Research Article

A Novel Bioinspired Vision System: A Step toward Real-Time Human-Robot Interactions

Abdul Rahman Hafiz,[1] Fady Alnajjar,[2] and Kazuyuki Murase[1, 3]

[1] *Department of System Design Engineering, Graduate School of Engineering, University of Fukui, Fukui 910-8507, Japan*
[2] *Embodiment and Consciousness Unit, Brain Science Institute, BTCC RIKEN, Nagoya 463-0003, Japan*
[3] *Research and Education Program for Life Science, University of Fukui, Fukui 910-8507, Japan*

Correspondence should be addressed to Kazuyuki Murase, murase@synapse.his.u-fukui.ac.jp

Building a human-like robot that could be involved in our daily lives is a dream of many scientists. Achieving a sophisticated robot's vision system, which can enhance the robot's real-time interaction ability with the human, is one of the main keys toward realizing such an autonomous robot. In this work, we are suggesting a bioinspired vision system that helps to develop an advanced human-robot interaction in an autonomous humanoid robot. First, we enhance the robot's vision accuracy online by applying a novel dynamic edge detection algorithm abstracted from the rules that the horizontal cells play in the mammalian retina. Second, in order to support the first algorithm, we improve the robot's tracking ability by designing a variant photoreceptors distribution corresponding to what exists in the human vision system. The experimental results verified the validity of the model. The robot could have a clear vision in real time and build a mental map that assisted it to be aware of the frontal users and to develop a positive interaction with them.

1. Introduction

Building a human-like robot controller that is inspired by the principles of neuroscience and can resemble living organism behaviors in some specific characteristic is recently one of the main challenges faced by robotic researchers [1]. The difficulty in such a system can be summed up into three main points as diagrammatically shown in Figure 1. (A) A mechanism for human-robot interaction, which mainly relies on robot's vision, speech recognition, sensor-motor interaction, and so forth. (B) A mechanism for learning and memory, which gives the robot the feature to learn and/or teach. (C) A mechanism for homeostasis, which gives the robot a degree of an internal stability. In this study, we are highlighting the issue of enhancing the robot vision toward advanced human-robot interaction. More specifically, we are introducing a novel dynamic edge detection algorithm that is inspired from the biological concept of "retina" and supported by the variant photoreceptor distributions and the robot's eye movements.

Edge detection is classified as a fundamental step in many machine vision systems and image processing applications [2–4]. The degree of its importance lies on the level of autonomy required in the image processing system [5]. It is mainly responsible for extracting accurate edges from the image, which paves the image for any farther processes, such as object recognition, feature extraction, and 3D environment constructing [6].

So far, many works have been done to develop a unique algorithm that can guarantee, to some degree, high-quality edge detection with less noise and less computational time [7]. Most of these works were basically relying on designing a static mask that sequentially moves through the pixels in the image toward extracting edges [8, 9]. Despite the success of these works, the idea of predesigning the mask for specific tasks, however, could limit the performance of these models, especially when dealing with the complexity of the real-world applications.

In recent years, researchers have been investigated the biological concept of "retina" to try to overcome the above
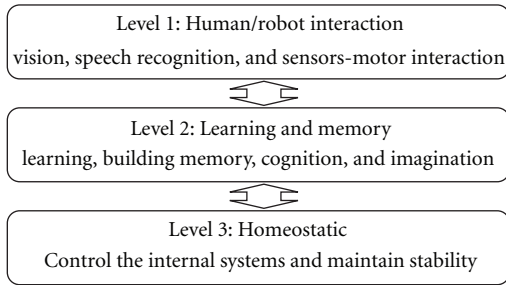
FIGURE 1: The desired human-like robot control system.

FIGURE 2: The structure of center surround neural connection in retina—adopted from (Moran eye center 2010 [20]).

problem, since it is now widely accepted that the biologically inspired technology is a powerful source to achieve an accurate model with simple structure and less computational time [1].

A cognitive vision research group in Hungarian Academy of Sciences [10], for instance, has designed a model for edge detection based on the center surround structure of the receptive fields that are presented in retina. They have simulated the eye tremors and drifts to enhance the output image. The filter, however, was static and cannot distinguish between the noises and the edges.

Becerikli and Engin [11], on the other hand, have tried to solve the problem by building a neural network and trained it by the back propagation to ignore the noises. Their work, however, does not guarantee the ability to detect different edge formations. Metta [12] has worked in an attention system for a humanoid robot based on robot eye movements and space variant vision for motor control. This work has missed the nature of the variety of photoreceptor in the human retina.

In Gifu University in Japan, some researchers have tried to imitate the natural eyes movement by recording the eyes and head movement of real humans and simulate the movement to the humanoid robot "YAMATO". Although they succeeded to simulate human eye movement, the robot missed the real advantage behind these behaviors [13].

Along this line of research, we are here proposing a bioinspired vision system that would help to develop advanced human-robot interaction in an autonomous humanoid robot. First, we enhance the robot vision accuracy in real time by applying a novel dynamic edge detection algorithm abstracted from the rules that the horizontal cells play in mammalian retina. Second, we improve the robot's tracking ability, to have it works as a supporter for the first algorithm, by designing a variant photoreceptors distribution similar to what exists in the human vision system. The proposed model is constructed by an artificial neural network and applied in parallel to the robot view. The experimental results proved the validity of the model. The robot was able to achieve accurate edge detection in real time and could build a mental map that helped it to be aware of the frontal users and to achieve a positive interaction with them.

