International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 3, ISSN No. 2455-2143, Pages 55-61
Published Online December - January 2016 in IJEAST (http://www.ijeast.com)

GRASP PERFORMANCE IMPROVEMENT USING ACTIVE LEARNING IN VISION BASED HAND PROSTHESIS

Jayant Kashyap
Student
Computer Science and Engineering
G. L. Bajaj Institute of Technology and Management

Sahil Agarwal
Assistant Professor
Computer Science and Engineering
G. L. Bajaj Institute of Technology and Management

***ABSTRACT* - Vision based grasping is one of the important aspects of human computer interaction. But it has proved intangible without the control environment. One approach towards building more flexible and domain-independent robot grasping systems is to employ learning to adapt the robot's perceptual and motor system to the task. However, one drawback in robot perceptual and motor learning is that the cost of gathering the learning set may be unacceptably high. Active learning algorithms solves this shortcoming by intelligently selecting actions so as to decrease the number of examples necessary to achieve good performance and also avoid separate training and execution phases, leading to higher autonomy. We are describing a prosthetic hand which rapidly learns to select the grasp approach form given target sets and can work efficiently when a new system is interfaced to it. For the experimental purpose the system is trained on few datasets and objects. Active learning algorithms are used to direct the system. The C4.5 decision tree learning algorithm along with Interval Estimation (IE) Statistical methodology is used to convert input-action pair into functional output that results in grasp pre-shaping that mimics the biological mechanisms for the control of grasping. The system also makes use of computer vision and the decision parameters include the measured distance to target object and the dimensions of the target object calculated by the methods of computer vision. Pressure sensors are deployed to measure the weight and hardness of the object. The Output of the system is an effective hits and efficient grasp on training sets and objects with controlled movement of prosthesis.**

GENERAL TERMS
Machine Learning, Pattern Recognition, Image Processing, Prosthesis

KEYWORDS – *Trans radial Prosthesis, Computer Vision, Active Learning, Raspberry Pi, Robots, Grasping*

I. INTRODUCTION

Grasping is an important skill in a robot system and has major applications in robotics in all kind of areas such as industrial, medical service and space robotics. Grasping refers to controlling and immobilizing objects with the fingers. Vision-based grasping and retrieval of objects is an important technique in many tasks. Here robotic systems have the ability to perceive target object features, and the ability to select viable grasp approach for an efficient grasp and would be able to carry out many useful functions. Applications for such a system range from the handling of toxic materials in dangerous environments to the assistance of people with physical disabilities or amputees in relatively benign household environments. In this paper we focus on Grasp performance improvement using active learning methodology and computer vision. Significant development has been made in machine perception techniques, highly dextrous robotics hands, and robotic grasp analysis and synthesis, however, due to limitations in the state of the art in computer vision and perception, robotics still preclude the availability of effective and efficient robotic systems with truly effective vision and manipulation capabilities. The primary reason is the limitations in the reliability of machine perceptual systems in varied environments, as well as the difficulties in the control of dextrous manipulators. The operating environments into which a robot system is placed may have a variety of unfamiliar objects, materials and lighting conditions, making the perceptual and manipulation task difficult without the use of a significant number of domain-dependent techniques.

We are using the sensorial, mainly visual, information to perceive the object features. We then develop a learning framework to apprehend the features of the environment that predict the outcome of the actions of the robot. In this paper we focus on the later. In this paper we introduce a more formidable approach that tries to use experience of real grasping actions to tune the behaviour and the reliability assessment capabilities of the grasping system. More specifically we follow an active learning approach.

We present a practical system for controlling the grasping under transradial prosthesis. The system is mounted on the artificial hand and comprises hardware and software that are convenient for real time implementation. The hardware consists of Raspberry Pi 2, a PiCam and a Standard webcam, Arduino UNO and Ultrasonic sensor and pressure sensor.

1.1. Related Works

In this section, we present an overview of the current research and previous research in the area of robotic grasping and relate it to our work.

