A Robotic Interface for Retrieval of Distributed Multimedia Content

Ellen Lau
Interactions Lab
Dept. of Computer Science
University of Calgary, Canada
laue@cpsc.ucalgary.ca

Ehud Sharlin
Interactions Lab
Dept. of Computer Science
University of Calgary, Canada
ehud@cpsc.ucalgary.ca

Ariel Shamir
The Interdisciplinary Center
Efi Arazi School of Computer
Science, Israel
arik@idc.ac.il

Abstract

In a social setting, it is often beneficial to have a partner that is alert of the environment and the circumstances and is able to communicate information that otherwise does not come to our awareness. Typically, these are human companions, however a robotic partner may be an alternate solution. For a robotic partner to be successful, it must naturally interact with its user, and be able to retrieve data and analyze its content. In this paper we present Iolaus, a robotic partner for social settings. Iolaus can be implemented as a robotic parrot perched on the user's shoulder, or as an attentive companion similar to a seeing eye dog. When the user enters a new social setting the robot can be asked and should attempt to answer questions regarding people and surroundings, very similar to a human partner. In this phase of the project, we are using the Sony Aibo robot dog to implement Iolaus's information retrieval component demonstrating how a human-robot interface can help the user access a vast amount of distributed information and multimedia content in a socially acceptable manner.

Keywords: human-robot interaction, distributed information retrieval, multimedia human computer interaction.

1. Introduction

The growth of mobile computing devices sees with it a growth in distributed computing. Gone are the days when mobile appliances were limited to performing simple tasks. Today, mobile devices are capable of performing complex functions, taking command of tasks that were previously restricted to desktop computers or other non-mobile systems. Such tasks may include performing searches on the Internet, image acquisition and image analysis and comparison. In addition, mobile devices have physically expanded beyond the boxed design of personal digital assistants. Mobile devices currently include a variety of cell phones, entertainment centers and game consoles as well as wireless robots.

Robotic devices reveals a wealth of new utilities and new possibilities for human computer interaction [1,2]. A robot can be viewed as a physical computer interface that can act in both the physical and virtual worlds. A robot can contribute to a task a set of physical abilities as well as digital capabilities, and distributed computing functions. A robot is capable of pairing audio and visual components of interaction, with physical actions and movements to further enhance the interaction experience.

People often find themselves in situations where they need to acquire specific information about new social or physical surroundings. For example, when entering a new social setting people sometimes need to approach a person whom they never met, or meet a person whom they previously met but fail to recall specific information regarding that person (for example, the person's name). Such shortcomings are sometimes relieved using a partner, spouse or a friend who can retrieve and deliver the required information in a manner that is sociably acceptable. Reliance on information appliances such as PDAs for these situations might not be socially acceptable (for example, trying to search for a person's name or image on a PDA while facing them in the middle of a social event). Further, it is not always possible to have a knowledgeable human companion at every social setting; it might however be possible in the near future to be escorted by a personal robot companion. This is assuming the robot can support socially acceptable interaction and can retrieve information dynamically and efficiently for the user.

This paper describes *Iolaus* (pronounced EE-oh-lus), which is a persistent robot partner that can convey important messages back and help the user navigate through new environments and unfamiliar social settings. In Greek mythology *Iolaus* played crucial role in Hercules' labours and helped him slay the Hyrda [3]. We use *Iolaus*, the obliging partner, as a metaphor motivating our human-robot interaction (HRI) approach to distributed information retrieval.

The *Iolaus* project goal is to develop a robot that serves as a socially acceptable information-

Figure 1. *Iolaus*-a conceptual drawing

supporting partner. The user and the robot are to interactively communicate with each other, with the robot able to sustain sufficient knowledge and awareness of the environment in which it acts. *Iolaus* is designed as a companion and acceptable partner that will behave in a socially unobtrusive manner. *Iolaus* is designed to provide assistance and explore the distributed information realm as naturally as a seeing eye dog negotiates a street corner.

2. Designing Iolaus

Our overall design task was to develop a robot that is capable of serving as an autonomous distributed information retrieval device which can act in the physical world. Much like a human partner, the robot is able to take in requests from the user, search for the appropriate information, and formulate a response back to the user. The robot can also act in the physical world based on the information it retrieved, for example by pointing to the direction of a person the user is seeking. To allow users to access information quickly and inconspicuously the robot must be designed as a socially acceptable interface.

