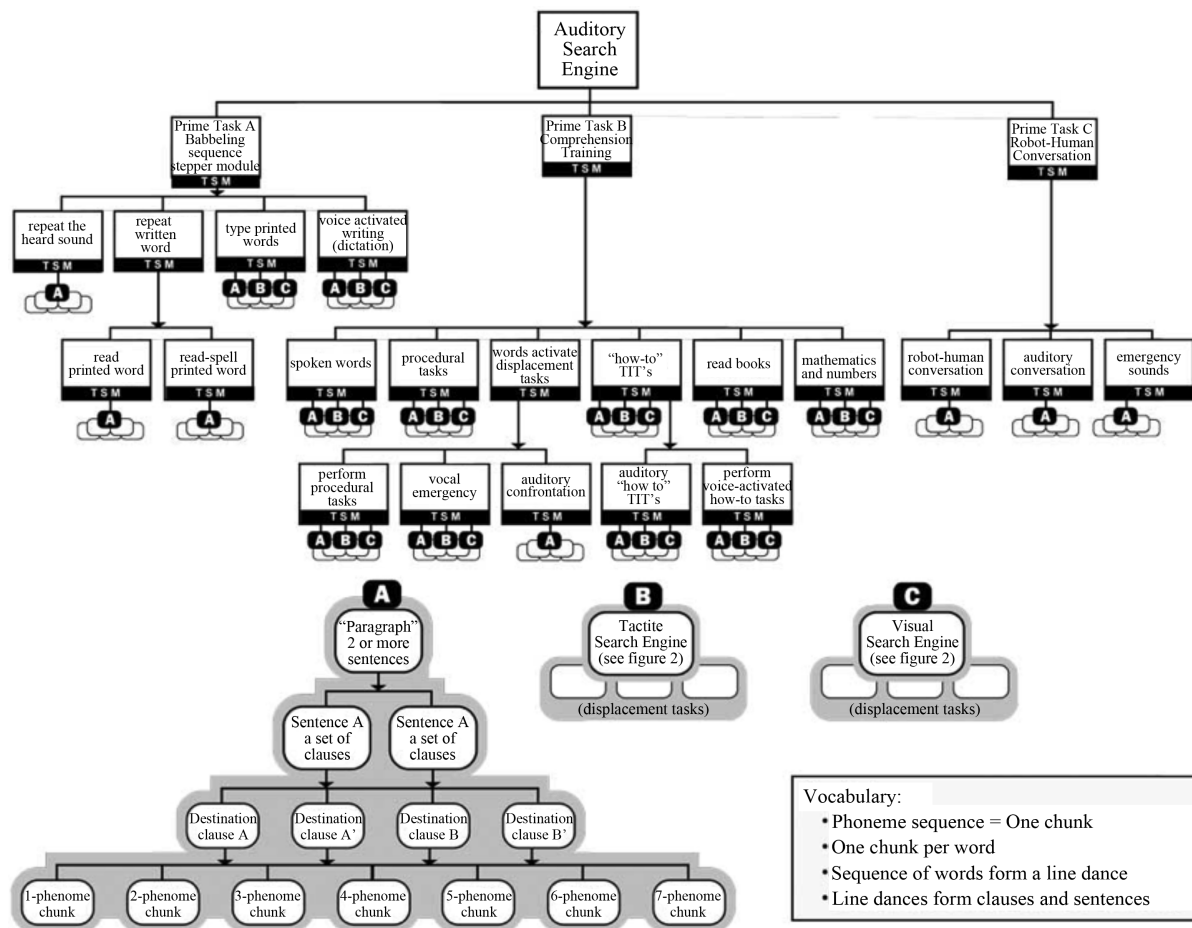


Figure 2. The declarative HTD: The HTD is the top-level specification of the system. This figure shows the TSMs of the auditory search engine. Those TSMs form a declarative memory system within the RRC.

in **Figure 1**. The Task Selector Module (TSM), associated with each task and subtask, is shown as a black bar at the bottom of each task or subtask shown in **Figure 2**. Each TSM consists of pattern recognition circuits that are programmed to activate only one TIT-task of the number of TIT-tasks that are directly below it. For example, the TSM of the top most Auditory search engine may generate TIT-a-f-t-words or sentence sound patterns that activate prime task A, or TIT-a-f-t-words that activate prime task B, or TIT-a-f-t-words that activate prime task C. For example, if the TSM of the auditory search engine, recognizes the a-f-t-pattern associated with the TIT-words “repeat the heard sound,” then the TIT-signal is transmitted to prime task A-1 for further processing. Note that after completion of training, if the robot is confronted with verbal phoneme sounds, but cannot recognize either the speaker or the words, then the signal is transmitted to prime task C-2, (the auditory confrontation mode). Note that TSMs are associated with each level of the HTD shown in **Figure 2**. Each TSM, at each level is activated by the TSM at a level above it. Each TSM consists of pattern recognition circuits that recognize TIT-words and sentences, and thereby transmit them to another TSM-level either for further processing or for activation of the verbal phoneme sound generator or robotic body or limbs. The steps for training/programming of the TSMs of the declarative memory system are shown in **Figure 3**. Note that the programmed/trained TSMs associated with the tasks in **Figure 3**, form a declarative memory system within the controller that remembers the procedure for performing the various subtasks listed in the HTD.

