IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Modeling and Recognizing Grounded response as events in Human Robot Interaction through iterative sensory data integration using semi autonomy algorithm and neuronally implementation with Kinematics Control using Deep Belief Network(DBN) in Deep Learning

Author: Dr. Ashok Kumar Ramadoss ,
MCA, MBA, MSc. Mat Sc., Ph. D, (International Award winner and Senior Member of Robotics)
Associate Professor, Coimbatore Institute of Technology, Anna University,
Scientist, Software engineer, Present Member in Society of Robotics Surgery , Orlando, Florida and Senior Member in UACEE New York and Hong Kong society of robotics and automation

Abstract

For Human Robot Interaction we believe that a robot should model and recognize a set of grounded responses that are built from knowledge about the nature of the interaction situation, and should also be able to ground responses that are found by semantics-free contingency detection.To carry out Research on Humanoid Robot effectively first in the absence of hardware support then building a Humanoid Robot for achieving data concurency by integerating data by the fusion of sensory datas in the humanoid robot .Multi Sensor fusion Studies for Dynamics, Kinematics thereby achieving Human Robot Interaction by sensory data integeration by sensor fusion for range detecting by vision and navigation using simulink model and humanoid robot prototype. For this a Simulation Platform could be designed which is easy to operate by using Virtual Reality Modeling Language technology and can be 3D simulated and the simulation results are more intuitive. Within the current framework, we can model and recognize grounded responses as events. As future work, we will investigate how to attribute semantics to ungrounded responses through iterative interactions and finaly by sensor fusion using pattern recognition and predicting with accuracy and reliability by bringing the source data closer to the analyzed data and predicted data with that of actual data using edge computing.An ideal contingency detector should be able to accept a variety of sensory cues, because certain perceptual and social cues are more informative for some interactional situations than for others.

In this Research, we could present a simple and reliable approach of creating humanoid robot platform based on the ROBO OS and modeling language using ubuntu Linux.Another goal is to investigate the general potential of SFA for using it within sensorimotor loops which to our knowledge has not been considered until now. The application of SFA within sensorimotor loops is motivated by pointing out its relation to second order Volterra filters. Our experiments show that the overall reactivity of the gait pattern increases without any profound loss in stability, and that SFA appears to be suitable for the usage even at such levels of sensorimotor control that are directly involved into motor activity regulation.

This work is concerned on sensitivity analysis of semi autonomy algorithm of humanoid robot to environmental sensors' failures. The construction of the robot, semi autonomy algorithm and used sensors have been described. The algorithm bases on a reactive hybrid approach that merges data from different types of sensors and calculates resulting velocities. This algorithm takes also into account environmental sensors' damage by modifying the behavior of robot in accordance to actual sensors' set state of health. Simulation research using ROBOS/Simulink package and experimental tests' results of semi autonomy algorithm were presented. The experimental tests were carried out in outdoor conditions. The research and tests were performed for normal environmental sensors' operation and for selected sensors' damage. On that basis, sensitivity of semi autonomy algorithm to selected environmental sensors damage was tested.

Keywords- Human Robot Interaction, Contingency/Response detection, Sensory data, Sensor fusion Integration, Vision, Data Concurrency, Sensorimotor,DBN,Introduction

Making Humanoid Robot Research is very challenging. The Cost factor could be optimized for building a humanoid robot prototype it could be better to analysis humanoid robot through simulation before building a humanoid Prototype[1-4].
On the basis of the effective simulation the final Protype can be made, for doing this design of the simulation platform is a very essential factor which makes to study the dynamics, kinematics and control method on simulation platform and then validate to prototype model[5-9]. So that the adverse financial loss can be avoided on building the model. It has the high order,coupled,variable structure, variable parameter and nonlinear.When a robot understands the semantics of a human's activity in a given context and has a specific expectation over a set of known possible actions, then checking for a contingent response might simply entail matching the human's executed action against the expected action set. This strategy of a contingent behavioral change by a human can occur in one or multiple communication channels. For example, a robot that waves to a human to get his attention may receive the speech of "Hello!" with a simultaneous wave motion as a response. Here , we consider the problem of human contingency detect with multimodal sensor data as input when forming our computational model. We validate the multiple cue approach using multimodal data from a turn taking scenario and show that modeling a response using multiple cues and merging them at the appropriate levels leads to improvements at the appropriate levels leads to improvements in accuracy in contingency detection. accuracy in contingency detection. A contingent behavioral change by a human can occur in one or multiple communication channels. For example, a robot that waves to a human to get his attention may receive the speech of "Hello!" with a simultaneous wave motion as a response. Here, we consider the problem of human contingency detection with multimodal sensor data as input when forming our computational model. We validate the multiple-cue approach using multimodal data from a turn-taking scenario and show that modeling a response using multiple cues and merging them at the appropriate levels leads to improvements in accuracy in contingency detection. Collectively with our previous work , the results in this paper demonstrate that our behavior-change-based contingency detector provides a highly indicative perceptual signal for response detection in both engagement and turn-taking scenarios.

1. Need of the Study

In this paper, we make the following contributions.

a. We present a contingency detection framework that integrates multiple cues, each of which models different as-pects of a human's behavior. We extend our prior work, which only uses the visual cues of motion and body pose, to create a more generic contingency detection module. In our proposed framework, the cue response can be modeled as either an event or a change[10].

b. We propose three different levels of cue integration: the frame level, the module level, and the decision level. We show that for change-based detection, integration of visual cues at the module level outperforms integration at the decision level[11].

c. We examine the effects of selecting different timing models and referent events. In particular we show how selecting the minimum necessary information referent event, improves detection and requires a smaller amount of data, increasing the tractability of the real-time detection problem[12].

d. We provide a probabilistic method for measuring the re-liability of visual cues and adaptively integrating those cues based on their reliability[13,14].

e. We evaluate our proposed contingency detection framework using multimodal data and demonstrate that multi-cue contingency detection is a necessary component for interactions with humans and their multi-modal responses[15].

In a recent paper, we have successfully shown that SFA can handle many kinds of sensory qualities by applying it to abstract visual features, acceleration sensor and motor position data from humanoid robots (Spranger et al., 2009). SFA extracted meaningful components from the multisensory input data stream, which were employed for detecting and classifying postures of humanoid robots.

The remaining paragraphs will cover the following topics: We begin with a short introduction to the Slow Feature Analysis and Volterra filters, pointing out why quadratic filters are a well-motivated choice for our purposes. Next, we pass over to the experiments section, presenting the robot platform that was used for our experiments, and describe the examined gait pattern and our modifications to it. In the last section we present our results and show that the modifications performed prove useful for increasing the robot's reactivity without destabilization of the gait. We conclude this article with a summary of the obtained results and by giving insights into future work.

2. Training the algorithm for data patterns in the Framework plan for the humanoid

Within the current framework, we can model and recognize grounded responses as events. As future work, we will investigate how to attribute semantics to ungrounded responses through iterative interactions. An ideal contingency detector should be able to accept a variety of sensory cues, because certain perceptual and social cues are more informative for some interactional situations than for others. Here we extend the modalities supported by our framework to include the audio channel and detail an approach for the integration of multiple cues[16-18].deep belief network in deep learning using multi layer perception for transformation and data.

3 . Proposed Modelling and Analysis of Humanoid Mechanical System Model

This provides a large number of real time system components such as bodies joints, constraints, coordinate systems, actuators and sensors. This can be utilized to analyze a mechanical system separately or to perform a Comprehensive simulation between any Simulink controller and other dynamic systems. Simulink can be done with the basis of sim mechanics and it is the analysis and research environment of the controller and the system is multidisciplinary[19].

4 . 3D Modelling

3-D Modelling is the graphics and the descriptive modeling language used to create a real model [20,21].

Its objects contain 3-D, midi data, mpeg image and interaction diagrams. Objects are called the node which contains the basic element of field and event. The field is the parameter in the node and the event is utilized for the parameter transfer[22].

