Proposed Enhanced Object Recognition Approach for Accurate Bionic Eyes

Mohammad Shkoukani[1], Hesham Abusaimeh[2] , Rawan Abu Lail[3] and Mays Khawaldeh[4]

[1] Computer Information Systems, Applied Science University
Amman, 11931 Jordan

[2] Computer Networks Systems, Applied Science University
Amman, 11931 Jordan

[3] Computer Information Systems, Philadelphia University
Amman, Jordan

[4] Computer Information Systems, The Arab Academy for Banking and Financial Sciences
Amman, Jordan

Abstract

AI has played a huge role in image formation and recognition, but all built on the supervised and unsupervised learning algorithms the learning agents follow. Neural networks have also a role in bionic eyes integration but it is not discussed thoroughly in this paper. The chip to be implanted, which is a robotic device that applies methods developed in machine learning, consists of large scale algorithms for feature learning to construct classifiers for object detection and recognition, to input in the chip system. The challenge however is in identifying a complex image, which may require combined processes of learning features algorithms. In this paper an experimented approaches are stated for individual case of concentration of objects to obtain a high recognition outcome. Each approach may influence one angle, and a suggested non-experimented approach may give a better visual aid for bionic recognition and identification, using more learning and testing methods. The paper discusses the different approached of kernel and convolutional methods to classify objects, in addition to a proposed model to extract a maximized optimization of object formation and recognition. The proposed model combines variety of algorithms that have been experimented in differed related works and uses different learning approaches to handle large datasets in training.

Keywords: *Bionic eyes, image recognition, image processing, object recognition, learning machine.*

1. Introduction

Cybernetic organism (cyborgs) are systems combining biological and engineering parts, bionic eyes are prosthetic devices implanted in the humans to retrieve their visual insight. The chips and cameras connected to the brain processes images and through the learning agents, it recognizes the images and sends the data to the brain for humans' decision making or sensing awareness.

However, technology is not to be reliable in function at all times, in terms of accuracy, for machines malfunctions sometimes and may lead to more damage than what we thought. And computer logic has never equalized with humans and matched all decision making a normal human may take.

Although all of those, Advanced Technology has been significant in improving human health and alters health problems. One of these technologies is Nano-technology, which changed the practice of experts to excel more in curing and performing sophisticated surgeries. These advanced technologies have helped starting from the prosthetic devices, which are used to substitute the missing parts of the human bodies, to the organs, which are replaced with artificial organs. These artificial organs have been trained to act as human functions. Many researchers have extended researches and techniques to find solutions for human visual replacement of the blinds, until the innovation of bionic eyes has occurred recently. Artificial intelligence (AI) has created waves of interest in the field of artificial vision which creates a revolution in the field of medicine. AI relies on object recognition, where in some context it has been mentioned the motivating application include biometric identification, content-based image retrieval, and handwriting recognition [1, 11].

In addition, learning machines uses several agent designs to process an image, feedback, and representation of what has been captured. For example, the agent learns a function (input) from images to produce a Boolean output (object recognition) one of the learning agents are logical agents that consist in its component propositional and first-order logical sentences (decision trees). During the object recognition, we may face several challenges such as the following:

- Image segmentation; an image may contain multiple objects, therefore framing the image into subset of pixels in order to correspond it to a single object at a frame.
- Pixel values; brightness, color, angle, and size of an image effects value of an image. It will affect the recognition of the object.

2. Overview

In order to explain our proposed research we should understand how the human eye works and image is formed. The human eye consists of many components as shown in Figure 1. Light rays have a role in forming a clear image on the retina, and certain processes contribute in light rays, which are: refraction of light rays, accommodation of the lens, constriction of the pupil, and convergence of the eyes Right after, an image shall be formed on the retina in order to stimulate its receptors (rods and cones). This leads to produce nerve impulses which must be conducted to the visual areas of the cerebral cortex for interpretation [2, 13].