This paper is organized as follows. The following section highlights the biological concept of human vision system. Section 3 describes in details the proposed algorithm. S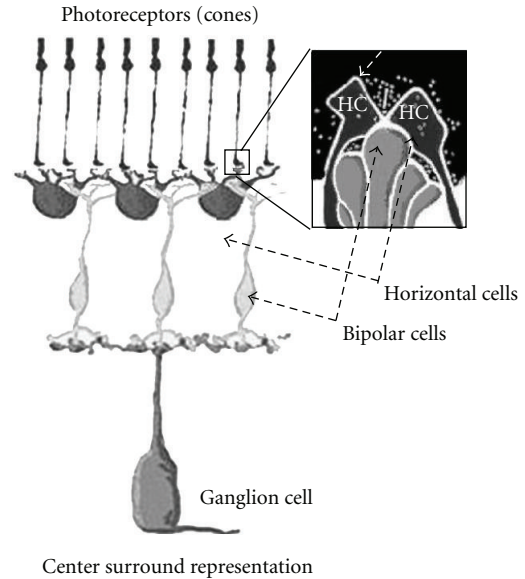ection 4 shows an experimental setup and results. Finally, Section 5 concludes the work and gives the direction for the future work.

2. Vision System: Biological Review

The retina, as well as the eye movement (saccades and pursuit), are working with each other in the human vision system in such a way that allows us to obtain better vision and awareness of the surrounding environment. The retina, which is a part of the brain, is responsible for performing the first stage of the image processing, for example, edge and motion detection, before passing its signals to the brain for any farther processing [14]. Saccades and pursuit, in contrast, can be considered as a retina vision supporter. They are responsible for various kinds of voluntary or involuntary movements of the eyes, which help to track objects and/or to direct the attention [15].

2.1. Retina's Neural Connection. The neural connections between ganglion cells and bipolar cells in the retina are responsible for edge detection in cold-blooded vertebrate, by performing a processing called the center-surround representation. However, they are not the only source for edge detection in mammalian vision, since the horizontal cells play also an important role in enhancing this mechanism [16].

As can be seen in Figure 2, the horizontal cells are between the photoreceptors (cones) and bipolar cells. When the light is absent, the horizontal cell releases the neurotransmitter gamma aminobutyric acid (GABA) that acts on GABA receptors [17]. This phenomenon has an inhibitory effect on the photoreceptors. Therefore, when the light is shined onto a photoreceptor, the photoreceptor hyperpolarizes and reduces the release of glutamic acid (glutamate), which excites the horizontal cell to reduce the release of GABA [17].
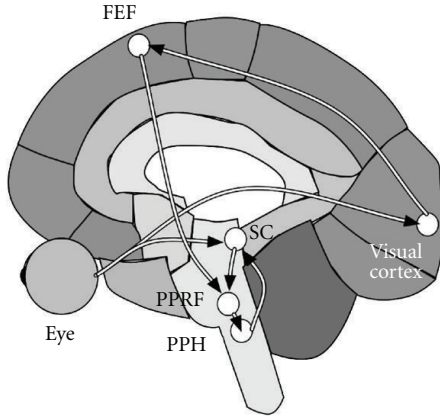
FIGURE 3: Neural signaling pathway for saccade eye movement.

FIGURE 4: The proposed dynamic two-layered neural network mask.

This reduction of inhibition leads to a depolarization of the photoreceptors. Such as complex process, however, is still a subject of hot debate in the community of retina scientists [18].

The functionality of the horizontal cells can be summarized by two main points: (1) A single ganglion cell constructs a center surround representation from the output of many bipolar cells, and transfers this representation to the brain. (2) Horizontal cells improve this center surround representation by changing the output of the bipolar cells, according to the input pattern [19]. According to Verweij et al. [17], the reduction of GABA varies according to the brightness of the light that shines onto the photoreceptors and the time that this light is present. This gives the center surround representation the accurate information about the edges [17].

From the above phenomenon, we have got the idea of designing a dynamic edge detection technique, which is self-adapted during the time based on the given input pattern.

2.2. Human Eye Movement. Saccades and pursuit are two types of the eye movement. They work jointly to support the retina by constantly placing the image of the object of interest on the center of its fovea.

Saccade, on one hand, operates to rotate both eyes to the same direction so that the desired image always falls on the fovea. Since the vision is poor during the saccade, it operates at high speed (up to 500 degrees per second). Pursuit, on the other hand, operates at low speed to smoothly follow a moving object [21].