The information requested from *Iolaus* could vary over a wide range of topics from asking directions to a point of interest, to finding the name of an associate's spouse. Generally, such information can be retrieved either by talking with knowledgeable individuals or by performing a search on the Internet. In a social setting, information should be requested and received in a socially acceptable and non-intrusive manner. Interaction with typical information appliances interfaces or wearable computers can often lead to behaviours that are perceived to be unsociable and often unacceptable.

2.1 Goal

Our near term design goal is to be able to ask *Iolaus* to locate a person in a room in a socially acceptable manner. The person *Iolaus* is asked to seek can either be physically in the room or be represented by a picture. The robot must then decipher the request in such a way that it can search for the information requested. The search for information will include an autonomous gathering of virtual data from a distributed information source (practically by performing a search on the Internet) and gathering of physical data from the robot's visual, audio and tactile sensors. The data gleaned from the robot's sensors is used to compare the images retrieved to those around the room. The robot should compare the visual images in the room to the image it found on the Internet and decide if a match is made. Lastly, the robot will formulate a response back to the user of its findings (for example, by pointing to the right person or walking towards them)

The robot acts as a physical mobile distributed computing device capable of retrieving information about its surroundings. In a sense, the robot is a distributed search engine and email client, capable of deciphering requests, and then formulating a response back to the user.

Our vision for such a distributed information retrieval robotic device would be similar to a robotic parrot that is capable of sitting balanced on the user's shoulder as seen in Figure 1. The close vicinity of the parrot to the user's ear makes it ideal for voice interaction from the user to the robot and vice versa. This is advantageous since both the user and *Iolaus* can communicate discreetly between each other by simply having the user turn their head toward *Iolaus*. The embodiment of *Iolaus* in a parrot shape can allow the robot to interact with the user in various forms: for example, the parrot can physically move its head or wings to show the user where the person is located. The physical movement can also be paired with an audio component where *Iolaus* verbally respond back to the user with some comments or instructions. From a robotic design point of view the physical requirements from the parrot are quite simple as all it needs to do is maintain balance on a shoulder (that is, the robot is not required to walk, navigate corridors, traverse stairways, etc.).

Another possible metaphor for *Iolaus* is an attentive information-seeking dog. This embodiment has the advantages of mobile abilities in the environment on top of the distributed information capabilities. These can be used for instance to help fetch or place something in the physical world. The physical appearance of a dog would also make it more suitable to deal with navigational requests as the robotic dog could in part respond by guiding the user to the desired location, very much like a real seeing-eye dog.

2.2 Query Request

Iolaus must be made aware of a query request. This can be done by issuing voice commands to the robot, in a similar manner to verbal queries between two people. The verbal interaction between the robot and the user then allows for a high level and more natural means of interaction.

2.3 Information Retrieval

Iolaus should be capable of autonomously retrieving information. It should not be dependent on any computational source other than what is contained within itself. Therefore, the robot is not tethered to any devices and is capable of being completely mobile. The task of information retrieval then is delegated completely to the robot. First, *Iolaus* must perform a search to find the correct answer for the user query. In the task of attempting to correctly identify a person in a room, the robot must be able to find the correct image of the person. Such an image can be obtained from the Internet through various web engine searches and the robot should be capable of going online, performing the search and obtaining the results from the search engine. Second, the robot must gather information from its surroundings through its visual sensors. Once it has successfully obtained an image from the Internet, the robot must look around the room and match the retrieved image with one that it sees through its vision system.

2.4 Formulating a Response

Finally, after the information retrieval phase, *Iolaus* should formulate its response into an appropriate form and deliver it to the user. This can be done in a number of different ways depending on the situation and application. For example, the robot can send navigational directions back to the user or physically guide the user to the location of the desired person using visual, verbal or physical interaction techniques.