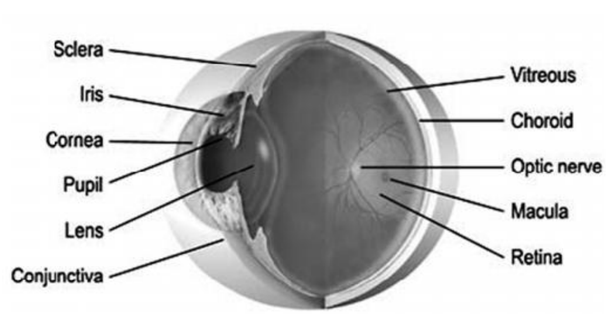

Fig. 1: Anatomy of Eye [3].

Therefore, some are born blind, where others had sight then lost it due to damages on the optic nerve, the brain, or any part of the visual pathway between them. Tissue damage for example resulting from a stroke may hit the color area recognition.

This condition is called acquired cortical color blindness. A corneal transplant may be a replacement to cure color blindness. Thus most damages may not always result in total loss of sight. This is how vision impaired.

Firstly, we should explain what Bionics is, the word bionics has been formed by mixing the words of "Bio: life" and "nics: electronics". It represents the role of science engineering in studying a natural living object and modeling it through complex methods to simulate it and model it with it’s fully function behavior. Secondly, Bionic Eyes are simply an embedded in the human eye connected to the brain. They are like eyes glasses but holds micro-cameras to fetch images and process them, before send them to the brain in order to get an understanding of the image [12, 13].

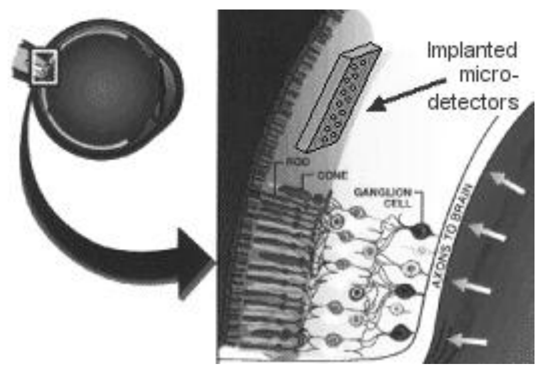

Fig 2: A schematic diagram of the retina with implanted detectors [4].

In 1988, directing an electrical current to the nerve ganglia behind the retina started, when artificial retina component chip helped a blind person to see points of light. This artificial retina comes with implanted micro-detectors as shown in the previous Figure 2. Proving that, the degenerated retina may have nerves behind it which may function well. After that, Scientists start thinking of replacing the retina with a device in order to convert images to electrical impulses, which will be composed the vision. However, the eyes environment such as the salty conditions of the eye could encourage corrosion in the delicate electronics required for this technology. Therefore, Researchers have designed a microchip that uses an external laser as power supply, so there is no need to use a battery in such a wet environment.

The bionic eyes system consists of:

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012
ISSN (Online): 1694-0814
www.IJCSI.org

a) External parts (Camera and electronic image processing chip), these external parts placed on the eyeglass frame. They are responsible about capturing and converting the scene to pixels. The eye array of pixels consisting about l00 million rows and 5 million columns.
b) Internal parts (retinal microchip); after the image array reach this part via laser or radio frequency. It converts the array to electrical current that can be represented as two dimensional grids of electrodes placed near the retina [5, 13].

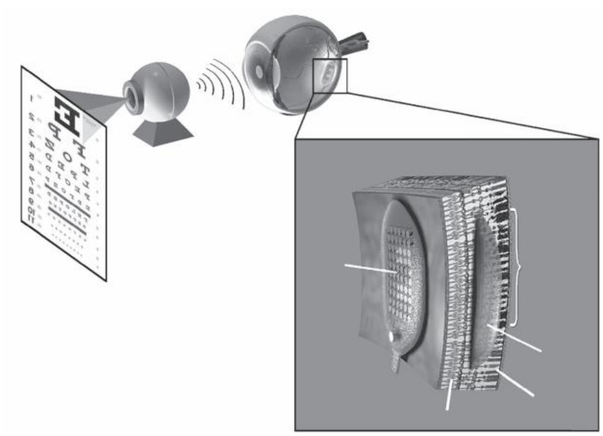

Fig 3: The process of capturing an image until generating the electrodes grid over the retinal chip [6].