1.1.1. Analytical and Heuristic Approaches

Previous research on grasping approach results in two variations. One is analytical approach that model the mechanical interaction between the gripper and the object for efficient grasps [1][2]. When the contact between the object and the gripper is made the coefficient of friction between the two are known and thus the relative force or torque is applied. The other is the heuristic approach inspired by the research of neuropsychologists [3] and orthopaedic medicine.

Analytical approaches to grasping evaluate grasp selection based on mechanics; most often in terms of maximizing some objective function such as grasp stability, slip-resistance, or minimizing some objective function, such as internal forces [4]. This approach has given a great deal of theoretical insight into grasp planning and can be measured by tactile sensing. The grasp quality can be computed as objective function which can further enhanced by optimizing parameters of a dextrous hand.

Another approach to select grasp is to use a heuristic approach which employs rules based on objects and selections that are empirically observed. In this approach grasp generators include some major grasp pre-shape types like cylindrical, palmer, lateral, hook, fingertips and spherical grasps. These different categories were formed based on behavioural analysis of human grasping where different types of hand shapes are observed. As a result many rule-based systems for grasp selection [5] were developed on how individuals perceive objects.

1.1.2. Anomalies in Analytical and Heuristic approaches

These systems are given production rules based on how the human perceives objects and pre-shapes grasps. Both these approaches have a common pitfall in a way that both the systems assume perception to be perfect regarding the shape and other properties. Practically these systems require machine perception to explore descriptions that are not so viable for vision systems in the contemporary technologies. The other major pitfall with the analytical technique is that the *a-priori* assumptions will not be applicable in the changing environments over time due to mechanical wear or other factors.

1.1.3. Autonomous-Heuristic systems

One solution is to design an autonomous system with sufficient constraints and *a-priori* knowledge of shapes using visual representation of the scene. Stansfield [5] described a heuristic rule-based system for grasp selection that used a symbolic multi-component representation of object views derived from a range scanning system. If 2-D object contours can reliably be extracted from vision processing in a particular domain, analytical grasp point selection technique for optimal grasp on smooth closed 2-D contours can be used [6]. Alternative approaches such as Bard [7][8] compute ellipsoidal decompositions of voxel description that rely on fusing multiple stereo vision views in a space occupancy grid of voxels. The pre-shape generation process attempts to identify feasible ellipsoid parts for grasping according to heuristics based on part size and accessibility, followed by a physics-based simulation to verify a grasp before execution.

1.1.4. Anomaly in Autonomous-Heuristic approach

Designing autonomous-heuristic systems for making it general purpose in behaviour by selecting domain-dependent attributes such as object contours, colours, etc., feature extractors and special purpose grasping routines requires custom programming and engineering. Even the definition of new systems, environments, objects and shapes adds significant cost and workload on the programmer and engineers. The scenarios for each deployment can be combinatorially very large and thus reduces the reliability of the system in practical sense. This cost addition and significant re-engineering on every new deployment can reduce the flexibility of the system. The system needs to be adaptive to such deployments by its own.

1.2. Machine Learning approach in vision-based grasping

One approach to avoid the domain-dependent systems is to employ perceptual and/or motor learning techniques. Many researchers [8][9] have found this solution preferable for robotic systems to learn a grasping strategy. The benefit of learning approach is that the system can adapt to the characteristics of the object, its features, surrounding attributes of the operating environment and the motor systems providing the system a higher autonomy.

1.2.1. Supervised and Unsupervised learning approaches

Techniques in learning to grasp have taken both supervised and unsupervised approaches. Supervised approaches are exemplified by a rote learning or "teaching by showing" approached developed by Kuniyoshi et al. and Kang and Ikeuchi[10]. The past observations are used for planning action primitives that the system carries out. A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples. A good sum of *a-priori* knowledge is required for such a system and requires a controlled environment so as to enhance the speed and reliability of the vision system for object recognition in real time. Unsupervised learning approaches do not require human intervention and so are suited for high autonomy. Learner is provided with unlabelled data and its aim is to find patterns in that data. A number of systems have used the output of computer vision, primarily in the form of two dimensional edge contours, to drive either recognition or indexing of previously grasped objects, or to use the local object contour information as the basis of features for classifiers [12].