3.4. The Memory Systems within the RRC

Learning and memory is generally classified as procedural (or reflexive) if the learning or memory involves motor skills, and declarative if the learning or memory involves verbal skills. In the multi-tasking visual-RRC-robot, procedural TITs operate in the motor-joint Nodal Map Module, and procedural memory is the basis for all

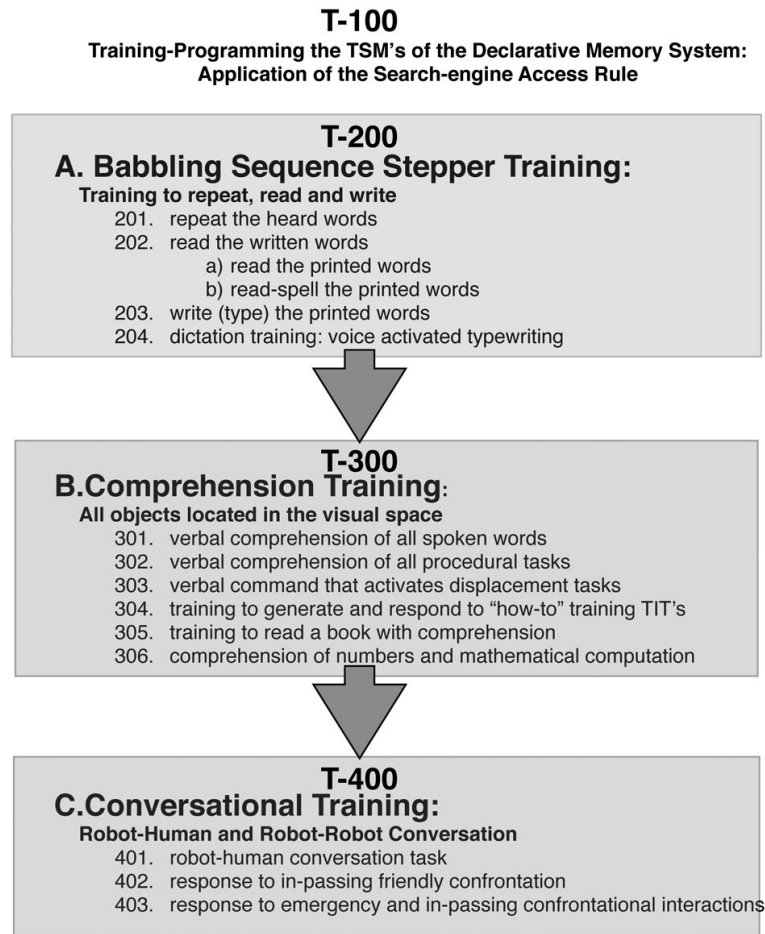


Figure 3. Training-programming the TSMs of the declarative memory system.

the control functions of the somatic motor system. In the Auditory RRC-Humanoid Robot declarative memory is the basis for all the control functions of the verbal-phoneme sound generator. **Figure 2** shows the TSMs associated with the declarative memory system. The programmed/trained TSMs shown in **Figure 2**, give the robot the capability to “remember how” to perform all the auditory sub-tasks listed in the figure. The declarative memory system includes a robotic capability to 1) repeat, read and write all words and sentences presented to the robot; 2) comprehend and identify and describe verbally all nouns, adjectives, verbs and adverbs that are presented to the robotic visual and tactile systems; and 3) perform robot-human conversation with comprehension.

4. Speech Processing: Recognizing the Acoustic Sequential Set of Phoneme-Signals as Phonetic Words and Sentences

4.1. The Problem

The problem of converting the perceived acoustic spectrographic (a-f-t) properties of language, shown in **Figure 1**, into an identifiable phonetic structure is an ill posed problem, similar to the 3-dimensional inverse optics problem [9]. There is not a simple one to one mapping between the acoustic properties of speech and the phonetic structure of an utterance. Co-articulation (the segmentation problem) is generally identified as the major source of the problem. Co-articulation gives rise to difficulty in dividing the acoustic signal into discrete “chunks” that correspond to individual phonetic segments. And it also gives rise to a lack of invariance in the acoustic signal associated with any given phonetic segment. Note: The usual methods for solving the problem includes lexical segmentation processing (co-articulation), word recognition processing, context effect processing, syntactic effects on lexical access processing, lexical information and sentence processing, syntactic processing, and intonation-structure processing.