2.5 Physical Interaction

In order for *Iolaus* to be an acceptable and useful human-robot interface it should be based on intuitive interaction techniques. The robot physical state must afford and reflect its current virtual status. For example, if the robot is busy searching for an image on the Internet, it should physically demonstrate that it is busy with a task and cannot be bothered. When *Iolaus* is waiting for the next user request it should physically reflect that it is willing to serve a new query. The robotic interface should not be restricted to visual and audio stimulation but rather enhance the experience by controlling and using the robot's physical state as a valid output and display tool.

3. Implementation

To formulate a robot capable of fulfilling the design tasks outlined, we used the Sony Aibo ERS-7M2 for *Iolaus's* prototype. The Aibo is an autonomous robotic dog that has a face plate LED panel, a 350K pixel image sensor, a speaker capable of producing MIDI sound, stereo microphone, head/paw/body sensors as well as IEEE 802.11b Wireless LAN capabilities.

Our Aibo-based *Iolaus* was implemented using OPEN-R—a C++ SDE for the Aibo robot dog available for free from Sony [4]. Aibo's OPEN-R supports coding most of the robot-dog functions as well as its IPv4 Internet protocol and TCP (Transmission Control Protocol). We used these capabilities to program *Iolaus* to retrieve emails and perform searches on the Internet.

Since we are looking for multimedia content, we decided that our Aibo-based *Iolaus* prototype will use a search engine that returns a good array of pictures such as www.picsearch.com [5].

3.1 Iolaus Prototype

The current early prototype of *Iolaus* is capable of performing a simplified but fully functional experiment which demonstrates the potential and usability of the new robotic interface as well as its functionalities in a distributed setting. Our current implementation allows the Aibo to autonomously search and find an image source on the Internet based on a user query. The requests from the user are sent via email to Aibo's private email account (Figure 2). Our Aibo-based *Iolaus* checks its emails regularly for new requests from users. Based on the user's request for information, Aibo will parse through the email, create a search engine query, connect to the Internet and perform a search on a public search engine to find an appropriate picture answering the user's query. After it has successfully extracted the results, Aibo will then package its response into an email and reply back to the user with the link to the online picture. Through this cycle, Aibo will perform various physical/visual movements that allow the user to be aware of its current status.

3.2 Distributed Information Retrieval

The act of connecting to the Internet and searching for an image is very similar to the

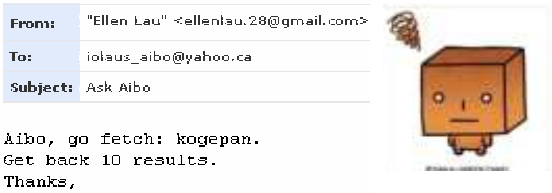

Figure 2. An email request sent to Aibo (left), and the image retrieved back (right).

distributed information retrieval task as outlined in the design plan.

Our Aibo-based *Iolaus* is capable of receiving and sending emails and so contains a fully functional independent email client. Using any email account with POP/SMTP capabilities, Aibo is able to consistently monitor and retrieve emails it receives. The Aibo-based *Iolaus* is also able to delete email messages that have been processed, so once Aibo has tried to search for a user query it will remove that email message from its inbox. This is one way that Aibo can manage its email messages and ensure that its email account doesn't overfill. Also, if the Aibo-based *Iolaus* sees that the subject heading on the email is not something it recognizes, it will automatically remove the message from the queue.

3.3 Robotic Multimedia Human Computer Interface

The key to *Iolaus* is to provide the user with a multimedia robotic interface that is natural to interact with in a social setting. This was achieved in the current prototype by having Aibo physically demonstrate its status. Our Aibo-based *Iolaus* assumes different physical doggish postures and movements reflecting its progress and virtual state (Figures 3,4).

4. Evaluation

To evaluate the effectiveness of the prototype in the original design task we present several complete request/response cycles that were tested using our current Aibo-based *Iolaus* prototype.

The requests made to Aibo are in the form of an email message sent to an email account created specifically for Aibo (*Iolaus* registered to a Yahoo email account). Currently, Aibo will only accept email messages in a specific format. The email message itself must contain the phrase 'Aibo, go fetch:', followed by the name of the item the user is looking for. A typical request may be in the form of: "Aibo, go fetch: bananas.". The message must conclude with an acknowledgement of Aibo's efforts by including, "Thanks, " at the bottom of the message. Figure 2 shows a request made from a user email to our Aibo-based *Iolaus*.