Employment of an unsupervised memory based learning approach with a vision derived two dimensional polygonal representation with grasp coded by relative location and orientation to the object [11]. Researchers [12] used a contour-based 2D representation of the perimeter of objects and measured grasp location and human selected heuristic quality parameters for the object. During on-line learning the system tried to apply previously attempted grasps to the current object and then compared their predicted fitness using a nearest-neighbour learning rule. If the fitness function finds no grasp, a randomized domain-specific heuristic was used to select a new grasp. Another implementation was done deploying batch-learning of association using connectionist representation with back-propagation learning [13]. Implementation based on full 3D information collected from laser range scanner and generating parameterized superquadric representations was also made and learning was performed as density adaptive learning [14][15].

1.2.2. Anomaly in Traditional learning methodology

One advantage of a learning approach to vision-based grasping is the increased autonomy with self-adaption to limits in sensing action. But in vision-based learning as in other learning methodology the cost of acquiring the learning data set is a big problem towards the application of learning. This task for each exemplar is particularly expensive since a significant amount of time is needed to execute the data processing task. Thus the computational complexity of performing such tasks increases non-linearly with increase in learning set.

1.3. Motivation of Active Learning

The increased cost of acquisition of learning sets is the main motivation behind the Active learning methodology. Traditional supervised learning passively accepts learning sets. But Active learning is a special type of semi-supervised learning which is able to interactively query the user to obtain the desired outputs at new data points. The situation in which unlabeled data is abundant but manual labelling is expensive the learning algorithm can actively query the user for labels. This results in large reduction in training sizes and thus the computational complexity. Active learning systems allow the learner to control where in the input space their exemplars are drawn [14].

They thus permit the learners to use strategies which balance the costs of gathering exemplars for learning (exploration) against the cost of misclassification during the execution of the task (exploitation).

1.3.1. Active Learning approach to Vision-based Grasping

Action selection approaches in active learning keep statistics relating to a particular discrete state or state action pairs. Different aspects have been given. For example, a Support Vector Machine (SVM) based learning approach was given [16] in which learning sets are accessed in the form of superquadrics. Learning algorithms were used to build a regression mapping between object shapes, grasp parameters and grasping quality. SVM regression generates functions whose outputs are scalar. The main motivation of SVM regression is to minimize a bound on expected error for future test data and inheriting interesting generalization properties and sparsity. The result showed the grasp was well chosen however was not always favourable.

Another approach was given in which ID3 algorithm was used along with Interval Estimation statistics [14]. This approach described IE-ID3 algorithm which extends the interval estimation active learning approach from discrete to real valued learning domains by combining IE with a classification tree learning algorithm (ID-3). The input information source is laser range scanner that generates superquadric representation of the object.

Another approach describes the use of mean-shift optimization technique [17]. In this system, at first the robot observes few good grasps by demonstration and learns a value function of these grasps using Gaussian process regression. It then chooses grasps which are optimal with respect to this value function using mean-shift optimization approach and tries them out on the real system. This method exhibits fast learning due to the data-efficiency of the Gaussian process regression framework and the fact that the mean-shift method provides maxima of this cost function.

Recent studies have defined the use of C4.5 decision tree algorithm in vision based learning [18][19]. This has resulted in the fast and efficient grasp.

1.3.2. Concurrent Approaches

The recent studies have shown a shift from vision based to EMG [20] and EEG [21][22] and bio-signals[23] based grasping and human-computer interaction. But collecting that amount of data and analyzing it poses a problem in real time execution. Some systems even used both vision-based as well as EMG concurrently [24].

II. PROPOSED SYSTEM

The whole system (Figure 1) executes in three phases: first, selection of the item from set of items; second, application of image processing techniques and learning and the last, actuation using Raspberry Pi[25][26] and Arduino UNO[27] microprocessor and microcontroller.