4.2. The Behavioral Speech Processing Methodology for Solving the Inverse Auditory Problem

Because of the complexity in the mapping between the acoustic signal and phonetic structure, an experiential, behavioral programming methodology was developed for “unpacking” the highly encoded, context dependent speech signals. “Unpacking” is performed in the Interface Circuit by programming the RRC to repeat and “remember” (in the TSM-memory modules) the ‘heard’ words and sentences of multiple speakers.

4.2.1. Repetition and Babbling the Words and Sentences Taken from a 50,000 Word Lexicon (Task 201 in Figure 3)

The first step for training the auditory RRC-robot is the requirement for a “babbling” sequence stepper module and an associated TSM that is trained to accurately and quickly repeat the sound of words, strings of words, or sentences heard by the robot. The trained repetition and babbling sub-task A-1 TSM, activates the total vocabulary of the robot. All the words or sentences spoken by the robot and activated by other prime task TSMs must access the sub-task A-1 TSM and form a compound TSM that does not necessarily repeat the sound but accurately enunciates other words and sentences (taken from the sub-task A-1 TSM) and associated with the compound TSM. Most of the design activities of the task T-201 shown in **Figure 3** are aimed at achieving enunciation accuracy in the repetition and babbling sub-task A-1 TSM. In order to achieve repetition accuracy it is necessary to refine the design of the phoneme sound generator, expand the number of phoneme sounds listed in the 120 phoneme sound combinations utilized in the preferred embodiment RRC-Humanoid Robot, and refine the tuning of the spectrum analyzer to the actual collective modalities present in the English language verbal input signal.

4.2.2. Additional Speech Processing

Further “unpacking” is performed by behavioral programming techniques that includes the following: First, by relating, correlating, associating and calibrating the heard verbal speech with the corresponding visual and tactile data obtained in the visual and tactile coordinate frames in which the robot is operating. Next, by training the RRC-Robot to be sensitive to such factors as acoustic phonetic context, speaker’s “body language”, speaking rates, loudness and “emotion laden” intonations. The Auditory RRC-Humanoid Robot takes into account the

acoustic consequences of such variations when mapping the acoustic signal onto the phonetic structure. The problems of speaker's "body language", "emotion laden" intonations, acoustic phonetic context, speaking rates, and loudness is solved in the Auditory RRC by coordinating the search engines of the visual and tactile systems with the search engine of the Auditory RRC-Humanoid Robot.

5. Behavioral Programming of the (Speech) Processed Words and Sentences

Once a sequence of words or a sentence is recognized as a TIT (in the TSM) and are projected to the Multi-dimensional RRC-Nodal Map Module, the RRC-Humanoid Robot is behaviorally programmed to control body and limbs, and the verbal phoneme sound generator in relation to the words and sentences applied to the Multi-dimensional Nodal Map Module. Words and sentences are identified, recognized, and comprehended by behaviorally programming the RRC-Robot to generate different words by the verbal phoneme sound generator or undertaking distinguishing body or limb actions based on the words or sentences applied to the Multi-dimensional RRC-Nodal Map Module.

5.1. The Programming of the Auditory RRC-Robot

Programming/training the RRC-robot is a process of training the pattern recognition circuits of each TSM associated with each prime level task and all the TSMs associated with the sub-tasks listed under the prime level task (see [Figure 2](#)). The pattern recognition circuits must recognize, identify and prioritize input-signal TIT patterns that initiate the prime level task and all the lower priority TIT-sub-tasks that are listed under the prime level task. The programmed TSMs associated with all the tasks in [Figure 2](#) and [Figure 3](#) give rise to a declarative memory system within the controller. Training the declarative memory system of the auditory RRC-robot is presented in the following sections.

5.2. The Search Engine Access Rule for Training the Declarative TSMs

The declarative memory system of the RRC-robot is made up of an array of TSMs with each TSM storing a large number of a-f-t-words phrases and sentences that represent the total vocabulary of the robot. In order to respond verbally with appropriate words and sentences the RRC must analyze the verbal input data, search through the memory TSMs, find the set of TSMs that have parts of the answer stored in them, form a compound TSM that has the total explicit word-answer stored in it, and activate the appropriate word answer that is stored in that compound TSM. The following programming rules have been devised in order to facilitate the search for an appropriate response to any auditory input signal.

- 1) Search the input signal to determine which TSMs are likely repositories of the appropriate verbal response.
- 2) Form a compound set of TSMs wherein the response may be stored.
- 3) Utilize the data present in the auditory input signal and in the compound set of TSMs to home in on an appropriate response.