From the biological point of view, the neural pathway that is responsible for generating the saccade eye movement can be illustrated in Figure 3. In the figure, the eye projects signals to both the visual cortex and the superior colliculus (SC). When the object is moving in the peripheral area of the vision, SC elicits a saccade via the paramedian pontine reticular formation (PPRF), so that the image of the object can be placed in the fovea and examined by the visual cortex. This, in consequence, gives the order to the frontal eye fields (FEFs) to remember the location of the object [22]. Therefore, the SC directs short latency involuntary saccades
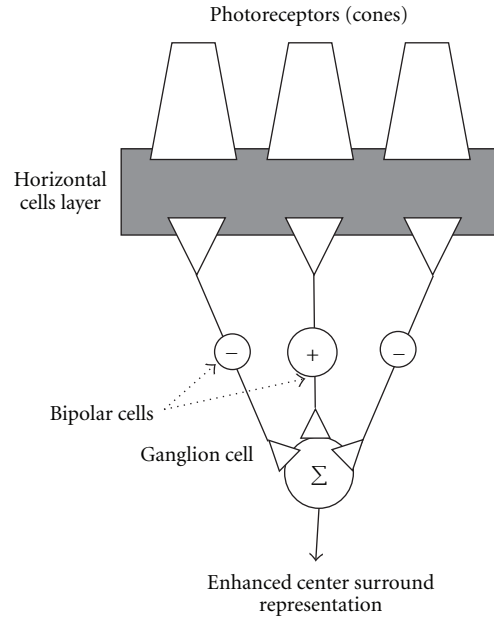
to an unexpected movement, and the frontal eye fields direct long latency voluntary saccades to the remembered target.

During this process, the low spatial frequencies are attenuated, while the higher spatial frequencies, which would otherwise be blurred out by the eye movement, remain unaffected. This phenomenon is known as saccadic masking [23].

The first step towards the initiation of pursuit is to follow a moving target. Signals from the retina activate neurons in visual cortex, which responds selectively to directions of movement. The processing of motion in this area is necessary for smooth pursuit responses [24]. It is likely involved in providing the signal to initiate pursuit, as well as selecting a target to track [25].

By reviewing the above functionality of both the retina and the eye movement (saccade and pursuit), our proposed model is designed.

3. The Proposed Model

3.1. Dynamic Edge Detection. In the retina, there are three layers of cells: photoreceptors, bipolar cells, and ganglion cells. Horizontal cells interconnect between photoreceptors and bipolar cells (Figure 4).

To simulate the effect of horizontal cells, we consider two functional layers: one is the input layer of photoreceptors and another is the output layer of ganglion cells. We treated the bipolar cells as static weights that perform the center-surround calculation. Horizontal cells are considered as dynamic weights (1). The horizontal cells are responsible for adjusting the synaptic weights between the photoreceptors and the bipolar cells. The above-proposed mask is a part of multi masks that are represented in parallel in the robot's vision (Figure 5).
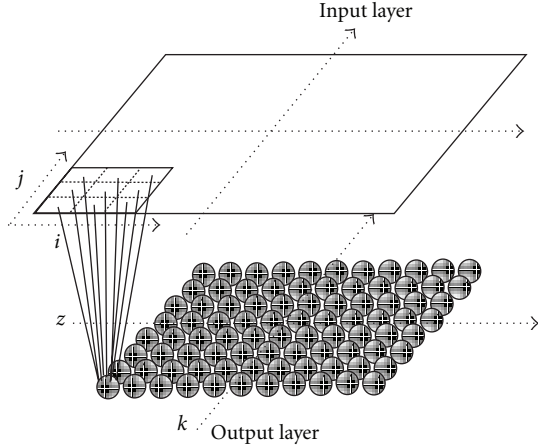
FIGURE 5: The proposed model, showing the connections between the input layer and the output layer, with overlap of two pixels in the input layer.
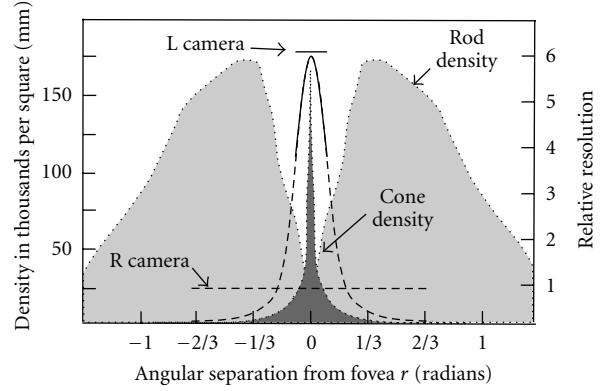

FIGURE 6: The method used to combine the two different images from robot cameras, which are similar to the photoreceptor distribution in human eyes—modified from (Metta [12]).

From the mathematical point of view, mapping between the input layer (In) and the output layer (Out) can be expressed by

$$\text{Out}_{(z,k)} = \frac{1}{9} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\text{In}_{(z,k)i,j} \times \left(w_{(z,k)i,j} + M_{i,j} \right) \right), \quad (1)$$

where w represents the synaptic weight (GABA variation) of the connection between photoreceptors and the bipolar cells, M is constant values of the bipolar cells' connection with the ganglion cells, which is responsible for center-surround representation. i and j are the location of the input neuron. k and z are the address of a particular mask in the robot's vision (Figure 5).