As shown in Figure 3; the image of an object at distance Z in the scene is produced at a fixed distance from the lens Z', where the relation between Z and Z' is given by the lens equation in which f is the focal length of the lens. Given a certain choice of image distance Z' between
The nodal point of the lens and the image plane, scene points with depths in a range around Z0, where Z0 is the corresponding object distance, will be imaged in reasonably sharp focus. This range of depths in the scene is referred to as the depth of field. Note that, because the object distance Z is typically much greater than the image distance Z' or f, we often make the following approximation:

$$\frac{1}{Z} + \frac{1}{Z'} \approx \frac{1}{Z'} \Rightarrow \frac{1}{Z'} \approx \frac{1}{f} \qquad (1)$$

Therefore, the image distances Z' FZ f. We can continue using the pinhole camera perspective projection equations to describe the geometry of image formation in a lens system. In order to focus objects that are at different distances Z, the lens in the eye changes shape, whereas the lens in a camera moves in the 2-direction.

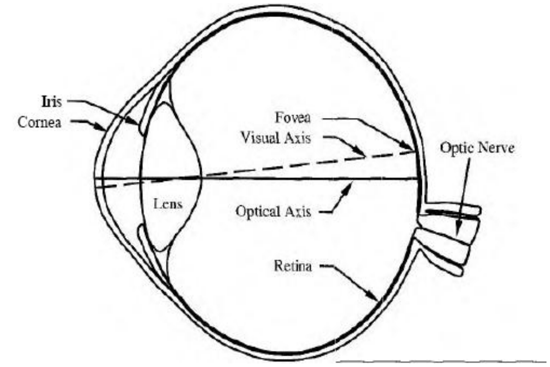

Fig 4: Horizontal cross-section of the human eye [1].

Identifying an object needs supervised learning, where the image segmentation is important to partition the image into pixels, where each subset pixels represent a single object. Then the pixels are moved into classifiers to label them using top-down process.
The geometrical transformations such as translation, scaling and rotation, or transformations of image brightness caused by moving light sources physically, have a different character than the intra-category variation such as exists between different human faces. Obviously, learning is the only way to learn about the different kinds of human faces, or objects.
Brightness-based and feature-based recognitions are easier to use than the shape-based object recognition. Brightness-based recognition checks which pixel brightness values are used directly. While feature-based recognition involves the use of spatial arrangements of extracted features such as edges or key points [1]. In order to extract the visual information of the image, we have to construct the tasks of manipulation, navigation, recognition, intermediate representations. Early vision image-processing algorithms extract primitive features about the image, such as edges and regions.
In case of object movement, the image must be first stabilized as in Figure 5 in order to extract features then recognized it accordingly.

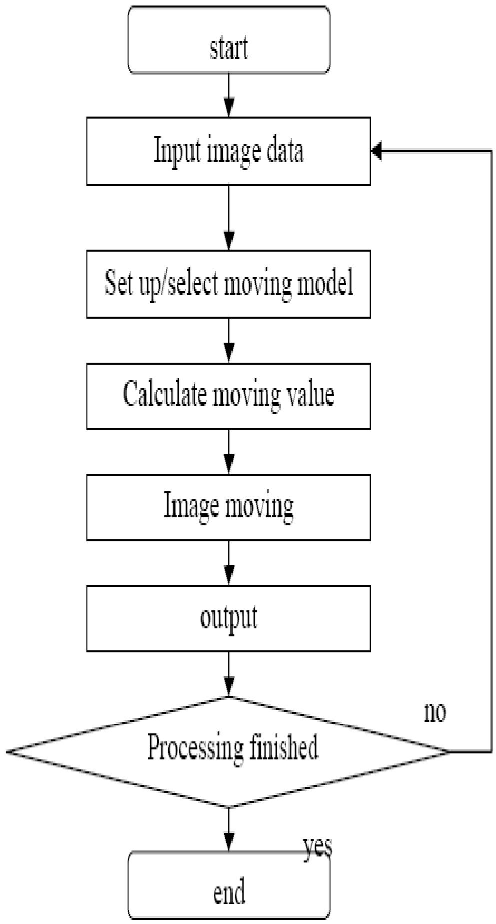

Fig 5: Flow of electronic image stabilization [7]