For example, the application of the auditory search engine access rule when the trainer-supervisor requests the Robot to "identify this visual image", leads to an identification of two TSMs and a compound TSM. The two TSMs are most likely the visual image pattern-TIT presented to the robot (Task T-301 in the B-1 TSM), and the repeat this sound—verbal word or phrase that describes the presented visual image (task T-201 in the A1 TSM). Note that the sub-task A-1 TSM stores and properly enunciates all the nouns, adjectives, verbs and adverbs taken from the 50,000 word lexicon. The compound TSM is formed in the programmatic development of the access rule and includes the phrase "I see an—," wherein the training selects that word or phrase from the A-1 TSM that describes the presented visual image.

5.3. Behavioral Programming Procedures

Behavioral programming procedures are performed on all the TSMs shown in [Figure 2](#). An example of behavioral programming procedure for the sub-task A-1 TSM, shown in [Figure 2](#), will be described in the following subsection.

Sub-Task A-1: TSM-Training to Repeat Phoneme-Sounds Spoken by the Trainer-Supervisor (sub-task T-201 [Figure 3](#)). The trained prime task A-TSM is a memory module that stores all the TITs that identify and properly enunciates all the words listed in the lexicon and the commonly used combinations of words, clauses

and sentences selected by the trainer-supervisor. All subsequent verbally generated tasks must access the TITs stored in the prime task-A TSM. The properly enunciated words and phrases taken from the Task A TSM are then associated with other TITs generated by the visual system, the tactile system, the olfactory system, the gustatory system, or other word TITs generated by the auditory system.

- Repetition and Babbling—The auditory RRC is trained to repeat via the phoneme sound generator, the words and sentences spoken by the trainer-supervisor.
- The high priority TIT that shifts the robot to Sub-Task A-1 is a simultaneous visual recognition image of the trainer, and the command “repeat this sound” spoken by the trainer.
- All the words and sentences repeated by the robot are taken from a 50,000 word lexicon that represents the total vocabulary of the robot.
- The lexicon or vocabulary of the robotic controller consists of the set of words and sentences that the sub-task A-1 TSM has been trained to repeat.
- The trained sub-task A-1 TSM is a memory system that properly enunciates all the words and sentences listed in the lexicon.
- Optimization of the sub-task A-1 TSM to properly enunciate all the words and sentences listed in the lexicon entails 1) refining the design of the phoneme sound generator to assure that the lexical segmentation and timing intervals between successive phonemes are optimized. 2) Expanding the selected number of phoneme sounds to optimize co-articulation problems. And 3) refining the tuning of the spectrum analyzer to the actual collective modalities present in the English language verbal input signal.
- The sub-task A-1 TSM memory system is always accessed by other TSMs in order to form compound TITs whenever verbal sounds other than the “repeat this sound” TIT are to be generated by the robot.
- Queries that access the sub-task A-1 TSM generally relate to the verbal enunciation of the words and phrases stored in the A-1 TSM. Therefore, in anticipation of such questions, acceptable and not acceptable grammatical structural forms of verbal enunciation should be programmed into the search engine for each word or phrase in the lexicon; and the specific answer to each anticipated query must be programmed into a compound TSM.

At this point the controller has performed all the speech processing that allows it to recognize and repeat, but not comprehend, all the phoneme constituents of words, sentences and clauses listed in the lexicon. The auditory RRC-monitor has thereby mapped the acoustic signal onto a linguistic representation that is amenable to declarative comprehension (in prime tasks B and C).

6. Reduction to Practice Experiments

Reduction to practice experiments were undertaken to validate the hardware and software developments of an RRC-Humanoid Robot that exhibits human-like levels of AI. In order to achieve human-like levels of AI the RRC-robot must adhere to the six requirements listed in 1)-6) (Section 1). Since the robot is trained/programmed behaviorally, the total mechanical robotic body and sensors must be built and operational before training/programming can begin. However, it is possible to validate that the Robot meets all the requirements listed above by dividing the reduction to practice experiments into three sequential phases:

Phase-1. The motorized robotic arm-wheelchair system: A motorized power wheelchair outfitted with a robotic arm and two human-like sensors, was fabricated and tested during the phase-1 sequence. **Figure 4** presents a photograph of the system. The mechanical robotic body and the bi-pedal legs of the RRC-Humanoid Robot are simulated by the motorized power wheelchair outfitted with a robotic arm and two human-like sensors. The sensors consist of pressure transducers distributed on the surface of the robotic wheelchair that simulate the human tactile sensors, and a set of video-visual cameras that simulate the human eyes. Behavioral programming techniques, described in phase-2, may be used to validate the performance of the system for the tactile and visual sensors.