At first object selection is performed by the object detection techniques [28]. Two dimensional contour based data is acquired. Ultrasonic sensor measures the distance between the object and the system. Pressure sensors are mounted on the finger tips are used to generate the applicable force for efficient grasping. These form the input descriptors for the object representation.

In this paper we are using the C4.5 decision tree algorithm. C4.5 algorithm is used for grasp classification in this paper. The algorithm builds the set of training data in the same way as ID3, using the concept of information entropy. The training data is a set $S = (s_1, s_2, s_3, ...)$ of already classified samples with each sample s_i consisting of p-dimensional vector $(x_{1i}, x_{2i}, x_{3i}, ..., x_{pi})$ where x_j represent attribute values or features of the sample, as well as the class in which s_i falls. The system approximates the 3D representation of a 2D image based on colour and contour descriptors. The dimensions are calculated with respect to the reference frame.

C4.5 algorithm applies to the input samples to classify them into various classes. The action vector has the four action parameters in its domain space Θ (Elevation), Ω (Angle between Fingers), Φ (Finger Elevation) and W (Torque Factor). The perceptual vector have has four perceptual parameters i.e., colour (c), length (l), width (w) and height (h).

Let P = (c, l, w, h) be the perceptual vector with continuous attributes color, length, width and height and A = (Θ, Ω, Φ, W) be the action vector (Figure 2) with attributes elevation, angle between fingers, finger elevation and torque factor, that are being continuously controlled by perception vector P. The mapping O = f(P, A) is over the domains P and A is the Binary Outcome. The learning is performed by optimizing the function by selecting the possible values of A for the probability of O = 1 given what is currently perceived in P. So the input space is perception-action space. C4.5 algorithm then performs classification over the perceived attributes and action parameters. The algorithm has an advantage over ID3 that C4.5 uses Information Gain concept from information theory. Information gain helps in using unknown values and a possibility to use continuous data. This results in increased accuracy.

Let $freq(O_i, S)$ stand for number of samples in S that belong to class O_i, out of k possible classes, and |S| denotes the number of samples in the set S. Then the Entropy of set S:

$$\mathrm{Info}(S) = -\sum_{i=1}^{k}\left(\left(\frac{\mathrm{freq}(O_i, S)}{|S|}\right).\log_2\left(\frac{\mathrm{freq}(O_i, S)}{|S|}\right)\right)$$

Let set T contains one or more samples all belonging to single class O_i. Then after T has been partitioned in accordance with n outcomes of an attribute test P_J

$$\mathrm{Info}_{P_j}(T) = -\sum_{i=1}^{n}\left(\left(\frac{|T_i|}{|T|}\right).\mathrm{Info}(T_i)\right)$$

and

$$\mathrm{Gain}\ (P_J) = \mathrm{Info}(T) - \mathrm{Info}_{P_j}(T)$$

Thus selection is made on highest gain value. This results in efficient classification.

Let set N represents action space partitioning by current decision tree subject to perceived attributes, P of the current object. Going through all possible values of A with P fixed all leaves are traversed which predicts the outcome for different action parameters. Since search results are to be reduced Interval Estimation [29] statistics is used. Thus the confidence intervals are calculated and the search results are reduced or search results have become selective. For each leaf i in N upper-bound and lower-bound probability of success is calculated in accordance with binomial confidence interval formula. And thus an intelligent selection is made using that confidence interval. With this grasp classification methodology we achieve the improvement in grasp performance.

