iRov: A Social & Interactive Robot Platform Using iOS Technology
Abdul Rahman Hafiz, Masaki Kimura, and Kazuyuki Murase
University of Fukui,
Graduate School of Engineering,
Department of Human and Artificial Intelligence System
{abdul, kimura-m, murase}@synapse.his.u-fukui.ac.jp

Abstract: This paper introduces an autonomous camera-equipped robot platform for active vision research and as an education tool. Due to recent progress in electronics and computing power, in control and agent technology, and in computer vision and machine learning, the realization of an autonomous robots platform capable of solving high-level deliberate tasks in natural environments can be achieved. We used iPhone 4 technologies with Lego NXT to build a mobile robot called the iRov. iRov is a desktop size robot that can perform image processing onboard utilizing the A4 chip which is a System-on-a-Chip (SoC) in the iPhone 4. With the CPU and the GPU processors working in parallel, we demonstrate real-time filters and 3D object recognition. Using this platform, the processing speed was 10 times faster than using the CPU alone.

1. Introduction
There is no general platform for developing embedded systems such as autonomous camera-equipped robot systems. The existing platforms are mostly based on pre-specified models, which are difficult to obtain for various applications due to complexity and imponderables of the environmental world. For designing and developing a general platform for autonomous camera-equipped robot systems we propose for the first time iRov robot, a platform that is based on iPhone 4 technologies with Lego NXT. This robot platform is excellent for broad range of active vision related researches, behavior-based robots, image and signal processing. More than that, it can be used as an educational platform, due to its low price and ease of use.

In this paper we introduce the design of iRov robot, and show how this robot can perform image filtering and recognize 3D objects in real-time. The next section reviews some of the existing robot platform that can be used in active vision. Section three explains the concepts of designing a mobile robot for active vision research and introduce iRov robot. Section four introduces an experiment using iRov robot. Section five shows the experimental results. Section Six discusses the performance and the significance of the platform, and finally section seven concludes this paper.

2. Background
Due to the advancement in digital camera technology, cameras have been used as sensors in many robots platforms, but it has many limitations, such as the size of the frames, the frames rate, in addition to the limitation of the parallel computation capability. These are some examples that available now a day.

2.1 Desktop Size Robots
The E-puck [1] is a desktop size robot and it is widely used for research purposes, especially in behavior-based robots. It is equipped with a camera with a resolution of 640 x 480 pixels. A simple processor like the dsPIC on the robot cannot process the full flow of information this camera generates. Moreover the processor has 8K of RAM, not sufficient to even store one single image. To be able to acquire the camera information, the frame rate has to be reduced as well as the resolution. Typically we can acquire a 40 x 40 sub-sampled image at 4 frames per second. This size of image is enough to study and realize insect like vision such as optical flow, but it is not useful in studying high-level visual functions that exist in mammalian vision systems.

2.2 Laptop Based Robots
Another type of robots is bigger in size and a laptop computer operates it. The base Pioneer 3-DX platform arrives fully assembled with motors with 19 cm wheels [2]. Adding a laptop equipped with camera this robot can be used as platform for active vision research. This type of robots does not have the flexibility to change according to the vision tasks. More over, because of its large size it is difficult to operate in the desk-side environment.

2.3 Robot Design for Active Vision Research
To design a robot with active vision, that can realize mammalian like visual behaviors, and help in studying different high cognitive and vision function, the robot has to be equipped with high resolution camera that can operate in
real-time and with high frame rate. In addition to that, if the image filters are performed by parallel microprocessors such as GPUs, the robot would have the ability to perform in real-time, and change its attention and prospective of vision to interact with the users and the objects. Moreover, the robot should have the flexibility to be configured for every different experiment.

3. iRov Mobile Robot
To reach our goal we decided to use iPhone 4 cameras and utilize the A4 chip for the vision research. Furthermore, iPhone 4 and new generation iPods are quipped with multi-touch display, speaker, external stereo microphones, accelerometer and gyroscope. These allow the robot to be able to interact with humans and objects in a natural manner. Moreover, We built the robot body with Lego NXT parts. It made the robot flexible to be configured differently for different tasks, taking the advantage of the 3 connected servos.

In the design phase, we took into consideration of the robot’s ability to change its prospective to the object and be able to manipulate it. To achieve that, iRov consists of two parts the head and the Body. The head can rotate 180° so the foveal vision (see Sect. 3.2 for detail) can face up and down, to the object or to the user, while the body helps the robot to move to change the robot prospective and to manipulate the objects. Fig. 1 shows the components of iRov robot.

