ELECTRICAL AND ELECTRONIC ENGINEERING

Bio-inspired system for gesture imitation applied in robotics

INGENIERIA ELECTRÓNICA Y ELECTRÓNICA

Sistema Bio-inspirado de imitación de gestos aplicado en robótica

§ Sandra E. Nope

Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia
§ sandra.nope@correounivalle.edu.co

Abstract

This work describes the necessary aspects for the programming of a robot through the paradigm of learning by demonstration and it proposes a system that keeps concordance with what is known about the functioning of biological systems. Four gestures performed by the arm of different human demonstrators were used in the imitation performed by a simulated arm. To achieve these gesture imitations, a Visual-Motor map was created which allows the robot to give its own interpretation of the observed gesture. The imitation performance was evaluated for its quality and quantity. New indicators are presented for the four gestures that make possible the imitation evaluation beyond the following of the trajectory usually performed in order to define the success of the imitation. The imitation was successful in 3 of the 4 gestures and scored an average of 62% for the poll and 78.6% for the indicators.

Keywords: Learning by demonstration, bio-inspired system, visual-motor map, metrics to evaluate imitation

Resumen

En este trabajo se describen los aspectos necesarios para programar un robot mediante el paradigma del aprendizaje por demostración y se propone un sistema que guarda concordancia con lo que se conoce del funcionamiento de sistemas biológicos. En la imitación se usaron específicamente cuatro gestos ejecutados por el brazo simulado de diferentes demostradores humanos. Para lograr la imitación de los gestos, se creó un Mapa Visuo-Motor que permite que el robot realice su propia interpretación del gesto observado. El desempeño de la imitación se evaluó de forma cuantitativa y cualitativa. Se presentan indicadores novedosos para los cuatro gestos que permiten evaluar la imitación, más allá del seguimiento de trayectorias realizado habitualmente para definir el éxito de la imitación. El éxito de la imitación fue bueno en el caso de 3 de los 4 gestos, lo que produjo una calificación promedio del 62% para las encuestas y de 78.6% para los indicadores.

Palabras clave: Aprendizaje por demostración, sistema bio-inpirado, mapa visuo-motor, métricas para evaluar la imitación.

1. Introduction

A new way of interacting and/or programming robots has been explored in the last years. These new ways do not require for the user or programmer to have a high level of expertise that is mainly related with the knowledge about the robotic platform and the need to foresee all possible operating conditions the robot would face.

Humans (Meltzoff, 1999) and some animals (Galef, 1988) use imitation for learning, whether confronting an unknown environment or task or when they want to improve their performance. This strategy applied to robotics is known as "learning by demonstration" or "learning by imitation" and makes possible the programming of robots with the ability of learning complex behaviors and interacting intelligently with the environment.

An important characteristic of learning by demonstration is that it opens the possibility for the master (human or robot) to program simultaneously many types of robots that could be morphologically different from him, just as he performs the tasks that he wants the robots to learn.

In this work, the imitation is oriented to the performance of a task done by the simulated robot in the same way as it was observed without necessarily following the same arm trajectory of the demonstrator.

Numeral 2 describes the principles that were used for the creation of each of the blocks that constitute the system and the evaluation methodology. Numeral 3 presents the results and a discussion on them. Finally, numeral 4 exposes the conclusions. Bakker and Kuniyoshi (1996) break downed a series of sub-problems that should be resolved in the steps involving imitation. These are: observation, recognition, and reproduction. The works done in learning by demonstration deal with sub-problems such as segmentation, the processing of relevant information that accompanies the demonstrated action and the selection of an appropriate representation for the actions. However, some of those problems such as the selection of an appropriate master, the automatic selection of the time for performing the learned action, and the quantitative evaluation of the imitation when is oriented towards reaching a goal and not as an exact copy of the actions performed by the demonstrator, are still unsolved. The present work deals with the last sub-problem. However, there is an aspect in robot programming by demonstration that was not included by (Bakker & Kuniyoshi, 1996) in the three steps above that is being investigated, it is that the robot will receive human feed-back to improve its performance (Grollman, Daniel & Billard, Aude, 2012) (Grollman, Daniel & Billard, Aude, 2011) (Pastor, P. et al, 2011).