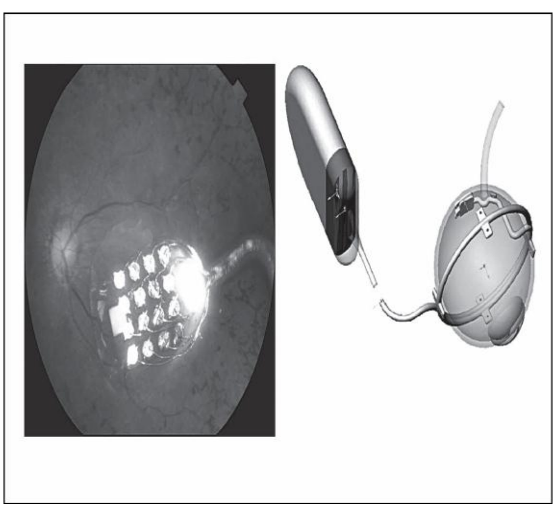

Fig 6: A way of simulate the electrode array in the eye using camera flash.

In Figure 6, sixteen platinum electrodes shielded with a silicon rubber are forming the prototype electrode array, the array is held in place by a retina tack, which is obscured by reflection from the camera flash. After that, an extraocular stimulator is implanted behind the ear; a subcutaneous, multiwire cable runs along the temple into the eye and terminates at the electrode array [6].

3. Related Work

During the phase of recognizing moving target, two-dimensional image need to be reverted multi-dimensional information as shown in Figure 7. The process needs large volume of calculation which is difficult to ensure real-time. In order to make up lost information, we introduce additional knowledge which will create additional noise. An important part of research is to find how to image directly which utilize the target multidimensional information at the same time. In addition, subject to structural constraints, resolution of compound eye imaging is far lower than the human eye and the existing higher solution imaging. However, low-resolution imaging is precisely conformed to its limited brain processing capacity of insect which could detect targets with small computation [7].

Model used in experiment:

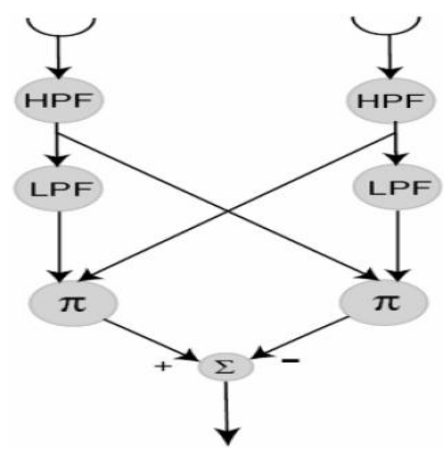

Fig 7: Hassenstein-Reichardt elementary motion detection model [7].

To ensure better performance and output recognition the chip system on bionic eyes is best to demonstrate number of key elements such as:

(i) Provide a large dataset including synthetic examples to learn from.

(ii) Use high resolution imagery camera to verify image detection.
(iii) Input multiple views of an image for better recognition [8].

Experiments:

$$\min\left(\frac{N_{12}(\mathcal{R}_1,\mathcal{R}_2)}{N_1(\mathcal{R}_1)},\frac{N_{12}(\mathcal{R}_1,\mathcal{R}_2)}{N_2(\mathcal{R}_2)}\right) \quad (2)$$

Where; N12 (R1, R2) is the number of SURF correspondences where the first point falls inside R1 and the point falls inside R2. N1 (R1) is the total number of SURF correspondences where the first point falls inside R1 (regardless of where its corresponding point is found in the second image), and similarly for N2. We discard low-scoring candidate correspondences, and then use a greedy algorithm to find the best pairing between detections in the two images [8].

A key question for object recognition is then how to measure the similarity of image patches based on the attributes of pixels within them, because this similarity measurement is used in a classifier. Techniques based on histogram features converts the individual pixel attribute values into bins and then compute a histogram over the discrete attribute values within a patch. The similarity between two patches can then be computed based on their histograms. Unfortunately, the binning introduces quantization errors, which limit the accuracy of recognition.

Handling a full size image, a complex –convolution-networks uses a small receptive field and share weights of hidden a visible layers among all locations in an image.

Neural networks are reliant on trainable hierarchal architecture that has been successfully applied to character detection, pose estimation, face detection, and generic object recognition. Additional experiment results explained that the depth kernel descriptors outperform existing approaches significantly. In future work, it would be interesting to investigate whether hierarchical kernel descriptors can boost the performance further [9].