Fig.1 Drawing of different parts of iRov robot

3.1 Characteristics of iRov
Because the iRov robot is small in size, it can perform in lab environment. Fig. 2 shows the setup consists of the user, the robot, and the stationary server, which establish the connection between the iPhone and the Lego NXT. This will be used for high computational load like object recognition.

3.2 Robot Cameras
Mounted in the top of the robot is a panorama camera (the view through the camera is shown in Fig. 4). This camera helps the robot localize itself in the environment and avoid obstacles.
The range of the back-facing camera in iPhone 4 is 30°. In order to have wider view of the periphery, while maintaining the sharpness of the center, we used a special lens as shown in Fig. 5. It generates this representation that helps the robot to be aware of any movement in the periphery (120°). At the same time it allows to examine the detail of the object in the center (fovea).

4. iRov Vision and 3D Object Recognition

By dividing the robot vision to foveal and periphery vision, we could implement an attention system similar to our previous work [3]. Since iRov has a high-speed image processing ability, iRov can recognize a simple 3D object and its orientation in real-time. To do that we implemented Hierarchical Chamfer Matching which is a parametric edge-matching algorithm (HCMA) [4]. By utilizing the capability of parallel processing of the A4 chip we could achieve real-time image filtering.

Before applying HCMA we need to generate edges of the 3D model templates and the real edges of the object. First the model templates are programmatically generated (Fig. 6). Second the image in the robot fovea has to pass through filters such as Gaussian Smoothing and Edge Detection as Fig. 7 (b). Third Build distance image pyramid (Fig. 7 (c))[5]. Finally by applying HCMA [6], we can found the local minimum of the best match between the template and the detected edges (Fig. 7 (d)).

5. Experimental Results

The default value for introduced parameter \( \tau \), threshold for selection of start points, is set to 0.7. Other parameters using HCMA were determined by suggestions from [6]: The factor \( \lambda \) is set to 2; the intervals in x-coordinate and y-coordinate, \( u_x \) and \( u_y \), are both set to 3, the radius \( r \) to define the neighborhood is 1; and the maximum level of distance pyramid \( L \) is set to 5, which means that image on the top of pyramid is 1/16 rescaled from the original image.

Table 1 is the result of the recognition rate for different prospective of the object from 0° to 90°.
Table 1 Result for different prospective of the object

<table>
<thead>
<tr>
<th>Prospective</th>
<th>Correct</th>
<th>Incorrect</th>
<th>Undetected</th>
</tr>
</thead>
<tbody>
<tr>
<td>0°</td>
<td>20%</td>
<td>9%</td>
<td>71%</td>
</tr>
<tr>
<td>30°</td>
<td>61%</td>
<td>11%</td>
<td>28%</td>
</tr>
<tr>
<td>60°</td>
<td>73%</td>
<td>15%</td>
<td>12%</td>
</tr>
<tr>
<td>90°</td>
<td>77%</td>
<td>4%</td>
<td>19%</td>
</tr>
<tr>
<td>Total</td>
<td>57.75%</td>
<td>9.75%</td>
<td>32.5%</td>
</tr>
</tbody>
</table>

Fig. 8 shows that by dividing the process to primary filters that perform in the robot and the secondary recognition stage that perform in the server we can achieve faster performance comparing with the work [8]. While the value of $\tau$ is increased, more candidate points are eliminated at the early stage.

Fig. 8 The relationship between threshold and computational time in both our work and the work done in [6]

6. Discussions

From Table 2, the 0° prospective of the object results in a low recognition rate. To overcome that the robot have to change it prospective to the object so that it can have multiple view points, by rotating its head vertically or its body horizontally.

As we can notice, the performance in this platform relies on the high-resolution cameras (up to 640 x 480 pixels) that can be processed for 60 fps. In addition to that, the accessibility to the hardware such as the CPU and GPU and other sensors like motion and rotation sensors, is much simpler than other platform thanks to iPhone SDK.

The significance of this Robot platform fills in two main fields: First, as research tool for active vision robots, it takes the advantage of the mentioned performances and the flexibility of the robot to be configured for different experiments. Second, as education tool, because of the publicity of both the Lego NXT and iPhone in the education environment, it provides the excellent programming experience to the students through this robot platform.

7. Conclusion

iRov robot is a platform for developing autonomous camera-equipped robot. Which can help to produce flexible systems at high degrees of sophistication. This system helps developing the future robots. More than that, it helps in studying the human vision and cognition by synthetic approach. We demonstrated that by performing simple 3D object recognition, and by dividing the process between the CPU and the GPU on the iPhone 4 and the server computer, we achieved real-time performance.

Now we are working in robot attention system and human robot interaction, which will demonstrate more of the iRov robot capabilities.

8. Acknowledgement

This work was supported by grants to KM from Japanese Society for Promotion of Sciences and from the University of Fukui.

References