5 . Platform Design - Prototyping Humanoid Model

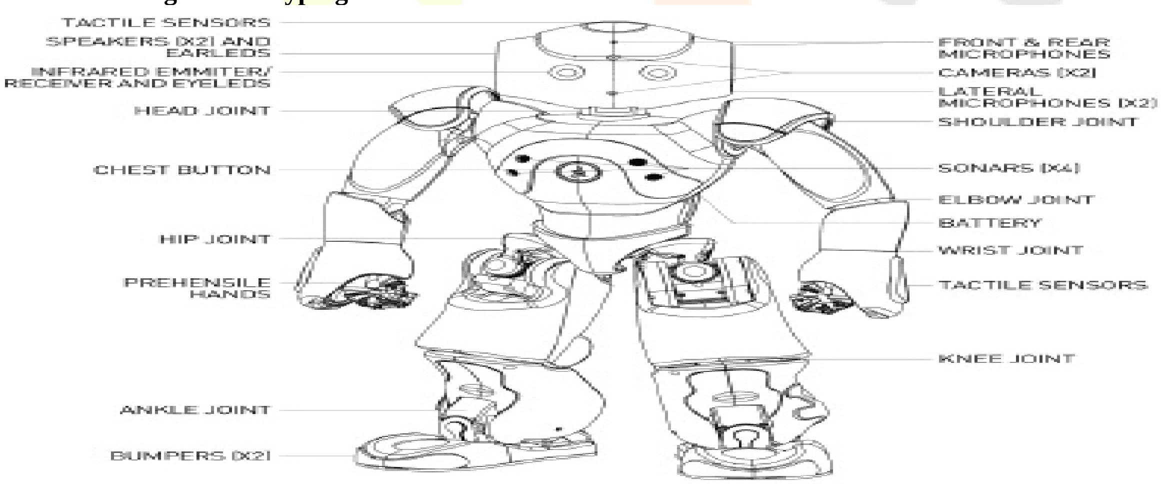


Fig.1. Humanoid Robot Prototype

To verify the result consistency, synchronizing Simulink data in 3-D Modelling of the virtual modeling language could be done. Initially we have to compile the Robots using Linux Ubuntu program of the algorithm and exchange the functional date with the Simulink data.

6. Create Humanoid Robot Model

[1]. technique is to utilize V-Realm Builder in the mat lab, [2]. Another is to create the whole robot model and then export the model as a VRML file for simulation, [3]. Since this technique is to shape the robot as a whole, therefore, any one link parameter (such as link length or (shape) changes, other link parameter will change automatically .This lead to modify the link, and the workload is enormous,[4].The last one is to create the each link of the robot respectively and stored as separate .wrl file, then to read these link files to form a complete robot . This simulation version is then set for the ROBOOS with Linux Ubuntu environment and

the settings for this simulation version of the robot for movement with respect to vision sensitivity and the data concurrency is taken at each step and validated using algorithm Kalman filter for avoiding the ungrounded data and the grounded data into the sensory state in analyzed with the experimental results. [5]. Due to the enormous workload, this research is proposed to utilize the first technique to create the robot model..

6.1 Plan for Creating Simulink Module

The procedure of create the Simulink module could be as follows:

(1) Selecting Ground, body and Joint module.(2) Setting the environmental parameter in the Machine Environment, such as the gravity vector. (3) Setting the base coordinates, could consolidate the robot system in the inertial coordinate system.
(4) Setting the link parameter in the Body module, such as weight, inertia tensor matrix. (5) Connect could be linked by the Revolute. (6) Connect, select and configure the Actuator and Sensor module. Actuator module could be utilized to input the Simulink signal to the joint. Sensor module is the measurement module, which could input the angle, speed, acceleration, and force of the joint to the Simulink.

6.2 Connecting the Robot Model and the Simulink

In order to realize the overall simulation platform of the humanoid robot virtual reality model and the Simulink, it could be needed to put the robot model created in the V-Realm Builder into the Virtual Reality Toolbox, then drive the model by Simulink. The procedure could be as follows:

(1) Setting a new Simulink file.(2) Adding the VR Sink and VR Signal Expander module to the file. The number of the VR Signal Expander module could be similar to the joints to be controlled.(3) Add the robot model to the VR Sink module.(4) Select the joints. This research could select the shoulder and hip joints, which could be showed in the arm1, arm2, R_leg, L_leg node.

6.3 Related Work to be done

Prior work as in literature survey has shown how contingency detection can be leveraged by a social robot to learn about structure in functional or social environments. Contingency detection has been used to allow robots to learn human gaze behaviors and to understand social interactions . Butko and Movellan et al [16] broke down the problem of contingency detection into two sub problems: response detection and timing interpretation. They developed a model of behavior that queried the environment with an auditory signal

7 Interpretation

Similarly, Gold and Scassellati focused on learning effective timing windows for contingency detection, which they did separately for an auditory signal and a visual signal.

There is evidence that humans use this contingency mechanism to learn some causal relationships. Watson found that infants use this mechanism for self-recognition and for detection of responsive social agents .A robot can also use contingency for recognition of self and others . Prior work has shown how contingency detection can be leveraged by a social robot to learn about structure in functional or social environments. Contingency detection has been used to allow robots to learn human gaze behaviors [5– 7] and to understand social interactions [1, 2, 8].Yang Yu-biao, [7] broke down the problem of contingency detection into two subproblems: response detection and timing interpretation. Thus, their focus was on the timing constraints of the contingency problem. Similarly, Gold and Scassellati [9] focused on learning effective timing windows for contingency detection, which they did separately for an auditory signal and a visual signal Sumioka H, Yoshikawa Y et al [14].There is evidence that humans use this contingency mechanism to learn some causal relationships. Watson found that infants use this mechanism for self-recognition and for detection of responsive social agents[[10, 11]. A robot can also use contingency for recognition of self and others [12, 13]. In this formulation of the problem, the presence of a visual change is directly mapped to the presence of a response. Thus, the presence of the response is assumed, but the source of the response is unknown and is attributed using timing interpretation.

In our work, we formulate the problem slightly differently. Because we are interested in human-robot interaction domains, such as engagement detection and turn-taking, we focus on a single response source: the human. We also can-not assume that visual changes always indicate responses, as visual changes frequently occur without correlating with contingent responses. Detecting contingency thus requires more complex analysis of what visual changes are observed. We implemented contingency detection using single vision-based cues and demonstrated the application of our contingency detection module as a perceptual component for engagement detection. Other work has focused on processing other individual channels, independently demonstrating the significance of gaze shift [14, 15], agent trajectory [16, 17], or audio cues [18] as contingent reactions.An ideal contingency detector should be able to accept a variety of sensory cues, because certain perceptual and social cues are more informative for some interactional situations than for others. Here we extend the modalities sup-ported by our framework in [3] to include the audio channel and detail an approach for the integration of multiple cues.Thus the multiple sensor sensitivity analysis is taken for the concurrency research.

8. Research Methodology

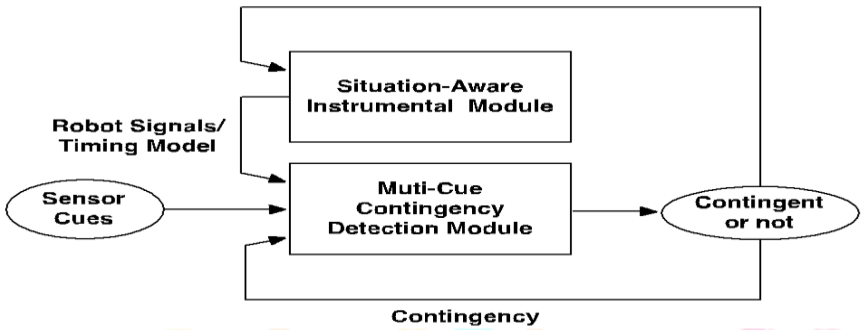


Fig.2. Situation-independent Contingency Detection.