Phase-2. Behavioral programming the phase-1 system: The training consists of programming experientially, itch-scratch-self knowledge into the system, and training the visual system to “visualize” objects in the FOV as does a human. All visual training is performed experientially, first by repeating the itch-scratch training to avoid visually observed obstacle along the itch scratch trajectory, then to identify and manipulate objects in the FOV of the system, and finally to drive/control the motion of the wheel chair in an urban environment. When the visual training/programming is completed the motorized robot will have the human-like intelligence to drive and navigate the wheelchair through sidewalks crowded with pedestrian traffic, and through heavy city traffic.



Figure 4. A photograph of the RRC-controlled motorized power wheelchair. A 4-degree of freedom robotic arm, mounted on the wheelchair, is shown in the photograph. The tactile system is represented by pressure transducers taped onto the wheelchair, with one pressure transducer located under each patch of tape. The visual system, presented in greater detail in [Figure 7](#), is also shown mounted on the wheel chair.

Phase-3. The auditory verbal system: Phase-3 includes the addition of two pieces of hardware to the phase 1 system. The hardware consists of pick-up microphone auditory sensors, simulating the human ears, and the addition of a verbal phoneme sound generator. The experiential programming/training of the system to hear and understand verbal speech and to converse intelligently is described in Section 5.

The following sections present a progress report for the performance of the 3-phase reduction to practice validation experiments.

6.1. Phase-1-Progress Report: Description of the Motorized Robotic Arm-Wheelchair System

A photograph of the motorized power wheelchair outfitted with a robotic arm and 2 human-like sensors is shown in [Figure 4](#). An illustrated block diagram of the proof of concept Robot is shown in [Figure 5](#). The experimental RRC-Humanoid Robot consists of 1) a motorized wheel chair modified by the addition of a metallic frame used to carry sensors, power supplies and the RRC controller. 2) Tactile pressure transducers that cover the surface of the motorized wheelchair (see [Figures 4-6](#)). Those pressure transducers simulate the mechanoreceptors distributed on the surface of the human skin. 3) The robotic arm is shown in [Figure 6](#) and is shown mounted on the wheel chair in [Figure 4](#) and [Figure 5](#).

The robotic arm is used to behaviorally program the self location and identification coordinate frame into the system, and to manually drive the motorized wheelchair by pushing the 2-degree of freedom joystick-controller of the wheelchair (see [Figure 6](#)). The wheelchair joy-stick is used to control the speed and direction of motion of the wheelchair. In addition, the robotic arm is programmed to control the wheelchair on-off power switch, the wheelchair safety brake, the wheelchair horn, and the wheelchair re-charging system. 4) The video-visual system used to simulate the human eyes is shown in [Figure 7](#) and is shown mounted on the wheelchair in [Figure 4](#) and [Figure 5](#).

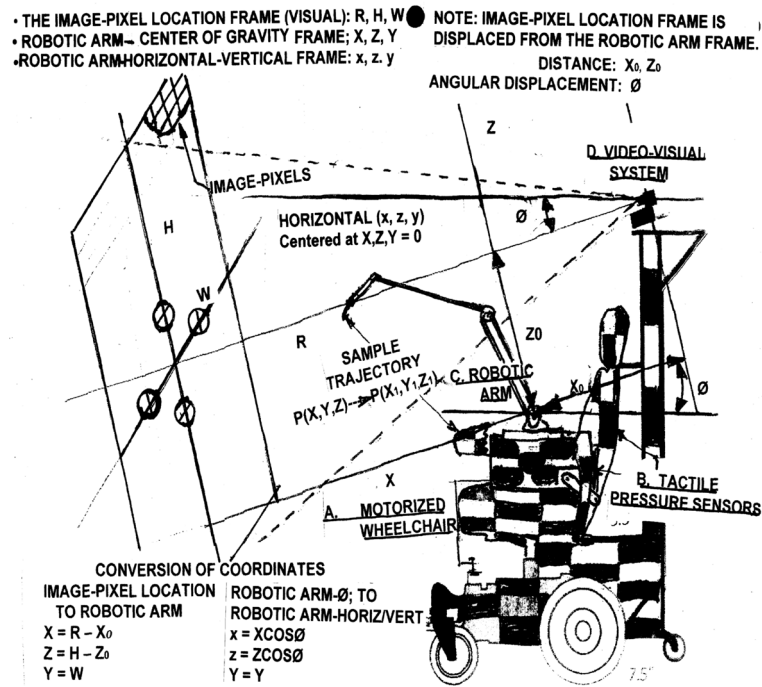


Figure 5. An illustrated block diagram of the proof of concept system showing the coordinate frames of the tactile and visual system.



Figure 6. A close up view of the robotic arm used to behaviorally program proprioceptive knowledge into the system.