Experiments:

Table 1: Comparison to existing category approaches [9].

Approaches	Depth	RGB	Depth+RGB
LinSVM [14]	53.1±1.7	74.3±3.3	81.9±2.8
kSVM [14]	64.7±2.2	74.5±3.1	83.8±3.5
RF [14]	66.8±2.5	74.7±3.6	79.6±4.0
This work	**78.8±2.7**	**77.7±1.9**	**86.2±2.1**

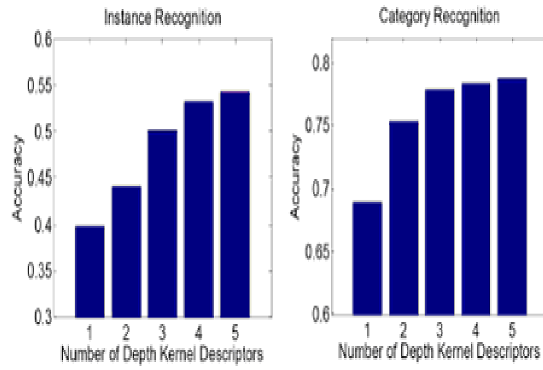

Fig 8: Instance and Category Recognition results [9].

Another challenge is computing the size and distance of an object to identify the edge and depth of an object. In addition to the brightness and light effects that may result in false recognition or cause an error, where by the eye receptive may result in confusing the brain decision and recognition. Therefore the feature descriptors must hold the processing of different inputs of an image such as size, shape, edges, colors, brightness, and more closer to real-time processing, through combining different algorithms for object recognition. And use a large dataset of learning, using supervised and unsupervised learning approaches.

4. Proposed Approach:

Kernel machines learning method:

Kernel machine learning method match the incoming pattern with templates extracted from a training set, followed by a linear combination of the matching scores. This method is combining between supervised and unsupervised training.

Kernel machines are non-parametric (allow the complexity of the solution to increase when more data are available) learning models, which make apparently weak assumptions on the form of the function f () to be learned.

There are some limitations of kernel machines, which involve visual perceptions, where studies have proved that kernel machines leads to fundamentally inefficient representations [10].

Convolutional nets architectures:

Convolutional nets are multi-layer architectures in which the successive layers are designed to learn progressively higher-level features, until the last layer which represents categories. All the layers are trained simultaneously to minimize an overall loss function.

Unlike with most other models of classification and pattern recognition, there is no distinct feature extractor and classifier in a convolutional network. All the layers are similar in nature and trained from data in an integrated fashion.

Convolutional nets are being used commercially in several widely-deployed systems for reading bank check, recognizing handwriting for tablet-PC, and for detecting faces, people, and objects in videos in real time [10].

Suggested Approach to Optimize Object Recognition:

1- Use Hassentein-Reichardt elementary motion detector for higher lens capturing.
2- Perform image stabilization procedure
3- Use multi algorithms to calculate the object distance such as, the lance distance algorithm in addition to finding the density of the object (determining depth features, and image features), EM algorithm (for deforming images, where this algorithm is repeated until convergence or until a maximum number of iterations has been reached.), and IDM kernel "Image discriminative Model" as shown in figure 9.

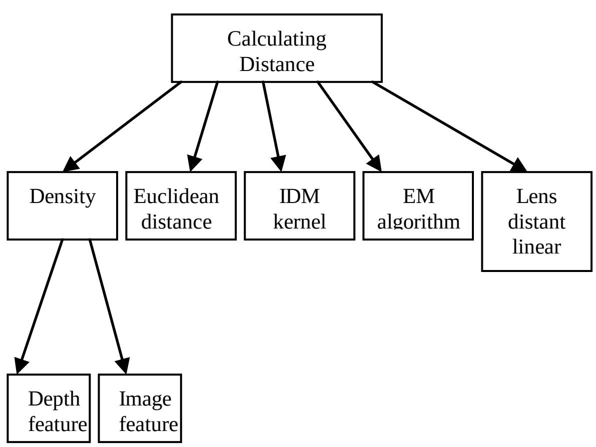

Fig 9: multi algorithms to calculate the object distance

4- As shown in Figure 10, combining more than one classifier method, such as Gaussian classifier and kernel machine model.
The Gaussian density classifier is a generative model approach which aims to represent data as accurate as possible. This model uses a probabilistic formalization method.
Whereas the kernel machine is a vector machine model, which aims to find parameters in order to categorize the data as accurate as possible. This model uses pixel based feature and log-linear interpretation approach as shown in Figure 11.