Figure shows information flow between a contingency module and other robot modules. To make our contingency detection module as situation-independent as possible, we assume that the expected timing of a response is determined by another of the robot's modules and is taken as a parameter to the contingency detection module

Contingency detection consists of two sub-problems: response detection and timing estimation. Figure 2 shows the causal relationship between a robot signal and the corresponding human response as well as time windows for detecting such a response. A robot generates some interactive signal to a human by gesturing to, speaking to, or approach-ing her. Some time after the robot initiates a signal and in-deed sometimes before the robot completes its action, the human may initiate a response that could last for a certain duration. Zenon Mathews et al [6]

The time interval during which a human may respond to a given robot signal needs to be estimated. We define its end- points to be the time at which the robot begins to look for a human response, tRef , and the maximum time for the hu- man's response delay after tRef , (tRef + MRD). To detect a response, we evaluate sensor data within that time window. Because we are not considering anticipation on the part of the human, tRef is defined to be after tS ; to be effective it also should precede the initiation time of the human's response, tR . Chao et al. [4] define the notion of the mini-mum necessary information (MNI) moment in time when an actor—human or robot—has conveyed sufficient information such that the other actor may respond appropriately. We will describe this concept in more detail within the con-text of our experiments in Sect. 7 and in the results we will show the MNI is a more effective referent than the time at which the robot completes its signal.To detect human responses within the estimated evaluation time window, we take two different
modeling approaches: event detection and change detection. As shown in Fig. 2, we define WB and WA as time windows
from which data are used to model the human behavior before and after the robot's referent signal, tRef , respectively. An event detection models a response as an event and that event is looked for in WA . On the other hand, change detection models a response as a change, which is measured by comparing observations between WB and WA. However, in the contingent cases the human changes her behavior to make a contingent response, thus WB and WA model different behaviors. To detect such changes, we measure how likely that sensor data in WA reflects the same behavior observed in the sensor data in WB .

9. Concurrency Representation in Human robot interaction using sensor signals

To model a given aspect of human behavior, we derive in-formation from observations from a single or multiple sensors, hereafter referred to as a cue. Depending on the characteristics of the sensors used, a cue is encoded as either a binary variable or a continuous variable. When underlying sensors produce low-dimensional observations, the derived cue is easily represented as a binary variable. For example, audio and touch sensor signals can be classified as either on or off by simply thresholding the magnitude of the raw observations. Sensors that generate high-dimensional observations, such as image and depth sensors, require more complicated preprocessing, and thus a derived cue would be en-coded as continuous and high-dimensional variable thus this is the procedures for extracting cues from sensors in detail.

10. Multi-Sensor Concurrency Integration

The extracted cues from sensors should be integrated in such a way that the contingency detection module reduces un- certainty and increases accuracy in its decision-making. We adapt the data fusion framework introduced by Hall and Lli- nas et al. to integrate cues. Here, we define a frame as a representation of cue information and as an input to the contingency detection module. We define three levels of integration: (1) the frame level, at which cues are merged into one augmented frame as an input to the contingency module;(2) the module level, at which cues are integrated within one contingency detection module; and (3) the decision level, at which outputs from multiple single-cue contingency modules merge. These levels are shown in Fig. 4.

Cues should be integrated at the right level based on characteristics of a cue, the underlying sensor's sampling rate and dimensionality, and encoded perceptual information. If two cues are encoding the same perceptual modality of a human and they complement each other, thus modeling behavior in a more discriminative way, then two cues should be merged either at the frame or at the module level rather than at the decision level. Zenon Mathews et al. [6]

This multiple sensor integration is used by taking sensitivity analysis of various sensors like range detection sensor using vision sensor for data concurrency so that the correct and accurate data is captured by the humanoid so that the data is reached simultaneously inside the microcontroller of the humanoid brain so that the action through vision can be taken in a one by one continuous simultaneously so that the data reaches one by one to process the data then the required data is analyzed and the range detection is evaluated for a particular distance then immediately after recognition of the response the humanoid performs and then the next data reaches the microcontroller so that the second action like climbing the staircase to the next step is arrived so that the contingency in falling the humanoid robot is monitored for the path of the robot using the vison sensor.

11. Research Methodology and validation

Present Humanoid robot simulation are done by ADAMS or Robot toolbox since this is complex process ,This research uses Virtual Reality Modelling language and sim mechanics to design a simulation porotype of humanoid platform to do further research. Algorithms could be compiled through MATLAB or Java Programming or ROBOOS Linux Ubuntu.

Language to Simulink on the platform and the validity of the algorithm can be verified. Here ROBO OS used. A Kalman filter could be designed with Algorithms to integrate sensory data from three different types of sensors viz,image sensor ,microphone sensor and range sensor . Here the Kinematics + Gyroscope and its velocity for the achieving of grounded resposes thereby recognized using this kalman filter. Mihelj (B) et al J [8]

The Humanoid robot with the human intereaction is measured with its velocity in m/s vs time in t seconds.

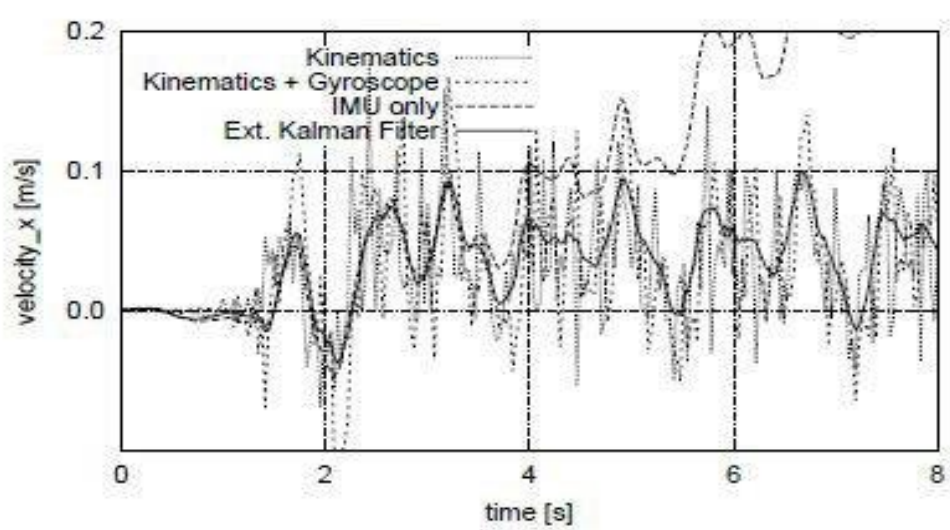


Fig .3 Kinematics and Gyroscope reading by Kalman Filter

The algorithm enables estimation of humanoid arm and leg posture, which can be used in trajectory planning for rehabilitation robots, evaluation of motion of robot with movement disorders, and generation of virtual reality environments, If such a distance metric is available, cues should be integrated in the frame level rather than at the Decision level. Mihelj (B) et al J Intell Robot Syst [8]

12. Response Detection

12.1 Event Detection

When the possible forms of expected responses are known to a robot a priori, responses can be modeled and be recog- nized as events. If temporal events are modeled with a high-dimensional cue, some vision-based methods such as gen- erative models (Hidden Markov Models and their (variants) and discriminative models (Conditional Random Fields and their variants) can be adopted Simple events such as touching a robot or making a sound can be modeled using low- dimensional cues, so they are easier to detect using sim-ple filtering methods.

12.2 Change Detection

Change detection measures a human’s behavioral change as a response to a robot’s signal. We look for a significant per- turbation in the human behavior by modeling that behavior before and after the signal and looking for significant differ- ences. We assume that any observed perturbation happening within the allowed time interval is the human’s response. A change is measured using a history of cue data between WB and WA as input. To measure the degree of behavioral difference in the response, we proposed a change detection framework [3], as shown in Fig. 5.
Change will be found out by the contingency detection for each sensory movements of the humanoid using the range finder sensor either by acqustic datas or by vision sensors in such a way that the ideas and the evaluation studies on the humanoid for all the datas will be found out in such a way that the data extracted from the multiple sensory datas is analysed and the same datas is used for contingency detection if anything happens and then the concurrency in the datas for one after one step by step sensory analysis is carried out for different readings and the evaluation studies are incorporated. Then, we calculate the dissimilarity score by measuring a statistical difference be-tween the graph nodes representing data from WB and WA. Finally, as an extension to [3] and as one of the contributions of this paper, we introduce a method that uses probabilistic models to evaluate a dissimilarity score S within our multi-cue contingency detection framework. Chao C, Lee J, Begum M, Thomaz A et al. [7].