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 3, ISSN No. 2455-2143, Pages 55-61
Published Online December - January 2016 in IJEAST (http://www.ijeast.com)

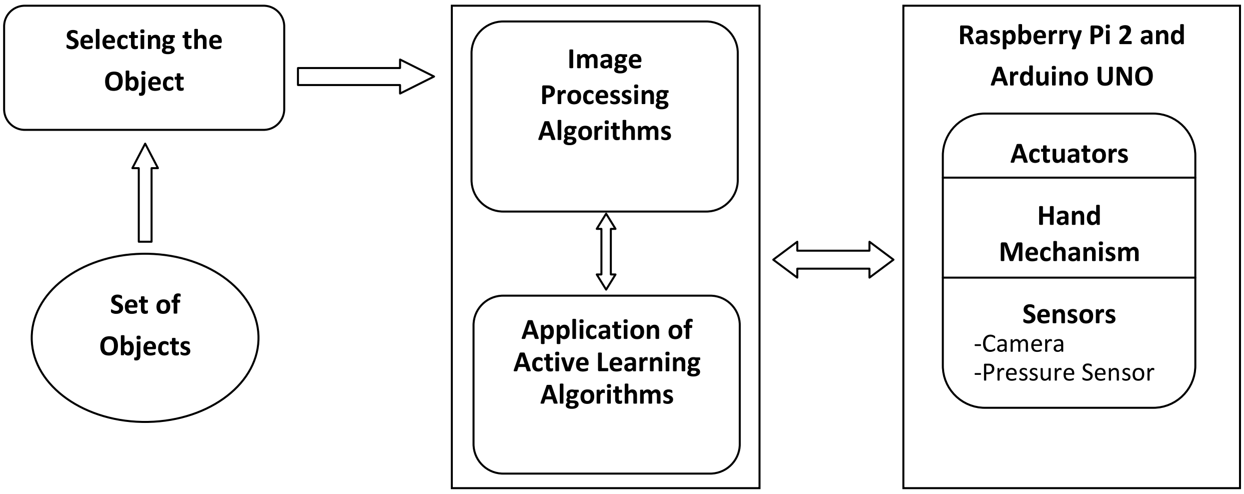

Fig. 1: Scheme for Proposed system

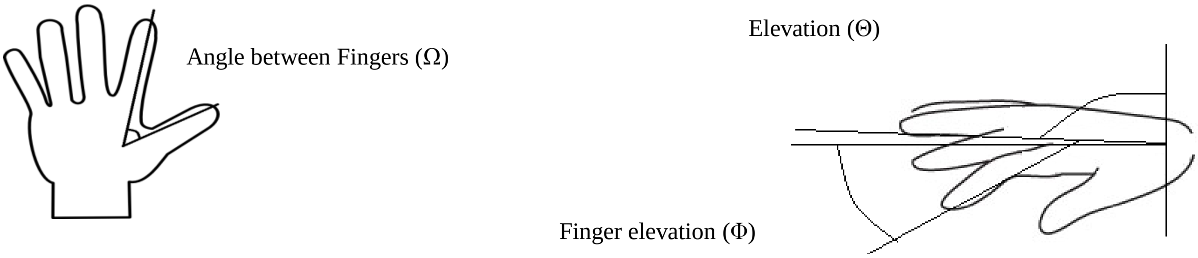

Fig. 2: Action Parameters

The system after executing the learning and pre-shaping operation is actuated using the motors. A pressure sensor is mounted that will help to get data about the weight of the object. An ultrasonic sensor is also mounted to determine the relative distance of the object to the system.

2.1. Experimental setup

The robotic system, anthropomorphic trans radial prosthetic hand, is designed by us and printed with 3D printer for implementation of the proposed model. A PiCam is attached to the prosthesis and is used to gather the information about the object under consideration. A webcam attached to the amputee helps to interact with the objects. Our image processing system will generate an approximation of the object in the 3D format. After the initial phase executes learning phase comes into consideration.

After the learning phase executes the respective class of the grasp is sampled to the actuation system. The microcontroller takes the command and actuates the model as necessary. The success or failure of a grasp is assessed by using the force-torque sensors on the gripper. During the grasp attempt, the system moves the gripper fingers inward and outward for making a near perfect grasp. After having grasp the prosthetic hand lifts the object as defined. It then measures the normal forces on the gripper fingers. If the force is below a threshold value, then the fingers are characterized as empty, otherwise if there is a large normal force, the grasp is labelled a success. This results in overall efficient prehension.