Autonomously parsing through the email, Aibo will isolate out the item to search for and formulate a well formed HTTP request [6]. A typical request may be: GET / search.cgi?q=kogepan HTTP/1.0\r\n. In this request, Aibo has isolated the item "kogepan" from the emailed request and inserted into the proper HTTP request form. The searches for images are done through the image search engine on Picsearch.com. The letters after the GET command in the HTTP request simply tell Picsearch that an image search is going to take place.

After the HTTP request is sent and a response is returned by the server, Aibo will proceed to parse through the web page looking for the first image link on the result page. The link will be the first answer that Aibo provides back to the user. In our example, the image link that Aibo has found (Figure 2) is:
http://images.picsearch.com/is?5261000945372.

Finally, Aibo creates a return email back to the user indicating that he has found an image link to the user query. The message also contains the link so that the user can access the image.

More examples of queries and resulting images are presented in Table 1. As seen, there are scenarios where the images retrieved can be informative, and in other cases there is a need for further processing of the results. For instance, in the case involving the 'fetch apple' request, it could very well be that the user was looking for the fruit, not a computer system. If the user is not satisfied with *Iolaus's* response the query can be easily repeated and extended.

The Aibo-based *Iolaus* physical posture and actions provide the user with insight into the interface state. When Aibo is about to search on the Internet, it performs a 'stretch' by flexing its legs out as shown in Figure 3. This appears as if the robot dog is physically getting ready to perform a complicated task.

Figure 3. Aibo 'stretching', indicating its readiness to search the Internet

Figure 4. Aibo's head turning waiting for a new query.

Once a response email is sent back to the user, the dog sits very much like a watch dog. Our Aibo-based *Iolaus's* head sways from side to side (Figure 4) as if looking for an interesting game to play while it searches and waits for user's next emailed request. As an added method of physically conveying its virtual state, Aibo also flashes its LED lights on its head, back and face when it is in the process of sending and retrieving emails. The lights will turn off when Aibo is awaiting a user request.

5. Related Work

Sociable robots need to perceive, recognize and interpret the behaviour of humans through multiple modalities including vision, hearing and touch. [1,2]. *Iolaus* must allow the user to interact with it in a natural manner in a social setting. A social robot should be able to naturally interact with humans and participate in human society in ways that will be perceived as natural and socially acceptable [1,7].

Robotic emotions can play a key role in the perceived quality of a human-robot interface [8]. The robot can use emotions as tools to understand and convey implicit messages to and from the user. In addition, *Iolaus* can use of synthetic emotions to elicit specific responses from the user. Emotions are crucial in formatting user impressions for reactions in different circumstances.

At first glance, the robot must possess qualities that make it appealing and socially acceptable and its appearance must match its function. A robot's appearance should provide clues to its purpose much the same as an interface should afford its behaviour and function [9]. It was demonstrated that users systematically preferred robots with human likeness features that match the sociability levels required in specific tasks [1,10]. Following, it is then easy to conceive *Iolaus* as a parrot perched on the user's shoulder or a robot puppy that guides users to a desired location similar to a seeing eye dog.

Iolaus needs to provide meaningful information back to the user. Since an autonomous robot has limited resources onboard, it will need to search for external information. This source could be a specialized database or the Internet. For multimedia content, very similar to textual queries, more advanced analysis should be performed on the query results to gain meaningful information. Previous work on image clustering [11,12] and image retrieval [13] already enable more advanced analysis of image results, and relevance feedback techniques which are suitable to human-robot interaction can further enhance its significance.

Table 1. 'Go Fetch' requests to the Aibo, and the images returned

Go Fetch:	Response back
Banana	
Aibo	
Beckham	
Apple	

6. Conclusion and Future Work

People often need to go to new places, meet new people and perform various social functions in new social settings. Often they do not have a companion that can accompany them on these new tasks. We design *Iolaus* to serve as a robot companion that can perform the function of a knowledgeable human partner without the social awkwardness associated with using various information appliances. *Iolaus* can fill requests from the user

regarding their surroundings and interact in such a way that conforms to social norms. *Iolaus* would ideally be placed on the user's shoulder, very much like a parrot but can also assume other forms such as an attentive dog.