12.3 Building the Distance Matrix

We define a cue buffer to be the set of cue feature vectors vt computed at each instant of time. From the cue buffer, we first extract cue data from two time windows WB and WA based on tRef and tC . WB is the time interval between tRef B and tRef , and WA is the time interval between .Building a distance graph from the distance matrix. The edge weight between Vi and Vj , w(Vi , Vj), is the distance between Vi and Vj , DMX (i, j). N (Vi , k) is the kth nearest neighbor of Vi in WB . The four nodes on the left are from V B , whereas the two on the right are from V AtC –A and tC . Let V B and
V A denote cue data extracted from the time windows WB and WA respectively:V B = {vt | t ∈ WB } = {vl+1, vl+2, . . ., vl+P } V A = {vt | t ∈ WA} = {vm+1, vm+2, . . . , vm+Q}.Let V = V B ∪ V A , so |V | = (P + Q). Let Vi denote the ith element of V. The distance matrix DMX of a cue X is calculated by measuring the pairwise distance.

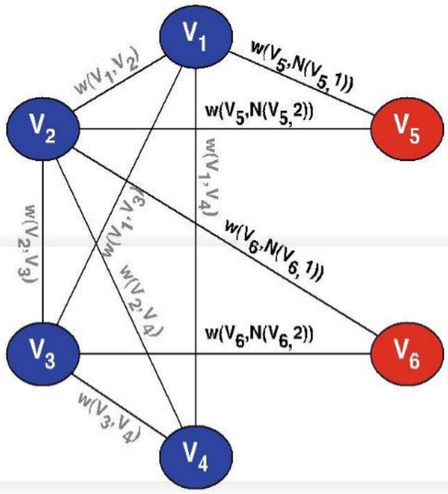


Fig. 4. Building a distance graph from the distance matrix. The edge weight between V_i and V_j , $w(V_i , V_j)$, is the distance between V_i and V_j , $DM_X (i, j)$. $N (V_i , k)$ is the kth nearest neighbor of V_i in W_B . The four nodes on the *left* are from V^B , whereas the two on the *right* are from V^A

t_{C-A} and t_C . Let V^B and V^A denote cue data extracted from the time windows W_B and W_A

respectively:

$V^B = \{v_t \mid t \in W_B \} = \{v_{l+1}, v_{l+2}, . . . , v_{l+P} \}$
$V^A = \{v_t \mid t \in W_A\} = \{v_{m+1}, v_{m+2}, . . . , v_{m+Q}\}$.

Let $V = V^B \cup V^A$, so $|V| = (P + Q)$. Let V_i denote the ith element of V. The distance matrix DM_X of a cue X is calculated by measuring the pairwise distance between cue elements in V ; $DM_X (i, j)$ describes the distance between cue vectors V_i and V_j using a predefined distance metric for the cue X. We will describe distance metrics for visual cues, motion , and body pose **. Kar et al [1]**

12.4 Building the Distance Graph

We construct a distance graph from the distance matrix. The distance graph has the following characteristics as shown in Fig. 6:

1. Nodes from W_B (e.g. V_1 to V_4) are fully connected to each other.
2. Nodes from W_A (e.g. V_5 and V_6) are never connected to each other.
3. Nodes from W_A are only connected to the κ nearest nodes from W_B .

The edge weight between two nodes V_i and V_j , $w(V_i , V_j)$, corresponds to $DM_X (i, j)$.

12.5 Calculating the Dissimilarity Measure

We measure dissimilarity by calculating the ratio of the cross-dissimilarity between V^B and V^A to the self-dissimi-larity of V^B . $N (V_i , k)$ denotes the kth nearest neighbor of V_i in V_B . Let E denote the number of dissimilarity evalua-tions. $CD(V)$ measures the cross-dissimilarity between V^B and V^A in the following equation: $CD(V)$

$$=\sum_{q=1}^{Q}\sum_{k=1}^{k}\sum_{e=1}^{E} w(VP+q, N(N(VP+q,k),e)), \qquad (1)$$

where P is$|V^B|$ and Q is$|V^A|$

SD(V) measures the self-dissimilarity within M^{B}_{T}

SD(V)

$$=\sum_{q=1}^{Q}\sum_{k=1}^{k}\sum_{e=1}^{E} w(N(VP+q,k), N(N(VP+q,k),e)), \qquad (2)$$

The dissimilarity of *V*, *DS(V)*, is defined as

$$DS(V) = \frac{CD(V)}{SD(V)} \qquad (3)$$

DS(V) is the dissimilarity score *S* for a given *V*

$$V^B = \{ v_t | t \in W_B\} = \{v_{l+1}, v_{l+2}, \ldots, v_{l+P}\} \qquad (4)$$

$$V^A = \{v_t | t \in W_A\} = \{v_{m+1}, v_{m+2}, \ldots, v_{m+Q}\} \qquad (5)$$

Let $V = V^B \cup V^A$, so $|V| = (P + Q)$. Let V_i denote the *i*th element of V. The distance matrix DM_X of a cue X is calculated by measuring the pairwise distance between cue elements in V ; $DM_X (i, j)$ describes the distance between cue vectors V_i and V_j using a predefined distance metric for the cue X. We will describe distance metrics for visual cues, motion & body pose Pitsch K, Kuzuoka H, Suzuki Y, Sussenbach L, Luff P, Heath C et al [2]

In [3], we learned a threshold value on the dissimilarity score from the training data and used that to classify the score as being contingent or not. This simple evaluation method can-not be used in our probabilistic model for multi-cue integration because it does not have the confidence on the decision made, and because it does not take into account how informative a used cue is. Triesch J, Teuscher et al [15].

We propose a new evaluation method that resolves the two problems described above. To determine that an ob-served change (i.e. a dissimilarity score) actually resulted from a human response and not from changes that occur naturally in the human's background behavior, we should evaluate the change not only under the contingency condi-tion, but also under the non-contingency condition. To this end, we model two conditional probability distributions: a probability distribution of the Time is shown as.

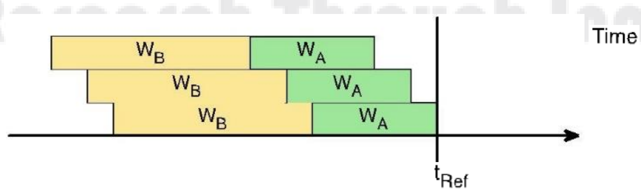


Fig. 5. A method for building null hypothesis. Dissimilarity score sam-ples are obtained by evaluating data in windows over time

The is the dissimilarity score *S* under the contingency condition *C*, *P (S|C)*, and a probability distribution .

We learn the distribution *P (S|C)* off-line from training data in which human subjects are being contingent to the robot's action. We estimate the distribution *P (S|C)*, the null hypothesis, on the fly during an interaction from observa-tions of the human's behavior before the robot triggers a signal. It is important to note that a null hypothesis is es-timated with on-line data, particularly the data gathered is estimated from dissimilarity score samples, each of which is obtained as if the robot's signal were triggered and enough data were accumulated at each point in time. Pitsch K, Kuzuoka H, Suzuki Y, Sussenbach L, Luff P, Heath et al [11] between cue elements in V ; DMX (i, j) describes the distance between cue vectors Vi and Vj using a predefined distance metric for the cue X. We will describe distance metrics for visual cues, motion and body pose .