2. Methodology

In order to conceive a simulated robot with the capacity to perform learned tasks from a demonstrator, the relevance of the visual information for gesture learning as perceived by the system and its relation with the motor actions done by the robotic arm was analyzed. This reaction results in the robot giving its own interpretation of the observed gesture which is validated by four gestures performed by the arms of different human demonstrators. These were: Gesture 1: rotating the arm counterclockwise back and forth as when saying "hello", Gesture 2: moving the hand up and down as fanning or calling somebody, Gesture 3: rotating the hand counterclockwise as when cleaning a surface, y, Gesture 4: Getting the hand near and away from the camera as to mean rejection and closeness.

The first three gestures were used by (Kuniyoshi et al., 2003) to evaluate its gesture imitation system. These gestures only involved 2 of the 6 DOF. For this work the gesture 4 was added. This gesture involves all the available DOFs.

140 videos were recorded with different demonstrators performing one of the four possible gestures, that is, 35 video interpretations for each gesture. From these, 48 (12 for each gesture) were used in the initial phase of training and 92 (23 for each gesture) for the validation of the system. The video recording conditions were not manipulated so that they could be constant. However, the extremely fast gesture performances, nocturnal recordings, and the covering of the object of interest (the hand) as others of similar color characteristics, were avoided.

2.1 The imitation concept

Imitation is a concept that generates controversy even in fundamental and ancient sciences such as psychology. Different works in engineering have adopted a variety of achievements over what could be considered an imitation. However, in general terms, imitation implies observing, recognizing, and reproducing somebody else's actions (Thorndike, 1988).

During the observation and recognition process, the imitator is concentrating in the most important aspects of demonstration and rejects details of little importance. This ability is especially important when the demonstrations, even those performed by the same demonstrator, present a high degree of variability in its executions.

On the other hand, in order to replicate a demonstration, it is necessary to have a certain way of measuring similarity between the observed actions and those the imitator could perform. This is known as correspondence problem and its difficulty could increase as the imitator and the demonstrator do not share the same physical characteristics, for example: arms of different length, different degrees of freedom, or different species. In any of these examples, imitating in exactly the same way is not always convenient, possible and/or necessary.

This is how fidelity or precision in the gesture reproduction could vary, and there is not a limit or absolute threshold that would define the success of the imitation except for the fact that whoever judges (the observer) the imitation could not associate the two executions. The association made by the observer between the demonstrator's execution and that of the imitator could correspond to the following: actions, states, and/or goals. Action imitation could be valued for example when the imitator replicates the same gesture with the same hand used by the demonstrator (in a similar or exact manner); although it may be worth to ask if one considers it correct the execution done with the other hand. The imitation goal could be valued when the same final effect is achieved, for instance: moving an object (even with executions and/or the use of completely different tools).

All the considerations over the possible variations in what one can consider or not imitation, are the main difference between the definition that each one can give to the concept of imitation.

2.2 System description

The proposed system has been divided into three big blocks: observation, recognition, and imitation. It is clear that it is possible to develop each one of these blocks with deterministic algorithms; however, biological systems have evolved towards simple and robust solutions that make them worth the effort of imitation and of study. For the later, the vast majority of developments exposed in this work propend to keep concordance with what is known about the functioning of biological systems.

2.2.1 Observation block

Neuro-physiological studies on the visual information processes in the brain start with the following the trajectory that the information makes from the eyes, where the retina transforms the fluctuant patterns of light in patterns of neural activity. The recording of a video with a monocular camera imitates for a single eye this representation. However, this process is only the beginning of a great number of transformations that are made in our central nervous system.

2.2.2 Recognition block

Biological vision systems are capable of extracting many types of information from the environment. Some could detect color or see parts of the infrared spectrum or detect changes in the polarity of light that passes through the atmosphere; others use various eyes to determine the information of depth. However, there is a type of information that is believed to be used for all biological vision systems: movement.

The path of movement process in the macaque's brain is composed by four areas: the striated area (V1), the middle temporal area (MT) the medial superior temporal area (MST) and the 7th area (Bruce and Green, 1990). The neurons in the V1 are activated from a particular movement direction and in at least 3 ranges of speed (Orban, Kennedy and Bullier, 1986). A high proportion of neurons in MT are activated in a similar way that those in VI while other proportion of neurons are selective in a particular angle between the direction of movement and the gradient of spatial speed (Treue and Andersen, 1996). Neurons in the MST area are activated by complex movement patterns such as: comprehension, expansion and rotation with receptive fields that cover the greatest part of the visual field (Graziano, Andersen, and Snowden, 1994) (Duffy and Wurtz, 1997).