The Radial basis function (RBF-kernel) has proved to give good classification results. The RBF-kernel is based upon the Euclidean distance.

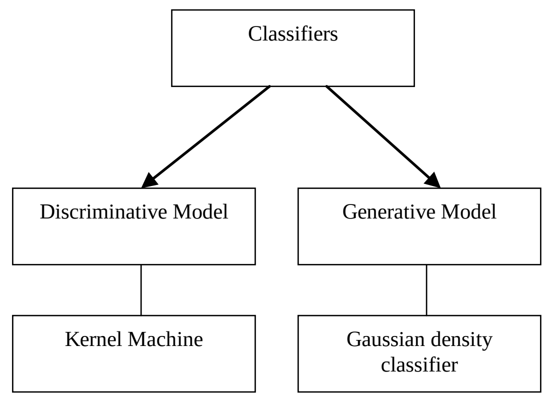

Fig 10: combining many classifiers to represent data

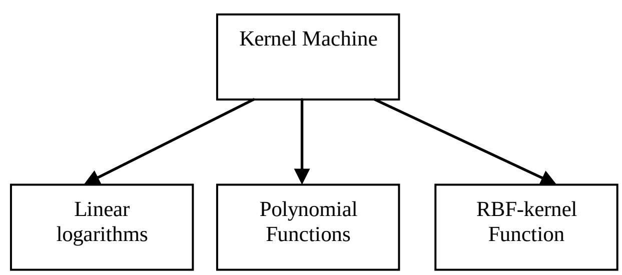

Fig 11: vector machine model

5- Training data using conventional neural networks which are trained using training data that was obtained by randomly deforming the training images and thus the neural network learns about deformed images in the same way as it learns about the initial training data as shown in the following Figure 12.

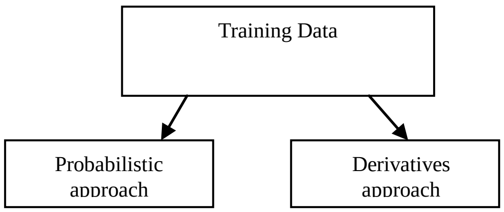

Fig 12: Training data using conventional neural networks

6- Find maximum optimization of object recognition is the final process of the proposed approach.

5. Conclusion

The AI with its' huge learning methods, has managed to replacing a human organs and parts with artificial computerized ones, While Human health is important, an accurate functionality should be considered in designing any artificial component which will be attached to the human body. Therefore, many measurements are to be taken into considerations. Complex algorithms and approaches are required in order to implement a chip and install in the human eyes then link it to the neural nerves in order to send electro pulses to the brain, which will be recognized as an understanding image. Supervised learning machines are consisting of learning agents and unsupervised learning to find complex features to extract of an object.
All techniques however require accuracy to simulate the images captured from the camera implanted on the iris.
The fact of implantation and electrical stimulation of the retina is stated to be safe. The human eye uses over 100 million photoreceptors to form an image, so an electronic device will never replace this resolution, at least not in the near future. However, it may be possible to restore a good amount of visual function that can be augmented by other assistive devices. Considering that so many blinding diseases have no cure and result in complete blindness, a partial restoration of sight will still be a major achievement and is within reach.
The work on enhancing object recognition is still being experimented using new techniques or new combined algorithms. Yet the bionic eyes for the time being is functioning in well graded manner, but to be enhanced in classifying unknown objects and reduce errors that maybe occurred due to brightness, edges, movement, or real-time difference.

To contribute in this study, a suggested model with combined algorithms has been clarified, to optimize the results in recognizing objects and classifying them. All in all with handling different image aspects such as distance calculations, object stabilization, and huge dataset training handling.
Experimenting this proposed approach needs many resources which are not available at the moment for our study therefore our future contribution is to implement our proposed approach in order to justify its result in the real world.