The synaptic weights in each mask are updated similar to that done by the glutamates and GABA effects [17], where the redistribution of the weights is done on the basis of the contribution of each input pixel to the location of z and k

$$C_{(z,k)i,j} = \frac{\left(\text{In}_{(z,k)i,j} \times w_{(z,k)i,j} \right)}{\left(9 \times \text{Out}_{(z,k)} \right)}, \quad (2)$$

where inputs with high/low contribution value gradually increased/decreased its related weights overtime (3). From this equation we can notice that the weights in the mask are decreased by 1/9 (this ratio is selected for simplicity, since the mask are constructed from 9 input and one output, and it is close to the real decay in the horizontal cell contribution [17]) of its previous value, by adding the contribution to the result we get the next distribution of the weights, this adaptation mechanism are similar to the horizontal cells contribution in the center surround representation [17]

$$\left[w_{(z,k)i,j} \right]_{(t)} = \frac{8}{9} \left[w_{(z,k)i,j} \right]_{(t-1)} + \left[C_{(z,k)i,j} \right]_{(t-1)}. \quad (3)$$

It is important to maintain the overall summation of the weights in each mask to be a constant in any given time, which is equal to the summation of the contribution (4).

This is important to the filter to produce a similar result with different elimination of the input

$$\frac{1}{9} \sum_{i=1}^{3} \sum_{j=1}^{3} w_{(z,k)i,j} = \sum_{i=1}^{3} \sum_{j=1}^{3} C_{(z,k)i,j}. \quad (4)$$

We believe by the above propose a model, each mask can gradually adapt its synaptic weight connections during the time to reflect the input pattern of the edges in its associated area.

Even though in this work we worked with luminance information of the view (gray images), this work can be extended in future to include colors detections. It is widely known that the receptive fields of the color channels in retina have a center-surround organization, that is, red-green and blue-yellow receptive fields. By introducing the proposed dynamic weights adaptation mechanism, we should be able to improve the representation of the colors in the robot view.

3.2. Human-Like Eye Movement. Since the proposed mask relies mainly on time to approach the best resolution, it is important for the robot to move its eyes (2 color cameras mounted in the robot's head with 2 degree of freedom each) to trace the object of interest and try always to place it in the center of its vision (fovea), similarly to that behavior found in the human eye (see Section 2 for details).

3.2.1. Variant Photoreceptor Distribution. To achieve such a phenomenon, an equivalent variant distribution of the actual photoreceptor in the retina (cone and rods) is designed in the robot vision (Figure 6). The cones are dense in the center of the retina (fovea) and responsible for edge detection and object recognition. The rods are absent in the fovea but dense in the peripheral area. Rods, therefore, has less resolution and it is responsible for motion detection.

Here, we applied the edge detection algorithm to the images captured from the left camera, while the right camera is responsible for motion detection. Since the left camera images are six times as higher in resolution as that of the right
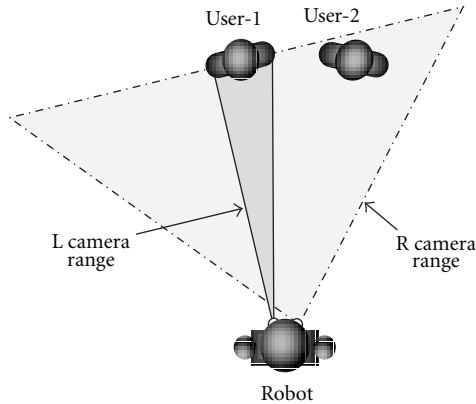
FIGURE 7: The top view of robot interacting with two users.

camera images (Figure 7). We use (5) to combine the two images in one single image

$$Result(r) = \begin{cases} Source_{Left}(\sqrt{r}) & |r| \le \dfrac{1}{8}, \\ Source_{Right}(\sqrt{r}) & |r| > \dfrac{1}{8}, \end{cases} \quad (5)$$

where r represents the radians from the center of the view to the periphery (Figure 6), $Source_{Left}$ is the source image from the left camera, and $Source_{Right}$ is the source image from the right camera (Figure 8).

3.2.2. Eye Movement Control System. The above-listed representation of the image is used to support the robot eye movements (Figure 9). The image in the SC is the foveal image from the retina. It has two parts: fovea and periphery. When the user approaches the robot and starts the conversation, the robot, at first, locates the user in its fovea and sends an inhibitory signal to PPRF. This allows the robot to give much attention to the target user by ignoring most of the movement surrounding him. The location of the user will be temporarily memorized in the FEF. When the user is interacting with the robot and at the same time starts to move, the FEF will provide the PPRF by the user direction and velocity, so the PPRF can send signal to rotate both eyes smoothly and keep the user always in its fovea.

In the PPRF stage, the robot will decide to keep looking at the user or change the attention to any object moving in the periphery depending on the conversation state and the characteristic of the object itself. If any moving object in the periphery attracts the robot attention, then the PPRF will send signals to motoneurons to generate saccades and move the eyes to the spot of the movement. Concurrently, the PPH will change the location of the SC in the mental map so that the new input will not overlap with the previous information.

When the robot moves its eyes to the new location, it will have the chance to examine that location in its fovea for another user. If another user exists, the FEF will store the location of the new user in its mental map. If the Robot notices that the second user is interesting to start interaction with it, the robot will try to interact with both users and update their locations continuously. Thus, the robot will have

a memory about the current users' location even though they are out of its eyes range.