12.6. Evaluating Dissimilarity Score

In [3], we learned a threshold value on the dissimilarity score from the training data and used that to classify the score as being contingent or not. This simple evaluation method can-not be used in our probabilistic model for multi-cue integration Csibra G, Gergely G et al [17]

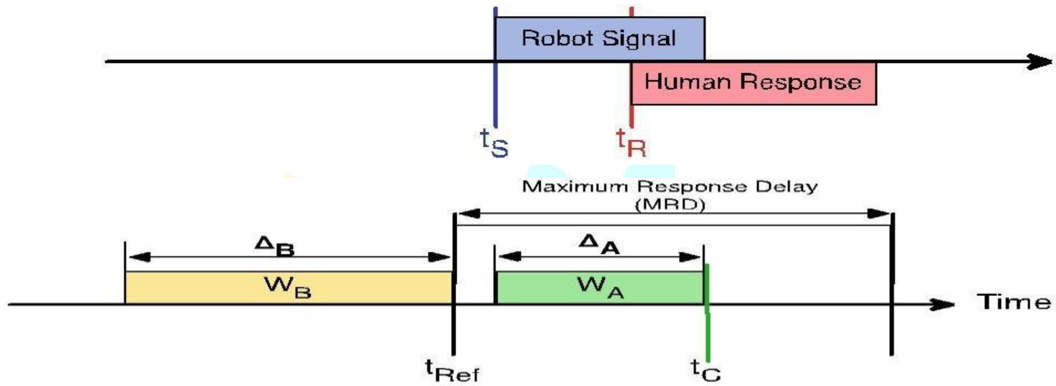


Fig. 6. Causal relationship between a robot's interactive signal and a human's response. *Top*: t_S is the start of the robot signal, and t_R is the start of the human response. *Bottom*: Two time windows, W_B and W_A , defined with respect to the current time t_C , are used to model the human behavior before and after t_{Ref} . W_B and W_A are of size $_B$ and $_A$ respectively. The time window starting at t_{Ref} and valid over *MRD* is examined for the contingent human response. Note that t_{Ref} may not be same as t_S

We propose a new evaluation method that resolves the two problems described above. To determine that an ob-served change (i.e. a dissimilarity score) actually resulted from a human response and not from changes that occur naturally in the human's background behavior, we should evaluate the change not only under the contingency condition, but also under the non-contingency condition. To this end, we model two conditional probability distributions: a probability distribution of the dissimilarity score S under the contingency condition C, P (S|C), and a probability distribution of S under the non- contingency condition, P (S|C). Assuming that a human changes her behavior when respond-ing, these two distributions need to differ for the cue to be considered informative.

We learn the distribution P (S|C) off-line from training data in which human subjects are being contingent to the robot's action. We estimate the distribution P (S|C), the null hypothesis, on the fly during an interaction from observations of the human's behavior before the robot triggers a signal. Lee J, Kiser J, Bobick A, Thomaz A et al [9] .

13. Hypothesis Formulate

Hypothesis is estimated with on-line data, particularly the data gathered immediately before the robot's signal is triggered. As shown in Fig. 7, this distribution is estimated from dissimilarity score samples, each of which is obtained as if the robot's signal were triggered and enough data were accumulated at each point in time. A method for building null hypothesis. Dissimilarity score samples are obtained by evaluating data in windows over time.

13.1 Senory data for Contingency Detection

The choice of cues for response detection should be deter-mined by the nature of the interaction. When a robot en-gages in face-to-face interaction with a human, a shift in the human's eye gaze is often enough to determine the presence of a response [14, 15]. In situations where a gaze cue is less indicative or is less reliable to perceive, other cues should be leveraged for response detection. Here, we are interested in modeling human behavior using three different cues: (1) a motion cue, the pattern of observed motion; (2) a body pose cue, the observed human body configuration; and (3) an au-dio cue, the presence of human sound.

13.2. Grounded response and data recognisition

The motion cue models the motion patterns of the human of interest in the image coordinate system. This cue is derived from observations from image and depth sensors. To observe only motions generated by a human subject in a scene, we segment the region in which a human subject appears. To do so, we use the OpenNI API [21], which provides the functionalities of detecting and tracking multiple humans using depth images. The process for generating the motion cue is illustrated in Fig. 9. . First, the motion in an image is estimated using a dense optical flow calculation [22]. After grouping motion regions using a connected components-based blob detection method, groups with small motions and groups that do not coincide with the location of the human of interest are filtered out. The motion cue comprises the re- maining motion regions. This extracted motion cue is used as an input into the motion contingency module and is accumulated over time. Triesch J, Teuscher C, Deak G, Carlson E et al [15].

enough to fill the past B amount of time and discards frames that are older. When triggered at time tRef , the data for WB stays fixed, and the detector transitions to accumulating cue data to fill WA until time t Ref + A .

$$P(C \mid S_i, S_j) = \frac{P(S_i, S_j \mid C)P(C)}{P(S_i . S_j)} \quad (6)$$

$$= \frac{P(S_i \mid C)P(S_j \mid C)P(C)}{P(S_i, S_j)} \quad (7)$$

$$P(C \mid S_i, \overline{S}_j) = \frac{P(S_i, S_j \mid \overline{C})P(\overline{C})}{P(S_i, S_j)} \quad (8)$$

$$= \frac{P(S_i \mid \overline{C})P(S_j \mid \overline{C})P(\overline{C})}{P(S_i, S_j)} \quad (9)$$

$$\frac{P(\underline{C} \mid S_i, S_j)}{P(C \mid S_i, S_j)} = \frac{P(S_i \mid \underline{C})P(S_j \mid \underline{C})P(\underline{C})}{P(Si \mid C)P(Sj \mid C)P(C)} \quad (10)$$

$$E = \sum_{t=1}^{n} e^2{}_t \Delta t \quad (11)$$

$$S = \sum_{t=1}^{N} \sqrt{(v_t - v_M) \;/(N^2 - 1)} \quad (12)$$

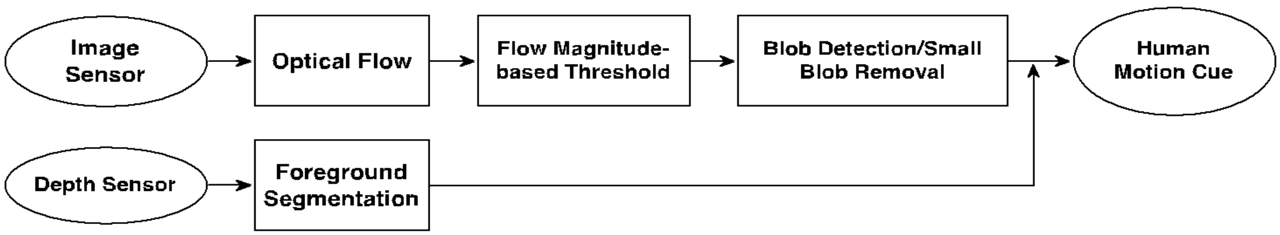
$$\eta(y) := \frac{T}{2\pi} \sqrt{\Delta(y)}, \quad (13)$$

13.3 Data Collection with gait pattern

Two important scenarios in HRI where contingency detection is applicable are engagement detection and turn-taking. In previous work, we demonstrated our contingency detection module with single cues in an engagement detection scenario using Simon, our upper-torso humanoid robot [3]. In that experiment, the robot attempted to get a human subject's attention while the human was performing a back-ground task, such as playing with toys or talking on a cell phone. Butko N, Movellan J (2010) et al [16]. In this paper, we validate multiple-cue contingency detection within a turn-taking scenario in which the robot plays an imitation game with a human partner. Importantly, this Sitation game was designed not for evaluation of a contingency detector but for generating natural interactions with human-like timings.