We have presented an early prototype of the project, where *Iolaus* is simulated using the Sony Aibo robot dog. The Aibo's task is very similar to what *Iolaus* is required to do. The user poses a request query for Aibo to find an image of an object. The query is directed in the form of an email sent to Aibo's email account. *Iolaus* checks its emails, parses new ones, and forms the request item into an HTTP request. This is done so that Aibo can autonomously take the request and wirelessly access the Internet via a search engine to retrieve a set of possible results. Once the results are returned, The Aibo-based *Iolaus* determines the appropriate answer by isolating an image link on the page and sends a response email back to the user.

We are currently working on ways to expand the human robot interaction with our implementation of *Iolaus*. Some of the tasks include:

- Developing an aural interface with Aibo so that requests and responses can be made via voice commands. Ultimately, we would like *Iolaus* to fully support voice interaction as well as the ability to communicate and understand basic emotions.
- Adding a physical interaction technique that will allow the user to cycle between multiple query results. For example, use the sensors on Aibo's paws and back to browse through the next image in the result set.
- Ultimately, *Iolaus* should be able to interact via physical means. This could mean gesturing of appendages or locomotion to a desired location. Whatever the means of interface *Iolaus* offers, it should be an amalgam of various multimedia outlets that best portray its response as natural for its social setting.
- Preprocessing the multimedia results in a more informative manner by clustering or analyzing the images. This will later enable for instance searching similar objects in the environment using the Aibo's visual system.

While the experimentation using the Aibo is very much a simplified form of *Iolaus's* planned functionality, there are important components that the Aibo prototype can help accomplish such as the wireless and dynamic retrieval of multimedia information which is accessed remotely. The wireless of these components will help advance the goal of realizing the completion of *Iolaus*.

Acknowledgments

This work was partially funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) and a University of Calgary startup grant.

References

[1] C. Breazeal, *Designing Sociable Robots,* The MIT Press. 2002.

[2] C. Breazeal, A. Brooks, J. Gray, M. Hancher, J. McBean, D. Stiehl, J. Strickon, "Interactive Theatre" in *Communications of the ACM,* July 2003, Vol 46, No. 7, http://robotic.media.mit.edu/Papers/Breazeal-etal-CACM03.pdf

[3] C. House, Hercules the hero: understanding the myth, 1998 . http://www.yale.edu/ynhti/curriculum/units/1998/2/98.02.06.x.html

[4] SONY AIBO Software Development Environment, http://openr.aibo.com/openr/eng/index.php4

[5] Picsearch, http://www.picsearch.com

[6] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee "Hypertext Transfer Protocol -- HTTP/1.1" *The Internet Society.* June 1999. http://www.w3.org/Protocols/rfc2616/rfc2616.html

[7] T. Kanda, H. Ishiguro, T. Ishida, "Psychological Analysis on Human-Robot Interaction," in *Proceedings. of IEEE Int. Conf. On Robotics and Automation*, 2001, http://www.lab7.kuis.kyoto-u.ac.jp/services/pub/04.pdf.

[8] D. A. Norman, *Emotional design: Why we love (or hate) everyday things.* Basic Books, New York, 2004.

[9] D.A. Norman, *The Design of Everyday Things.* Basic Books, New York, 1988.

[10] J. Goetz, S. Kiesler, A. Powers, "Matching Robot Appearance and Behavior to Tasks to Improve Human Robot Cooperation," in *University of California, Berkeley, Carnegie Mellon University.* 2003. http://www.contrib.andrew.cmu.edu/~rgockley/papers/AppearanceAndBehavior.pdf

[11] S. Krishnamachari and M. Abdel-Mottaleb, Hierarchical clustering algorithm for fast image retrieval. *In SPIE Conference on Storage and Retrieval for Image and Video databases VII*, pp. 427-435, San-Jose, CA, Jan 1999.

[12] Kobus Barnard, Pinar Duygulu, and David Forsyth, Clustering Art, *In Computer Vision and Pattern Recognition,* pp. II:434-439, 2001.

[13] Y. Rui, T. S. Huang, and S.-F. Chang, Image retrieval: Current techniques, promising directions, and open issues *Journal of Visual Communication and Image Representation,* 10, 39-62. 1999.