4. Experimental Results

We have conducted three experiments: the first is to examine the validity of the proposed dynamic edge detection algorithm. The second is to examine the validity of the proposed robot eye movement techniques to support the ability of the dynamic algorithm. The third is to show a complete scenario of a robot equipped with the above two models to examine the ability of a human-like robot to achieve a multiuser interaction in office-like environment. All the following experiments are conducted on a physical human-like robot "Robovie-R2" with two CCD cameras with resolution of 480×480 pixels, each camera can rotate horizontally and vertically to simulate human eye movement (ATR-Robotics [26]).

4.1. Extracting Edges in a Real Time. In this platform, we are examining the validity of the proposed dynamic edge detection to extract clear edges in real time from a user standing in front of the robot (Figure 10). Note that at this stage, the robot's eyes movement techniques to trace the user movement were not activated. Therefore, when the user changes his location, the learned edges were lost and the network needed to be retrained to adapt to the changes. From the figure, it can be seen that after passing 40 frames (2 seconds), the network converges to detect accurate edges from user's face.

We compare between the proposed algorithm and the well-known canny edge detection algorithm, obtained by a detector of edges similar to the work done by Boaventure and Gonzaga [27], and to do this comparison, we have to test these algorithm with ground truth images that represent the edges in the scene.

It is well known that the edge detection algorithms perform better with higher resolution images, so to generate the ground truth set of images that describe the location of the correspondent edges, we first use sequence of images with resolution of 260×260 pixels to produce ground truth images, and we tested the edge detection algorithms with lower-resolution images 120×120 pixels.

The performance of the edge detection algorithm can be obtained through a set of direct measurements, such as the number of correctly detected edge pixels, called true positive (TP), the number of pixels erroneously classified as edge pixels, called false positive (FP) and the amount of edge pixels that were not classified as edge pixel, called false negative (FN), From these measures, the following statistical indices can be obtained.

The percentage of pixels that were correctly detected (P_{co}):

$$P_{co} = \frac{TP}{\max(N_I, N_B)}, \quad (6)$$

where N_I represents the number of edge points of the ground truth image and N_B the number of edge points detected.
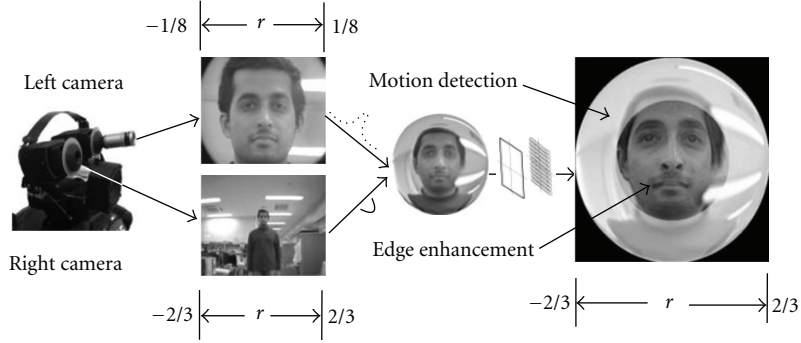
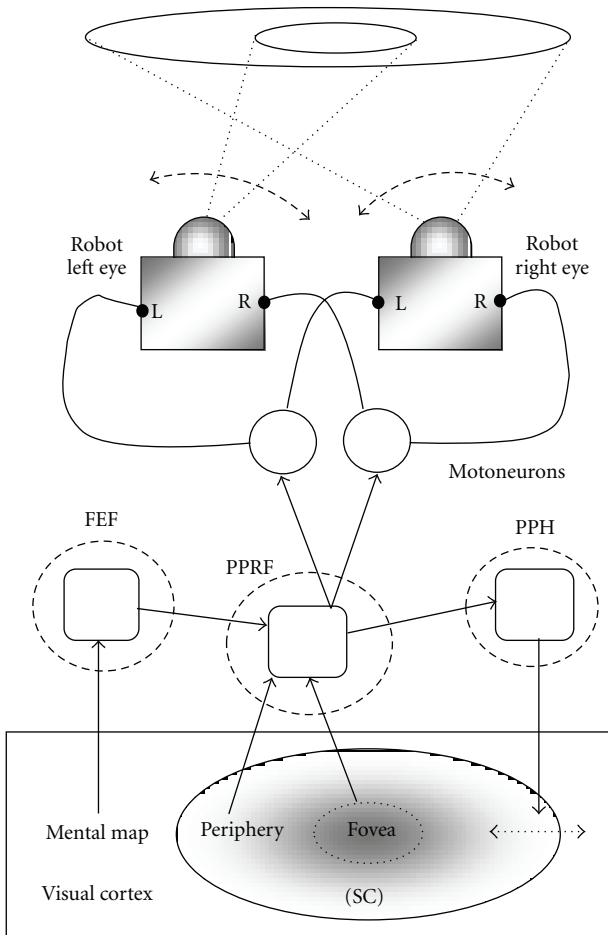FIGURE 8: The foveal representation of the view after applying (5).