13.4 Sensory Integration at the Decision Level

At the decision level of integration, we use the naïve Bayes probabilistic model to integrate dissimilarity scores obtained from cues. We chose this model because cues that are integrated at this level are assumed to be conditionally in-dependent of each given the contingency value; otherwise, they should be integrated at other levels. Assume that some function f_X is provided that summarizes the current and past cue observations into a dissimilarity score. For two cues Cue_i and Cue_j that are integrated at the decision level, the overall contingency C in terms of $S_i = f_i\ (Cue_i\ )$ and $S_j = f_j\ (Cue_j\ )$ is estimated by standard Bayesian mechan-ics, as described by the Integrating Motion and Body Pose Cues at the Module Level At the module level of integration, the cue data are integrated when calculating a distance matrix. From the accumulated cue data for both motion and body pose cues, two distance matrices are calculated independently and merged into a new distance matrix. Since the motion and body pose cues are represented in different coordinate spaces, an image space for motion and a 3D world space for body pose, the distance matrices need to be normalized. Crystal Chao Aaron F. Bobick Andrea L. Thomaz et al [4].



of motion cue from image and depth sensors

Fig.7. Extracting cues from different sensors: (a) audio cue, (b) motion cue, and (c) body pose cue has to be evaluated so that the event might occur suppose the event did not occur then the data concurrency is reached if the data concurrency for the event to be happened like the robot falls with the data reaching the mind of the memory is delayed the robot gets an fall to avoid this a contingency factor has to be evaluated. The way the data reaching the humanoid is calculated and then the range detection can be done to the robot in navigation. An implementation of the proposed contingency framework with the Humanoid motion is given with SFA.

14. Application of semi autonomy algorithm and neuronally implementation

In contrast to the IIR filter more acceler-ation sensor values were integrated, namely four sensors from both shoulders and another four sensors located at the robot's feet (overall four sensors directing to the coronal plane and four to the sagittal plane). All 16 sensors could have been used, but in order to keep computational cost low the number of sensors was reduced as long as no deterioration of the result-ing SFA(2) components was observed. Interestingly

Table. 1 Values of quality rates for Simulation

Grounded Resposes (values)	S1	S2	S3
E	7317	7681	7287
S	0.131	0.126	0.125
S	23.8	23.6	23.6
T	47.0	50.3	46.1
W	0.51	0.47	0.51

Table. 2 Sensitivity Analysis Results

	Frequency (Hz)	Amplitude(db/Hz)
No filter	1.76	-10.71
IIR	1.67	-7.11
SFA	1.47	-8.54

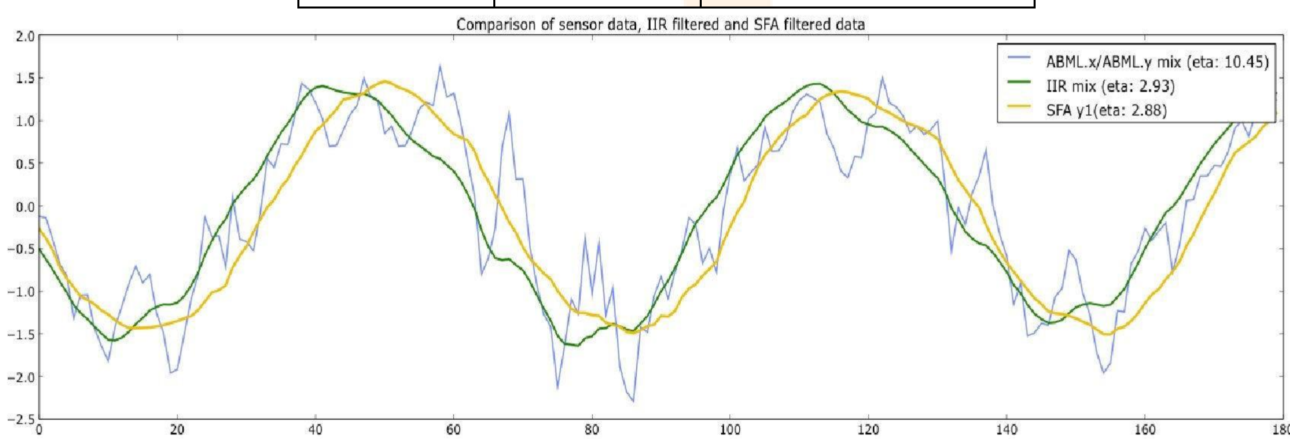


Figure 8. Comparing a weighted sum of the coronal acceleration sensors located at the shoulders to an IIR filtered signal and the slowest component extracted by SFA.

15. Results and Discussions

We conducted several experiments with our Humanoid, using the SFA implementation available from the open source in order to compare the obtained signals the η value proposed in (Wiskott and Sejnowski, 2002) was used:

A dedicated software environment has been created for autonomy method testing and evaluating. It is based on ROBOS/Simulink package. The software responsible for surroundings simulation and virtual sensors indications has been separated from semi autonomy method.

This work presents the simulation research results in which obstacles were placed on the robot's route to the target. The robot preformed the complex „go to the target omitting obstacles" behavior. The mobile robot's semi autonomy algorithm has been evaluated in terms of sensitivity to the failure of selected environmental sensors. In order to receive a conclusive assessment of the results, the following quality rates have been proposed:

the sum of squares of the robot's distance to the target Chao C, Lee J, Begum M, Thomaz A et al [13]
where e_t – the robot's distance to the target, N – the number of iterations till the robot reaches its target or the simulation ends before the target is reached,

where: v_t – robot's speed, v_M – robot's average speed, c) length of the route from the starting position to the target the time T it takes for the robot to reach the target, assuming the target is achieved for $e_t \leq 0.5$ [m], robot's medium speed v_M (within a time from 0 to T).

It should be noted that quality rates (a) – (d) should be minimized whilst (e) should be maximized. In the case the robot cannot reach the target within the assumed time T_{max} = 100 [s], the quality rates s and T reach the value of +∞. Remaining rates reach the values calculated for T_{max}.

All environmental sensors are working properly in Simulation 1The first presented simulation has been performed with all of the robot's sensors working properly. The results of this simulation are presented in Fig. 6-7 and Table 1. They are used as reference for the next simulations, in which influence of selected sensors' failure is presented.

In Fig. 6 and Table 1 motion trajectories of the robot and quality rates for all simulations. Fig. 7 illustrates robot's linear speeds v_R, its medium value v_M (a) as well as robot's distance e and angle $d\gamma$ to the target (b).

The relation (weighting) between robot's sub behaviors remains constant during the robot's movement. It changes only in case of detection environmental sensors' failure. A performed simulation shown that elaborated method allows omitting obstacles and getting to defined destination point. Robot follows a path with variable speed: the speed is decreased, as movement, the distance to the target is decreasing in Simulation.

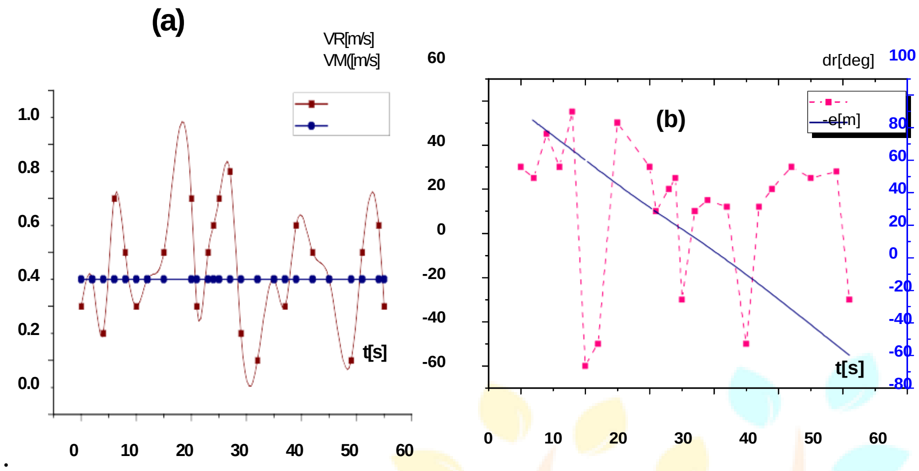


Fig .9. Result of Simulation 1

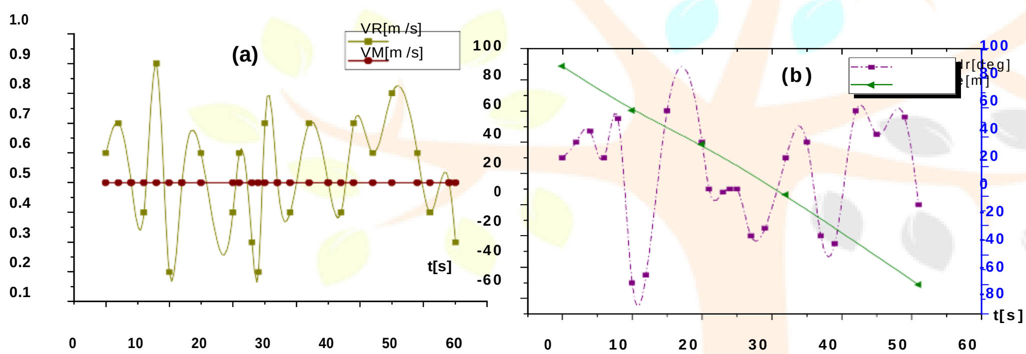


Fig.10. Result of Simulation 2

Although we described in (Spranger et al., 2009) that it is possible to obtain very smooth resulting SFA components by the application of several subsequent SFA steps and without time embedding, this method is inappropriate for this task. The reason is that a cascade of subsequent SFA components.