This following sequence of operations to be performed:

1. Webcam selects the object and send that to image processing block.
2. Data reduction and segmentation algorithms to refine the data and produce the descriptors and send it to image processing and learning block.
3. Implementing active learning using C4.5 algorithm and generating a decision tree.
4. Interval estimation algorithm selects best values for action parameters.
5. Grasp attempts are made and outcome is assessed.
6. Success attempt is recorded.
7. Repeat if command given.

This algorithm gives an optimal approach to our setup and grasp performance is much better and efficient in real-time operation. With the increment of the tests the system will get more flexible and reliable to the working environment and thus relieves the amputee for his more effort.

III. FUTURE SCOPE AND APPLICATIONS

The system we proposed extends its applications ranging from prosthesis to industrial robotics to remote handling of robots which cannot be otherwise possible. The system can be further implemented on the other prosthesis types such as lower limb prosthesis and craniofacial prosthesis with various other attributes such as temperature and pressure. This prosthesis system can be extended from vision-based to EEG and EMG based on signal handling. This system can also be applicable in autonomous robotics. The active learning technology we used here can also be used in integration with other technology such as text recognition and natural language processing where noise cutting is an important aspect.

IV. CONCLUSION

We proposed an anthropomorphic prosthetic hand that has the cognitive ability based on visual perception. We also presented the learning algorithm which combines the Interval Estimation statistics for exploration with decision tree algorithm C4.5. The combination of these algorithms gave a synergistic learning outcome. The system is also endowed with pressure sensor which helps to determine the gravitational effect as well as rigidity of the object. The system starts from a limited set of initial motor and perceptual and with gradual learning develops more sophisticated ways to interact with the environment. The perceptual parameters calculation is done on what is perceived visually. The first goal of this research is to find tradeoffs between good grasping capabilities of the device. The system is designed by considering the effort minimization of the amputee.

The robot system perceives the features and sends them to the learning block. The learning algorithm defines classes of grasps based on action parameters. The output decision tree from the learning block defines the control system of the actuators. The learning and image processing is performed on Raspberry Pi development kit and the actuation is carried on using the Arduino UNO development kit. The robotic systems we described here can be extended to complex environments and we believe that this is just a step in humanoid robotics but gives a necessary support to the respective field of computer vision, machine learning, prosthesis and grasp improvement.

V. REFERENCES

[1] Ulrich, Nathan T. “Methods and apparatus for mechanically intelligent grasping.”, U. S. Patent No. 4,957,320. 18 Sep. 1990.

[2] Cutkosky, M.R. (1989). On grasp choice, grasp models and the design of hands for manufacturing tasks. IEEE Journal of Robotics and Automation, 5(3):269–279.

[3] Jeannerod, M. & Biguer, B. (1981). Advances in the analysis of visual behavior, chapter Visuomotor mechanisms in reaching within extra personal space. MIT Press, Boston.

[4] Iberall, T. Jackson, J., Labbe, L.&Zampano, R.(1988). Knowledge-based prehension: Capturing human dexterity.

[5] Stansfield, S. (1990). Knowledge-based robotic grasping. In Proceedings, 1990 IEEE International Conference on Robotics and Automation, pages 1270–1275.

[6] Taylor, M., Blake, A.&Cox, A. (1994). Visually guided grasping in 3D. In Proceedings, 1994 IEEE International Conference on Robotics and Automation, pages 761–766.

[7] Bard, C., Laugier, C., Milesi, C. & Troccaz, J. (in press). Achieving dextrous grasping by integrating planning and vision based sensing. The International Journal of Robotics Research, 1995.

[8] Bard, C., Troccaz, J. & Vercelli, G. (1991). Shape analysis and hand pre-shaping for grasping. In IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 64–69.

[9] Kang, S. & Ikeuchi, K. (1993). Toward automatic robot instruction from perception-recognizing a grasp from observation. IEEE Journal of Robotics and Automation, 9(4):432–443.

[10] Kuniyoshi, T., Inaba, M. & Inoue, H. (1989). Teaching by showing: Generating robot programs by visual observation of human performance. In Proceedings of 20th International Syposium on industrial robotics, pages 119–126.