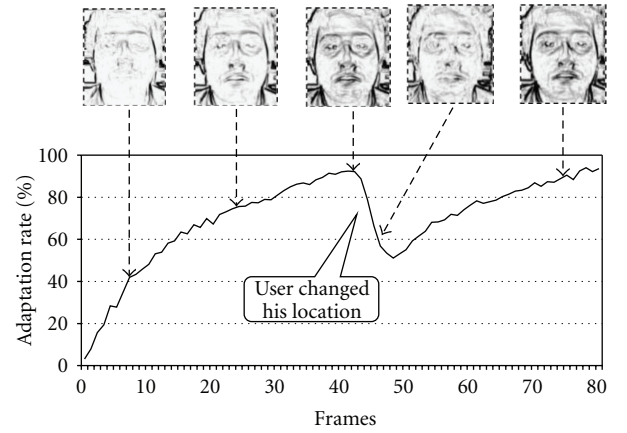FIGURE 9: Robot eye movement control system.

FIGURE 10: The convergence of the weights in the proposed network in 4 seconds (80 frame).

The values of statistical indices represented by (6), (7), and (8) ranging between 0 and 1, and reach ideal values in case 1 for P_{co} and 0 for indices P_{nd} and P_{fa}.

The distance to the ideal edge detector d varies between 0 and $\sqrt{3}$ (9), where the value 0 represents the perfect fit for this measure; that is, the best edge detector among several detectors will minimize this distance

$$d = \sqrt{(P_{co} - 1)^2 + P_{nd}^2 + P_{fa}^2}. \tag{9}$$

Equation (10) can calculate the performance of the edge detection algorithm (Figure 11)

$$\text{Performance} = \left(1 - \frac{d}{\sqrt{3}}\right). \tag{10}$$

From the figure, we can observe that our proposed model can detect clearer edges than the canny edge detection; in addition to this, it is interesting to say that the dynamic feature of the proposed model gives the ability of each mask to adapt itself to suit within the input pattern and allow it to overcome the noises, and we used (11) to compare between the edge detection algorithms with the present of noise, as shown in Figure 12

$$\text{Performance} = \left(\frac{\text{Performance}_{\text{noice}}}{\text{Performance}_{\text{original}}}\right). \tag{11}$$

The percentage of pixels that were not detected (P_{nd}):

$$P_{nd} = \frac{\text{FP}}{\max(N_I, N_B)}. \tag{7}$$

The percentage of pixels that were erroneously detected as edge pixels, that is, the percentage of false alarm (P_{fa}):

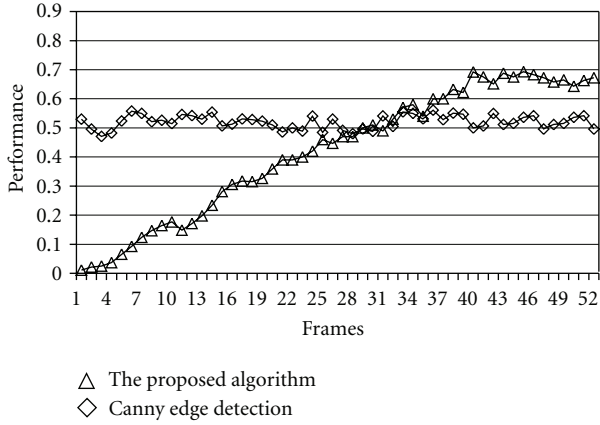$$P_{fa} = \frac{\text{FN}}{\max(N_I, N_B)}. \tag{8}$$

FIGURE 11: Comparing between the proposed algorithm and canny edge detection algorithm by the time passage (frames).
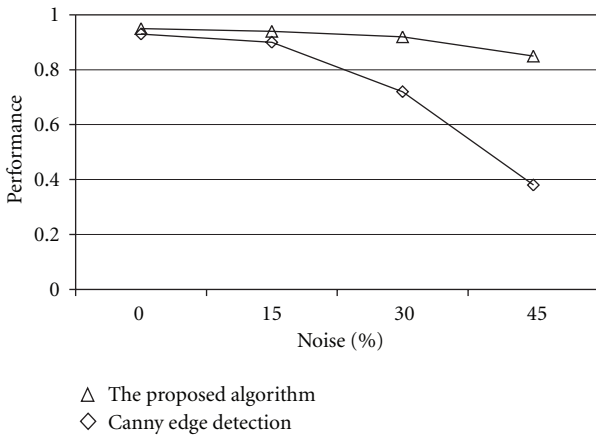

FIGURE 12: The performance of our algorithm and canny edge detection algorithm with the present of noise, showing the result of our proposed filters after its weights converges (after passing 40 frames).

4.2. Robot Eye Movement. As Figure 10 shows, whenever the user changes his location, the network needed a time to readapt to the new input, which can be considered a drawback in the system. To overcome the problem, the robot uses its eye movement techniques (saccade and pursuit) to always trace the user face and tries to center it on its fovea.

We borrow face detection algorithm from OpenCV Library (Face Recognition OpenCV [28]), which allows the robot to center the user in its fovea, as we can see in Figure 14. After applying the eye movement technique, the convergence rate kept approximately stable and the robot succeeded to keep the user in the center of the robot's vision (fovea), although the user changed his location twice (Figure 13).