Although the dynamics of the phase-1 system require improvements, the design is sufficiently complete that phase 2 behavioral programming of the tactile and visual sensors may be initiated, as described below.

6.2. Phase-2-Progress Report: Behavioral Programming of the Tactile and Visual Sensors

The primary reason for building the robotic-arm-wheelchair system is to validate the behavioral programming techniques performed with the tactile sensor data, and with the video-visual camera data.

6.2.1. Behavioral Programming of the Tactile Sensors

The methodology for programming a self-knowledge coordinate frame into the system is described in Section

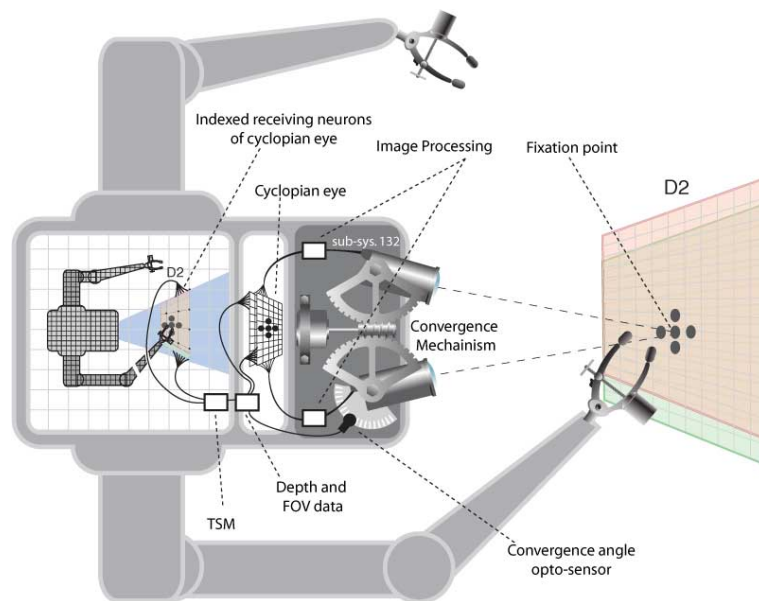


Figure 7. Detailed mechanical design of the video-visual system mounted on the motorized power wheel chair. The figure shows the convergence mechanism and correspondence matching of the system.

2.1. Behavioral programming of the tactile near-space is performed by first defining the coordinate frame in which the robotic limbs are controlled, and then programming the “itch-scratch” motion of the limbs in the near-space of the robot. **Figure 5** shows the tactile near space coordinate frame of the robotic arm-wheelchair system. The center of the XYZ coordinate frame, at the point $P(0,0,0)$, shown in **Figure 5**, is located at the base of the robotic arm and in the central region of the wheelchair. The coordinate scale size was chosen to be one centimeter per scale division. For reference, the location of the robotic finger is shown at point $P(X,Y,Z)$, and assuming a “itch” point at $P(X1,Y1,Z1)$, the trajectory of the robotic finger, necessary to scratch the itch point, is shown in the figure (from $P(X,Y,Z)$ to $P(X1,Y1,Z1)$).

The nodes in the coordinate frame, defined by integral values of XYZ at the $P(X,Y,Z)$, are particularly important. (Examples of integral nodes are $P(0,0,0)$, $P(1\text{cm},0,0)$, $P(1\text{cm},2\text{cm},3\text{cm})$ etc). The nodes quantitatively define the near space (the spatial extent of each robotic end-joint) and the self-location and identification coordinate frame of the system.

Programming the self-location and identification coordinate near space requires that the indexed location of each node be entered into the RRC, and the following data be stored at each indexed nodal position: 1) each node should include a table of 26 control vectors that may be used to transition an end-joint, such as a robotic finger, to one of the 26 adjacent nodal positions. (Note: The correct magnitude and direction of each control vector is first determined experimentally, and then refined and corrected during the training/programming process). 2) Each node must be programmed to detect whether an activation (itch) occurs at that point, or whether an end-joint is located at that point, and the magnitude of the itch, or location of the end-joint. 3) If an itch-point is activated at any given node, that node becomes a Task Initiating Trigger (TIT), and the final position (q-final) of the itch-scratch” trajectory. 4) The Task Initiating Trigger (TIT) generated by the RRC forms a command, applied to the end-joint, located at q-initial, that causes the end joint to move along the trajectory from q-initial to q-final. For example if the robotic finger is at $P(X,Y,Z)$, and a itch occurs at point $P(X1,Y1,Z1)$, then the RRC-controller generates a sequence of control signals that propel the robotic finger from $P(X,Y,Z)$ to $P(X1,Y1,Z1)$, as shown in **Figure 5**.