In Experiment 1 – failure of the laser rangefinder marked as FL In the second experimental test laser rangefinder FL failed. As a result (see Fig. 10, Fig. 12 and Table 2) robot reached the destination point in 101 [s]. All quality rates except of *S* were slightly worst in comparison to the first experiment.

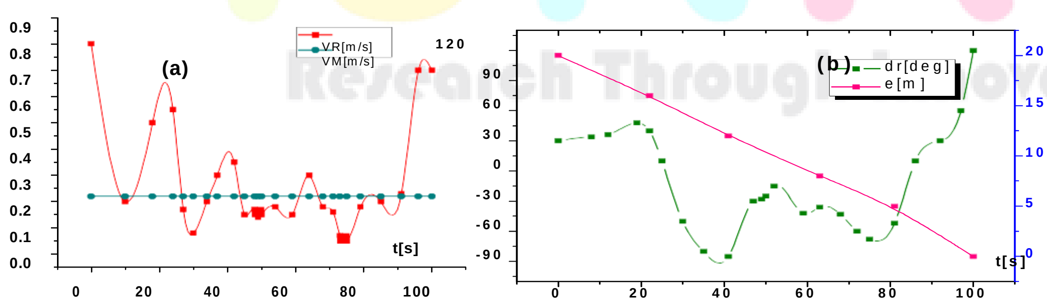


Fig.11. Result of Experiment 1

International Research Journal
Research Through Innovation

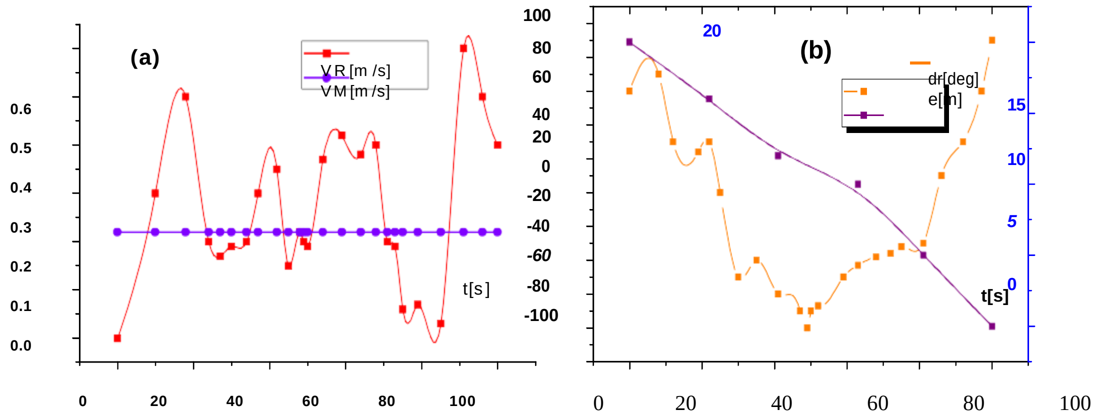


Fig.12. Results of Experiment 2

In Experiment 2 Investigation of the laser scanner

In this experiment the failure of laser scanner occurred. As one can notice (see Fig. 10, Fig. 13 and Table 2) the robot moves to the target by different route and with much less speed in comparison to previous two tests. Therefore quality rates E, T and vM are the worst and S and s are the best from all experiments. Robot moves slowly but with relatively constant speed and reached the target accidentally using shorter route.

Extracted SFA Components

Figure 3 plots data stemming from an extract of an SFA training sequence generated by the unfiltered gait network. The acceleration sensor data mix, the sig-nal obtained by the application of the IIR filter to the acceleration data mix and the slowest component extracted by the SFA module are depicted. All signals were whitened before plotting for better comparability and calculation of η values. The acceleration data mix's η value being at 10.45 is much higher than the values of the IIR and SFA filtered signals ranging both at about 2.9. It is obvious that the resulting slowest component is highly correlated to both the acceleration data mix and to the IIR filtered signal. However, a short delay in the SFA module compared to the other signals issuing from the time delay , In Experiment 2 Investigation of the laser scanner In this experiment the failure of laser scanner occurred. As one can notice (see Fig. 10, Fig. 13 and Table 2) the robot moves to the target by different route and with much less speed in comparison to previous two tests. Therefore quality rates E, T and vM are the worst and S and s are the best from all experiments. Robot moves slowly but with relatively constant speed and reached the target accidentally using shorter route.

Extracted SFA Components

Figure 3 plots data stemming from an extract of an SFA training sequence generated by the unfiltered gait network. The acceleration sensor data mix, the sig-nal obtained by the application of the IIR filter to the acceleration data mix and the slowest component extracted by the SFA module are depicted. All signals were whitened before plotting for better comparability and calculation of η values. The acceleration data mix's η value being at 10.45 is much higher than the values of the IIR and SFA filtered signals ranging both at about
2.9. It is obvious that the resulting slowest component is highly correlated to both the acceleration data mix and to the IIR filtered signal. However, a short delay in the SFA module compared to the other signals issuing from the time delay is observable.

16. Future Work

Within the current framework, we can model and recognize grounded responses as events. As future work, we will investigate how to attribute semantics to ungrounded responses through iterative interactions.

17. Conclusion

In this Research, we could propose a contingency detection framework that integrates data from multiple cues by simulation and thereby implementing in Humanoid experiment.Multimodal sensor data could be gathered from a turn-taking human-robot interaction scenario based on the turn-taking . This is done by imitation game of "Simon says." implement multiple-cue approach and evaluate it using motion and body pose cues from a structured light depth sensor and audio cues from a microphone."The results could be shown as integrating multiple cues at appropriate levels which offers an improvement over individual cues for contingency detection. It could be showed that using MNI as a referent event provides contingency detectors a more reliable evaluation window. From the proposed constrained experiment, it could be an important observation that humanoid response timings and effective integration of cues are important factors to detect contingency. We could believe that this observation could be important to understand other factors situated in more complex and ambiguous interactions. We could believe that our contingency detection.
To engage in multimodal interactions with humans then the semantics of the human's behavior are not known to the humanoid Robot.

To overcome this limitation, we believe that a robot should model and recognize a set of **grounded responses** that are built from knowledge about the nature of the interaction situation and should also be able to ground responses that are found by semantics-free contingency detection.

we have investigated how to attribute semantics to **ungrounded responses** through iterative current framework, we can model and recognize grounded responses as events.

In this Research, we had presented a simple and reliable approach of creating humanoid robot platform based on the Sim Mechanics and Virtual Reality Model Language using Robo OS

This approach could be low cost, easy to operate. Also, it has good scalability.

Therefore, put the robot platform into the different virtual environment, different studies were conducted, such as

a. walking control,

b. arms coordinated operation,

- ✓ multi-robot coordination manipulation.

DECLARATION

The author have no relevant financial or non-financial interests to disclose.
I am a Single author, There are no authors involved in this research.

So therefore the following declarations
The authors have no conflicts of interest to declare that are relevant to the content of this article.

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

The authors have no financial or proprietary interests in any material discussed in this article.
The authors did not receive support from any organization for the submitted work.

No funding was received to assist with the preparation of this manuscript.