4.3. Advanced Human-Robot Interaction. In this stage, we applied the above combined proposed models to the robot. The robot was set in an office-like environment with a number of students moving around. As early mentioned, in addition to the running models presented in this study, the robot was also run by a simple face recognition program
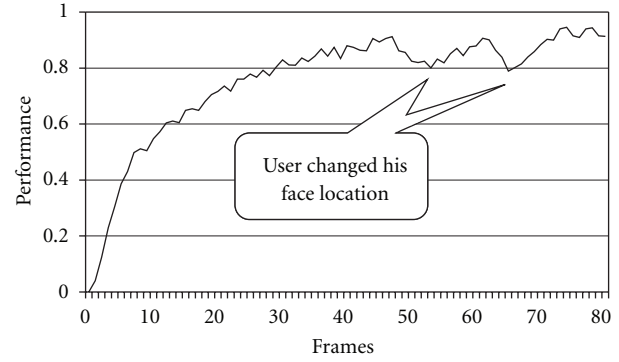

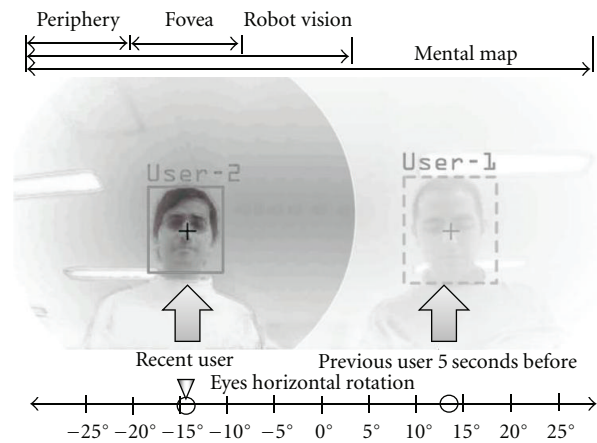FIGURE 13: The adaptation of the synaptic weights after applying robot eye movement.

FIGURE 14: Robot vision system, the robot is interacting with two users in the same time, by memorizing their location in its mental map.

(Face Recognition OpenCV [28]) and a speech recognition program borrowed from Microsoft Corporation (Microsoft-speech [29]).

At the initial time, the robot started to look around randomly giving priority to moving objects. When a user (user-1) approached the robot and started a conversation, the robot gave attention to the user by centering his face into its fovea and started responding to the conversation with the user. During the conversation, the robot always attempted to keep the user into its fovea, training its mask, as long as the user was giving it attention. The robot was also neglecting the other users who were moving around even though the robot was aware of them because of the advantage resulting by the proposed control system Figure 9.

After the passage of time, another user (user-2) approached the robot and engaged in the conversation. The robot could successfully interact with both users at the same time in a natural way. While giving attention to user-2, the robot kept memorizing the location of the user-1 in its mental map so that it could return to him and keep continuing the conversation between them. After a period of time the user-2 had left the scene, and the robot gave back its all attention to the user-1 (Figure 14).

As we can see from this scenario, the robot can maintain the user face in specific location of his vision (fovea) that allows our dynamic edge detection algorithm to converge and produce accurate edges from the user's face, which can be used for face expressions recognition and user identification, as higher level of human-robot interaction.

5. Conclusion

This work is part of a series of studies that aim to develop a human-like controller capable of resembling living organism behaviors in some specific characteristic (Figure 1). More precisely, this study is concerned with enhancing the human-robot interaction level by developing a bioinspired vision system.

In the first stage, we enhanced the robot's vision accuracy in real time by applying a novel dynamic edge detection algorithm abstracted from the rules that the horizontal cells play in the mammalian retina. The algorithm was constructed by multimasks, each of which was represented by a two-layer artificial neural network and applied in parallel on the robot's vision. Synaptic weights, in each mask, were updated gradually during the time based on the founded edges in the image.

In order to support the first algorithm, in the second stage, we improved the robot tracking ability in the attempt to keep the subject in a certain area in the fovea so that the weights in the mask maintain its stability by designing a variant photoreceptors distribution similar to what exists in the human vision system. In this distribution, edge detection processes lay more densely in the central region called fovea, while they become gradually sparser in the periphery and vice versa for the motion detection. The distributed nature of the network would also allow for a parallel implementation, making real-time frame-rate processing a definite possibility, which helps in implementing the robot eyes' movement.

Experimental results were focused to examine the validity of the proposed model in achieving edge detection in an efficient manner with less noise than those obtained by a static model. We believe that our proposed method would be efficient for any dynamic application, where the moving edges require always detecting. The robot could give more attention to the main subject and at the same time keep aware of the potential targets that are moving around. The robot is able to build a mental map that helped it to be aware of the frontal users and to build a positive interaction with them.

We believe that besides that this study focuses on building a model inspired from a biological concept toward new solutions to robotics, but more importantly is the goal of gaining a better insight of how the brain of living systems operates to solve such sort of problems.

Acknowledgments

References

[1] D. Floreano and C. Mattiussi, *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*, The MIT Press, Cambridge, UK, 2008.

[2] V. Manian and R. Vasquez, "Multiresolution edge detection algorithm applied to SAR images," in *Proceedings of the IEEE Geoscience and Remote Sensing Symposium*, vol. 2, pp. 1291–1293, July 1999.