At this time only about one third of the nodes in the near space of the robotic finger have been programmed. However experiments were undertaken to have the robotic arm control the motion of the wheelchair by pushing the joystick controller. This was accomplished by conditioning the desired control signal to an “itch” signal generated by the trainer and emanating from the joystick controller. The joystick thereby becomes the “itch” point, the TIT-q-final for the system. When the initial position of the robotic finger, q-initial is at a fully programmed

nodal position, and all the nodes between the itch-point and the initial position of the robotic finger are also fully programmed, then the robotic finger may be controlled by the RRC to touch/push the joystick so as to propel the wheel chair in a direction determined by the direction of the force exerted by the robotic finger. A video of the controlled motion of the wheel chair by the robotic arm is available for viewing at www.mcon.org.

6.2.2. Behavioral Programming of the Visual Sensors

The video-visual system of the RRC-robot system has been described and published in the peer-reviewed literature [7]. **Figure 7** taken from the referenced paper, shows the detailed layout of the system, whereas the mounting of the visual system on the motorized robotic arm-wheelchair robot is shown in **Figure 4**, and **Figure 5**.

Reduction to practice experiments to demonstrate behavioral programming of the visual sensors are performed in 3 steps:

Step 1. The FOV-space coordinate frame, shown as the image-pixel location frame in **Figure 5**, was defined and specified relative to the tactile-XZY coordinate frame.

Step 2. The two 2-dimensional (2-D) images obtained from the two cameras must be processed to form a 3D-photometric image that is a high fidelity representation of the object in the FOV that gave rise to the two 2D-images.

Note: That in our design, the image-pixels are distributed in 3-D in the tactile near-space coordinate frame. The image-pixels form a high fidelity representation of object size and object location, to the degree that step 1 is accurately programmed.

Step 3. All behavioral programming is performed based on the detection, identification, or observation of the characteristics of the 3-D photometric images.

At this time only step-1 and step-2 of the 3-step process have been performed. The following is a progress report of the status of our reduction to practice experiments.

Reduction to practice experiments are performed by first defining the FOV-space coordinate frame in which the imaged-objects are located, and then mechanically calibrating the FOV-space with the near space coordinate frame shown in **Figure 5**. **Figure 5** shows the FOV coordinate frame of the video-visual system mounted on the motorized wheel chair system. The following experimental tasks were performed:

1) The camera system was mechanically fabricated and mounted at the tilt angle θ (see **Figure 5**) so that the FOV of the visual system intersects a large segment of the self-location and identification coordinate frame described in Section 6.2.1. and shown in **Figure 5**.

2) The two video cameras were designed to converge at a point on the midline LOS. The point of convergence and the image plane associated with it is shown in **Figure 5** at a distance R from the two cameras.

3) The location of each image-pixel, shown in **Figure 5**, from each of the camera's CCD-arrays, is programmed into the RWH coordinate frame shown in the figure.

4) The object identification measurement involves identifying various patches of pixel activations as objects that are commonly encountered by the system. For example, other people, automobiles, curbs, stop lights, poles, fire hydrants, or just obstacles that may interfere with the system etc. A small number of objects have been identified in the laboratory, however the identification experiments and the visual cues depth measurements described in 5) and 6) are awaiting completion of the tactile experiments described in Section 6.2.1.

At this point the pixels on the image plane forms a 2D-image of objects in the FOV of the system. Note that only the image-pixel located at point $P(R, W, H) = P(R, 0, 0)$ is at a known correct depth. All other image-pixels displaced from the $P(R, 0, 0)$ central position may be located at greater or lesser depths than their position on the image plane.

The following steps, aimed at generating a 3D-photometric image, describe the experimental visual cues [7] that were programmed into the system during phase-2 in order to determine the depth of pixels that are offset from the $R, 0, 0$ position on the image plane. The binocular dispersion visual cue measurement is described in step 5), and the image size visual cue measurement is described in step 6).

5) The binocular dispersion visual cue measurement: For each value of R , that defines the location of an image plane and the position of the offset image-pixels on the image plane, the binocular dispersion (separation distance between the image pixels of the right and left camera) was measured as a function of W and a function of H . The dispersion as a function of R , W , and H was programmed into the system. Although at any image-plane distance R , the dispersion-depth of any offset pixel is not a unique nor strong function of W or H , it is possible to determine that any given patch of pixels are either not greatly depth-displaced or are greatly dis-

placed from the image plane.

6) The image size visual cue measurement: When an image-object is identified in 4) to be of known size in the self-space coordinate frame, the measured pixel-activated image size forms a strong function of the depth location of offset image-pixels. Image size measurements of a small number of identified objects have been performed, however the bulk of the image-size experiments are awaiting completion of the tactile experiments described in Section 6.2.1.