No funding was received for conducting this study.

No funds, grants, or other support was received.

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

References

[1]. Linear-time robot localization and pose tracking using matching signatures Kar et al Robotics and Autonomous Systems 60 (2012) 296–308, Elsevier.

[2] .Pitsch K, Kuzuoka H, Suzuki Y, Sussenbach L, Luff P, Heath C (2009) The first five seconds: Contingent stepwise entry into an interaction as a means to secure sustained engagement. In: IEEE international symposium on robot and human interactive communication (ROMAN)

[3]. Feature fusion for 3D hand gesture recognition by learning a shared hidden space. Jun Cheng J. et al. Pattern Recognition Letters 33 (2012) 476–484 , Elsevier . Can Xie a,b, Wei Bian a,b,c, Dacheng Tao c, Shenzhen Institutes of advanced Technology, Chinese Academy of Sciences, China b The Chinese University of Hong Kong c Centre for Quantum Computation and Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology, Sydney Broadway NSW 2007, Australia.

[4]. Multi-cue Contingency Detection,Jinhan Lee · Crystal Chao · Aaron F. Bobick Andrea L. Thomaz. 7 November 2011 Springer Science & Business Media BV 2012,Int J Soc Robot DOI 10.1007/s12369-011-0136-5.

[5].Lohan K, Vollmer A, Fritsch J, Rohlfing K, Wrede B (2009) Which ostensive stimuli can be used for a robot to detect and maintain tutoring situations? In: ACII workshop

[6].PASAR An integrated model of prediction, anticipation, sensation, attention and response for artificial sensorimotor systems Zenon Mathews a,, Sergi Bermúdez i Badia a, Paul F.M.J. Verschure et al ,Information Sciences 186 (2012) 1– 19,elsevier . A SPECS, Technology Department, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain b Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08010 Barcelona , Spain

[7].Yang Yu-biao, He Han-wu. A study of VRML-JAVA Based Robot Motion Simulation. *MECHANICAL SCIENCE AND TECHNOLOGY*; 2004, 23(1), p. 120–126

[8].Inverse Kinematics of Human Arm Based on Multisensor Data Integration. Mihelj (B) et al J Intell Robot Syst (2006) 47: 139–153 Elsevier ,Faculty of Electrical Engineering, University of Ljubljana . 25, SI-1000 Ljubljana, Slovenia, J Intell Robot Syst (2006) 47: 139–153,DOI 10.1007/s10846-006-9079-8,e-mail: J Intell Robot Syst (2006) 47: 139–153 DOI 10.1007/s10846-006-9079-8Stoytchev A (2011) Self-detection in robots: a method based on detecting temporal contingencies. In: Robotica. Cambridge University Press, Cambridge.

[9] . Lee J, Kiser J, Bobick A, Thomaz A (2011) Vision-based contingency detection. In: ACM/IEEE international conference onhuman-robot interaction (HRI).

[10]. Lohan K, Vollmer A, Fritsch J, Rohlfing K, Wrede B (2009) Which ostensive stimuli can be used for a robot to detect and maintain tutoring situations? In: ACII workshop

[11]. Pitsch K, Kuzuoka H, Suzuki Y, Sussenbach L, Luff P, Heath C (2009) The first five seconds: Contingent stepwise entry into an interaction as a means to secure sustained engagement. In: IEEE international symposium on robot and human interactive communication (ROMAN)

[12]. Lee J, Kiser J, Bobick A, Thomaz A (2011) Vision-based contingency detection. In: ACM/IEEE international conference on human-robot interaction (HRI)

[13]. Chao C, Lee J, Begum M, Thomaz A (2011) Simon plays Simon says: The timing of turn-taking in an imitation game. In: IEEE international symposium on robot and human interactive communication (ROMAN)

[14]. Sumioka H, Yoshikawa Y, Asada M (2010) Reproducing interaction contingency toward open-ended development of social actions: Case study on joint attention. In: IEEE transactions on autonomous mental development

[15]. Triesch J, Teuscher C, Deak G, Carlson E (2006) Gaze following: why (not) learn it? In: Developmental science
[16]. Butko N, Movellan J (2010) Infomax control of eye movements. In: IEEE transactions on autonomous mental development
[17]. Csibra G, Gergely G (2006) Social learning and social cognition: The case for pedagogy. In: Processes of changes in brain and cognitive development. attention and performance. Oxford University Press, London[18]. Gold K, Scassellati B (2006) Learning acceptable windows of contingency.In: Connection science
[18] Ashok Kumar R, Marimuthu K (2014) Prediction of stochastic cognitive neural schema with neural network paradigm of latent semantic nodes in autobot—humanoid robot using nonlinear regression of Gaussian sigmoidal curves in Boltzmann normali- sation. International Information Institute, Tokyo, Japan. Inf Int Interdiscip J 17(8):3971–3984
[19] Ashok Kumar Ramadoss , Investigation of Data Concurrency of Sensory fusion in Humanoid Robot, IJAMAE - International jounal of advancement inmechanical and aeronautical engineering,vol 1, pp 45-49, IRED, Institute of Research engineers and Doctors, NEW YORK,USA..ISBN: 978-981-07- 6261- 2doi:10.3850/ 978- 981-07-6261-2_61: ISSN: 2372-4153, 2013
[20] Ashok Kumar Ramadoss & Marimuthu Krishnaswamy , Embedding inference engine in fuzzy expert robotic system shell in a humanoid robot platform for selecting stochastic appropriate fuzzy implications for approximate reasoning, Artif Life Robotics, Springer, SG Germany, vol. 20, no. 1, pp. 13-18.ISSN:14335298,2015 doi.org/10.1007/S10015-014-0189-2 https://link.springer.com/article/10.1007/s10015-014-0189-2
[21] Ashok Kumar Ramadoss & Marimuthu Krishnaswamy, Cognitive Potential impact Minimization by Mitigating through the implementation of correction factor in the TED algorithm for the concurrency control in anticipation error during different movements in the Ramda Humanoid robot, International Journal of Digital content Technologies and its applications, Advanced Institute of Convergence Information Technology Research Center, South Korea, vol. 2, no. 4, pp. 16-21, ISSN :19759339, 2015
[22] Dr.Ashok Kuma Ramadoss,Fuzzifing using Approximation with Fuzzy to Extract Structure and Motion Environments in Dynamic Humanoid Robotic System from the Opted and Uncertainty by Quantization in non monotone neural networks, International Journal of Universal association of computers and electronics engineers IJUACEE,VOL 1, pp 5-9, Institute of Research Engineers and Doctors, Switzerland ISBN No. 978-1-63248-131-3, 2017

About Author

Dr.Ashok Kumar Ramadoss
Associate Professor
*EX-Scientist and Specialist (SIV Industries Ltd.,),
*Project LEADER (Society Generalle' Software Asia French MNC ITPL Bangalore and LG soft Korean MNC)
*Senior Global Member Universal Association of Computers and Electronics Engineers, Sr.MUACEE (New York US- IRED).
*Best Paper Award Winner in Humanoid Robotics at International Conference - KL, Malaysia.
*Appreciation Award received at International conference in Neural Schema for Humanoid Robotics in Computing ,Communication and Information Technology @ ZURICH, SWITZERLAND.
*Invited Speaker on "Recent Cognitive trends and Activities on Humanoid Robots" @ TOKYO, JAPAN
*Editorial Board Member in journal "Technoarete Transactions on Industrial Robotics and Automation Systems"
*International Scientific and Technical Committee Member and Editorial Board Member in Information and Communication Engineering Committee, for International conferences and journals in World Academy of Science, Engineering and Technology, having i10-index:1775, headquarters at London UK.
*International Advisory Board Member Scope Database.
*Senior member in the HONG KONG Society of Robotics and Automation (HKSRA)
*Member in SRS Robotics surgery Illinois, USA and Orlando Florida
* Best paper award winner at Bangkok Thailand, for Robotics conference.
* Honor to be awarded as member in the Annual meeting held at Orlando, Florida in Society of Robotics Surgery, and selected to attend and give speech at France. Strasburg.