[11] Dunn, G.B. & Segen, J. (1988). Automatic discovery of robotic grasp configuration. In Proceedings, 1988 IEEE International Conference on Robotics and Automation, pages 396–401.

[12] Kamon, I., Flash, T. & Edelman, S. (1994). Learning to grasp using visual information. Technical Report, The Weizmann Institute of Science, Rehovot, Israel.

[13] Diaz, R., et al. "Comparison of three algorithms in the classification of table olives by means of computer vision." *Journal of Food Engineering* 61.1 (2004): 101-107.

[14] Salganicoff, Marcos, Lyle H. Ungar, and Ruzena Bajcsy. "Active learning for vision-based robot grasping." Machine Learning 23.2-3 (1996): 251-278.

[15] Selganicoff, M. (1993). Density-adaptive Learning and forgetting. In proceedings of the Tenth International Conference on Machine Learning. Morgan-Kauffman.

[16] Pelossof, Raphael, et al. "An SVM learning approach to robotic grasping."Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on. Vol. 4. IEEE, 2004.

[17] Kroemer, Oliver, et al. "Active learning using mean shift optimization for robot grasping." Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on. IEEE, 2009

[18] Glette, Kyrre, et al. "Comparing evolvable hardware to conventional classifiers for electromyographic prosthetic hand control." Adaptive Hardware and Systems, 2008. AHS'08. NASA/ESA Conference on. IEEE, 2008.

[19] Bermejo, Pablo, Jose A. Gámez, and Jose M. Puerta. "A GRASP algorithm for fast hybrid (filter-wrapper) feature subset selection in high-dimensional datasets." Pattern Recognition Letters 32.5 (2011): 701-711.

[20] Cipriani, Christian, et al. "On the shared control of an EMG-controlled prosthetic hand: analysis of user–prosthesis interaction." Robotics, IEEE Transactions on24.1 (2008): 170-184.

[21] Hazrati, Mehrnaz Kh, and Abbas Erfanian. "An online EEG-based brain–computer interface for controlling hand grasps using an adaptive probabilistic neural network." Medical engineering & physics 32.7 (2010): 730-739.

[22] Dornhege, Guido. Toward brain-computer interfacing. MIT press, 2007.

[23] Martin, H.; Donaw, J.; Kelly, R.; YoungJin Jung; Jong-Hoon Kim, "A novel approach of prosthetic arm control using computer vision, bio-signals, and motion capture," in Computational Intelligence in Robotic Rehabilitation and Assistive Technologies (CIR2AT), 2014 IEEE Symposium on , vol., no., pp.26-30, 9-12 Dec. 2014 doi: 10.1109/CIRAT.2014.7009737

[24] Gardner, M.; Woodward, R.; Vaidyanathan, R.; Burdet, E.; Boo Cheong Khoo, "An unobtrusive vision system to reduce the cognitive burden of hand prosthesis control," in Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on , vol., no., pp.1279-1284, 10-12 Dec. 2014 doi: 10.1109/ICARCV.2014.7064500

[25] Neves, Rui, and Anibal C. Matos. "Raspberry PI based stereo vision for small size ASVs." Oceans-San Diego, 2013. IEEE, 2013.

[26] Brahmbhatt, Samarth. "Embedded Computer Vision: Running OpenCV Programs on the Raspberry Pi." Practical OpenCV. Apress, 2013. 201-218.

[27] D'Ausilio, Alessandro. "Arduino: A low-cost multipurpose lab equipment."Behavior research methods 44.2 (2012): 305-313.

[28] Butko, Nicholas J., and Javier R. Movellan. "Optimal scanning for faster object detection." Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on. IEEE, 2009.

[29] Carlson, Andrew, et al. "Coupled semi-supervised learning for information extraction." Proceedings of the third ACM international conference on Web search and data mining. ACM, 2010.

[30] Gonzalez, Rafael C. Digital image processing. Pearson Education India, 2009.

[31] Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley & Sons, 2012.