6.3. Phase 3-Progress Report: Design-Development of the Auditory Verbal System

The phase-3 hardware consists of a verbal-phoneme sound generator and a set of pick-up microphones feeding a programmable spectrum analyzer. Those systems have been designed, fabricated, but not tested. The testing procedure is described in Section 3.1., and requires the completion of all phase-1 and phase-2 tasks. The installation, and testing of the phase-3 hardware, and the behavioral programming of the auditory RRC-wheelchair system, will be described in a future follow on paper.

7. Conclusions

The purpose of the reduction to practice experiments was to validate the claimed visual and verbal-auditory intelligence levels of the RRC-Humanoid Robot. The visual intelligence level of the motorized visual wheelchair is validated when it demonstrates that it “sees” the environment as does a human, and is capable of using the robotic arm to control and navigate the robotic wheelchair system through all difficult traffic environments.

The behavioral programming of the motorized visual wheelchair is only partially complete, however the performed experiment, and the experimental results validate that the visual system of the robotic wheel chair, does indeed “see” the environment, as does a human. And based on that visualization it is likely that follow-on validation experiments will prove that the system may be successfully programmed to navigate through all difficult traffic environments, as does a human.

The navigation of the robotic wheel chair is unique because it utilizes only a visual system that simulates the human eyes to navigate the robotic system. Most other systems generally utilize radar, lidar, sonar, or GPS data, in addition to visual camera data, to navigate driverless robotic systems. The validation experiments described in the previous sections prove that the human-like visual system gives the robot human-like data that the system internalizes (has “self-knowledge of”, or has human-like “awareness of”), and most likely, has a discrimination or discernment capability that equals or exceeds the discrimination capability of data generated by a combined radar, lidar, sonar or GPS data systems, used in other driverless motorized systems.

The visual wheelchair system data are human-like for the following reasons (demonstrated by the validation experiments described in the previous sections):

- 1) The data gathering equipment and methodology is humanlike:
- 2) The data are imaged on two CCD-arrays that exhibit the binocular disparity of the two retinas of the human eye.
- 3) A 3D-photometric image is formed within the controller that is a high fidelity representation of the objects that gave rise to the image.
- 4) The image visualized by humans is a 3D-illusion formed in the brain that is a high fidelity representation of the objects that gave rise to that image.
- 5) Therefore, by symmetry arguments, the 3D-photometric image is identical to the 3D-illusion formed in the brain. Thus, if a human “sees” the 3D-illusion, then the robotic system behaves as if it too “sees” the 3D-photometric image,

The partially performed validation experiments prove that the robot detects aspects of the photometric image in a human-like manner with a high level of discrimination and resolution.

Validation experiments, to be reported in a follow on paper, will attempt to prove all the high IQ-verbal intelligence claims for the auditory RRC-Humanoid Robot.

Acknowledgements

The authors are grateful to MCon Inc. for allowing us to publish their proprietary data, which were accumulated during the past 20 years. This paper presents an overview/introduction to the design of smart RRC-Humanoid

Robots. The reader may obtain technical details relating to the design in the references included in this paper.

References

- [1] Turing, A.M. (1950) Computing Machinery and Intelligence. *Mind*, **59**, 433-460.
<http://dx.doi.org/10.1093/mind/LIX.236.433>
- [2] Turing, A.M. (1953) Solvable and Unsolvable Problems. *Science News*, **31**, 7-23.
- [3] Rosen, A. and Rosen, D.B. (2007) The Turing Machine Revisited: ASSC-e Archive. www.mcon.org
- [4] Rosen, A. and Rosen, D.B. (2013) A Behavioral Programming Approach to the Design-Development of Humanoid Robots. *Proceedings of the 2013 International Conference on AI, World Comp'13*, Las Vegas, 22-25 July 2013.
- [5] Rosen, A. and Rosen, D.B. (2006) An Electromechanical Neural Network Robotic Model of the Human Body and Brain. *Lecture Notes in Computer Science*, **4232**, 105-116.
- [6] Rosen, A. and Rosen, D.B. (2006) A Neural Network Model of the Connectivity of the Biological Somatic Sensors. *International Joint Conference on Neural Networks*, Vancouver, 299-306.
- [7] Rosen, D.B. and Rosen, A. (2007) A Robotic Neural Net Based Visual-sensory Motor Control System That Reverse Engineers the Motor Control Functions of the Human Brain. *International Joint Conference on Neural Networks*, Orlando, 12-17 August 2007, 2629-2634; *Journal of Financial Economics*, **7**, 197-226.
- [8] Rosen, A. (2005) The Design of the NCC-Circuit for Audition and Sound Generation. An Internal MCON Inc.
www.mcon.org
- [9] Marr, D. (1962) Vision. Freeman Press.

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or **Online Submission Portal**.