Acknowledgments

The authors are grateful to the Applied Science Private University, Amman, Jordan, for the financial support granted to cover the publication fee of this research article.

References

[1] Stuart Russel and Peter Norving, "Artificial Intelligence A Modern Approach" third Edition, Prentice Hall, 2009.

[2] K. Harish Kumar, M. Santosh Kumar, T.Sanjeev Kumar, " Bionic Eyes - Hope for the blind", http://www.seminarprojects.com accessed on 7/1/2012.

[3] V.P.Vaidya, S.A.Ingole, L.K.Gautam, R.N.Gadbail & V.S.Gulhane, "Bionic Eye, Restores Sight to the Blind", International Journal of Computer Science & Communication Vol. 1, No. 2, December 2010, pp. 251-253.

[4] NASA Science development news, http://science.nasa.gov/science-news/science-at-nasa/2002/03jan_bioniceyes/. January 3, 2002.

[5] M. Humayun , E. de Juan Jr., J. Weiland, G. Dagnelie , S. Katona , R.Greenberg, and S. Suzuki ,"Pattern electrical stimulation of the human retina," Vision Research., vol. 39, pp. 2569–2576, 1999.

[6] J. Weiland and M. Humayun . "Intraocular Retinal Prosthesis", IEEE engineering in medicine and biology magazine, 2006.

[7] Lei Yan, Pengqi Gao, Huabo Sun and Hongying Zhao, "A New Method of High Temporal Resolution, Remote Sensing Imaging: Moving Target. Detection with Bionics Compound Eye." Source: Aerospace Technologies Advancements, pp. 492, January 2010.

[8] Adam Coates, Andrew Y. Ng. "Multi-Camera Object Detection for Robotics", Stanford University, 2010.

[9] Liefeng Bo, Xiaofeng Ren, and Dieter Fox, "Depth Kernel Descriptors for Object Recognition". IEEE International Conference on Computer Vision and Pattern Recognition , 2011.

[10] L. Bottou, O. Chapelle, D. DeCoste, J.Weston, "Large-Scale Kernel Machines", MIT Press, MIT press, Sep.2007

[11] Haipeng Chen, Xuanjing Shen, Xiaofei Li, and Yushan Jin, "Bionic Mosaic Method of Panoramic Image Based on Compound Eye of Fly, Journal of Bionic Engineering vol. 8 p.p. 440–448, 2011

[12] Huabo Sun, Haimeng Zhao, Peter Mooney, Hongying Zhao, Daping Liu, and Lei Yan, A Novel System for Moving Object Detection Using Bionic Compound Eyes, Journal of Bionic Engineering vol. 8 p.p. 313–322, 2011

[13] N.Sharmili, P.S. Ramaiah, and G.Swamynadhan, "Image Compression and Resizing for Retinal Implant in Bionic Eye", International Journal of Computer Science & Engineering Survey, Vol.2, No.2, May 2011

Mohammad Shkoukani received his B.Sc. degree from Applied Science University, Amman, Jordan in 2002, M.Sc. and PHD degrees from The Arab Academy for Banking and Financial Sciences, Amman, Jordan, in 2004, and 2009 respectively, all in Computer Information Systems. His research interests include Agent Oriented Software Engineering, System Analysis and Design, and Electronic Commerce Applications.

Hesham Abusaimeh received his B.Sc. degree from Applied Science University, Amman, Jordan in 2003, and M.Sc. degree from New York Institute of Technology in 2004, both in computer His Ph.D. degree in computer science in the field of wireless sensor networks communication and routing protocols from Loughborough University, UK in 2009. His research interests include Network and Controls, Routing Protocols, Network Lifetime and Consumption Energy, wireless sensor networks, and web applications security.

Rawan Abu Lail received her B.Sc. degree from Philadelphia University, Amman, Jordan in 2002, M.Sc. and Ph.D. degrees from Arab Academy for Banking and Financial Sciences, Amman, Jordan, in 2004 and 2009 respectively all in Computer Information Systems. Her research interests include Artificial Intelligence, Decision Support Systems, and Information Security.

Mays Khawaldeh is currently pursuing her M.Sc. degree at Arab Academy for Banking and Financial Sciences, Amman, Jordan, in Computer Information Systems. Her research interests include Knowledge Based Systems and Knowledge Management.