[3] A. Rydberg and G. Borgefors, "Extracting multispectral edges in satellite images over agricultural fields," in *Proceedings of the International Conference on Image Analysis and Processing*, pp. 786–791, 1999.

[4] A. H. Tewfik, F. A. Assaad, and M. Deriche, "Edge detection using spectral estimation techniques," in *Proceedings of the 6th Multidimensional Signal Processing Workshop*, pp. 34–35, 1989.

[5] J. C. Stewien and T. L. J. Ferris, "The asterisk operator. An edge detection operator addressing the problem of clean edges in bone X-ray images," in *Proceedings of the 2nd International Conference on Knowledge-Based Intelligent Electronic Systems (KES '98)*, vol. 3, pp. 28–31, April 1998.

[6] K. Karantzalos and D. Argialas, "Improving edge detection and watershed segmentation with anisotropic diffusion and morphological levellings," *International Journal of Remote Sensing*, vol. 27, no. 24, pp. 5427–5434, 2006.

[7] L. Meylan, D. Alleysson, and S. Süsstrunk, "Model of retinal local adaptation for the tone mapping of color filter array images," *Journal of the Optical Society of America A*, vol. 24, no. 9, pp. 2807–2816, 2007.

[8] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing, presented at a talk at the Stanford Artificial Project," in *Pattern Classification and Scene Analysis*, R. Duda and P. Hart, Eds., pp. 271–272, John Wiley & Sons, 1968.

[9] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[10] B. Resko, A. Roka, A. Csapo, and P. Baranyi, "Edge detection model based on involuntary tremors and drifts of the eye," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 11, pp. 648–654, 2007.

[11] Y. Becerikli and H. Engin, "Alternative neural network based edge detection," *Neural Information Processing*, vol. 10, pp. 193–199, 2006.

[12] G. Metta, "An attentional system for a humanoid robot exploiting space variant vision," in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, pp. 359–366, Tokyo, Japan, 2001.

[13] M. Shamoto, K. Kato, and K. Yamamoto, "A implementation of Humanoid Vision-analysis of eye movement and implementation to Robot," in *Proceedings of the Annual Conference of the Society of Instrument and Control Engineers (SICE '07)*, pp. 744–747, September 2007.

[14] M. Kolesnik, A. Barlit, and E. Zubkov, "Iterative tuning of simple cells for contrast invariant edge enhancement," in *Biologically Motivated Computer Vision*, vol. 2525 of *Lecture Notes in Computer Science*, pp. 91–106, 2002.

[15] R. J. Krauzlis, "The control of voluntary eye movements: new perspectives," *The Neuroscientist*, vol. 11, no. 2, pp. 124–137, 2005.

[16] R. Nelson and H. Kolb, "Synaptic patterns and response properties of bipolar and ganglion cells in the cat retina," *Vision Research*, vol. 23, no. 10, pp. 1183–1195, 1983.

[17] J. Verweij, E. P. Hornstein, and J. L. Schnapf, "Surround antagonism in macaque cone photoreceptors," *Journal of Neuroscience*, vol. 268, pp. 1053–1056, 2003.

[18] D. M. Dacey, "Primate retina: cell types, circuits an colour opponency," *Progress in Retinal and Eye Research*, vol. 19, no. 5, pp. 647–648, 2000.

[19] F. Werblin, "Synaptic connections, receptive fields, and patterns of activity in the tiger salamander retina," *Investigative Ophthalmology and Visual Science*, vol. 32, no. 3, pp. 459–483, 1991.

[20] John Moran Eye Center, University of Utah, http://webvision.med.utah.edu.

[21] B. Cassin and S. Solomon, *Dictionary of Eye Terminology*, Triad Publishing, Gainsville, Fla, USA, 1990.

[22] I. Israel, "Memory-guided saccades: what is memorized?" *Experimental Brain Research*, vol. 90, no. 1, pp. 221–224, 1992.

[23] S. J. Judge, R. H. Wurtz, and B. J. Richmond, "Vision during saccadic eye movements. I. Visual interactions in striate cortex," *Journal of Neurophysiology*, vol. 43, no. 4, pp. 1133–1155, 1980.

[24] X. J. Wang, "A neural circuit basis for spatial working memory," *Neuroscientist*, vol. 10, no. 6, pp. 553–565, 2004.

[25] J. D. Coltz, M. T. V. Johnson, and T. J. Ebner, "Population code for tracking velocity based on cerebellar Purkinje cell simple spike firing in monkeys," *Neuroscience Letters*, vol. 296, no. 1, pp. 1–4, 2000.

[26] ATR-Robotics, http://robotionary.com/robotics/communication-robot-development-platform-robovie-r2.php.

[27] I. Boaventure and A. Gonzaga, "Method to evaluate the performance of edge detector," in *Proceedings of the The Brazilian Symposium on Computer Graphics and Image Processing*, Sibgrapi, Brazil, 2009.

[28] Face Recognition OpenCV, http://opencv.willowgarage.com/wiki/FaceRecognition.

[29] Microsoft-speech, http://en.wikipedia.org/wiki/Microsoft_Speech_API.