A computing algorithm that emulates these four areas is described in (Pomplun et al., 1997), it starts from the estimation of optic flow vectors. The optic flow was strictly described by (Nakayama and Loomis, 1974), it is defined as a field of speed in the plane of image, that describes the pixel movement. Although there are many techniques for the estimation of optic flow, the affine optic flow was selected so that, as its name indicates, it combines the optic flow restriction equation and the corresponding equations to the image formation model for the prospective projection or affinity model. Affine optical flow is determined by the θ values in Eq. 1. This technique was implemented as it is described by (Santos-Victor and Santini, 1996) and it is calculated on the pixels

identified as belonging to the demonstrator's hand. The affine optic flow performs a smoothing of the resulting vectors and it was the optic flow technique with major strength in its estimation for the case of an object with little texture such as a hand and under uncontrolled illumination conditions.

$$\lfloor I_y \quad xI_y \quad yI_y \quad Ix \quad xI_x \quad yI_x \rfloor \, \theta = -I_t \qquad (1)$$
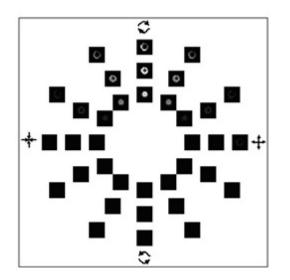
where,

$$\theta = \begin{bmatrix} v_o & v_x & v_y & u_o & u_x & u_y \end{bmatrix} \qquad (2)$$

The information of movement direction used by neurons V1 and MT is found in the angle of the optical flow vectors while the information about speed is found in the magnitude. Since each of these neurons are selected at different directions and speeds of movement, the information provided by the optic flow passes through Gaussian filters that emulate the neuron's tuning taking into account that in the macaque's brain the Receptive Fields (RFs) of the neurons in V1area are circular, are uniformly distributed in the visual field and have an approximate overlap of 20%; they were simulated calculating the mean of all points inside the fixed overlapped circles.

The selection of the type of movement made in the neurons MTS is achieved according to the movement angle observed related to the angle of the magnitude gradient, in that way, it is possible to identify between movements of expansion (0°), of contraction (180°), clockwise rotation (90°) and counterclockwise rotation (270°). Intermediate angles mean that the movement is a combination of two of the basic movements. On the other hand, a pure translation will produce a non-determination of this angle. It is important to note that this identification is too strong to changes in the point of view during the recordings. The response to a two-dimensional filter $G(s_k, \theta_p)$, tuned to a speed s_k and a direction of movement θ_k is determined by the Eq. 3 in which I_s and I_0 are the RF's responses to the optical flow magnitude

and to the angle respectively, and σ_s and σ_θ, are the standard deviation of the same.

$$G(s_k, \theta_p) = e^{\frac{-(I_s - s_k)^2}{\sigma_s}} e^{\frac{-(I_\theta - \theta_p)^2}{\sigma_\theta}} \qquad (3)$$

Figure 1 represents the complete set of neuronal responses for an octagon that rotates counterclockwise. The interior ring corresponds to the lower speed and it increases toward the exterior ring. The three squares in the right side of the horizontal correspond to a 0°angle that increases to 30° in a counterclockwise direction. It is appreciated in the figure that indeed the more brilliant images (neuronal response of greater value) correspond to the clockwise **rotating movement** (superior vertical images). The brightness of these images is related to the tuning speed of the neuron.

Figure 1. *Complete set of neuronal responses to the counterclockwise rotation of an octagon.*

The temporal information of the movement provided by the emulation of the areas involved in the movements previously described are integrated in time, it is made through the motion history image (MHI) proposed by (Bobick & Davis, 2001). The MHI is an image that records the neuronal responses during a period of time that depends on the programming memory. In this representation, the neuron responses will appear strongest depending on how recent a strong neuronal response has happened and on how long that response is kept (Nope, Caicedo & Loaiza, 2008). The memory used in the imitation was of 5 video segments. In order to reduce the dimensionality of 102x76x36 to 36 and the redundancy of the MHI information, this is converted to a histogram that condenses the neuronal responses and is calculated according to Eq. 4.

$$h_\alpha(s_{k=i}, \alpha_{p=j}) = \frac{1}{M} \sum H_\alpha(s_i, \alpha_j, t) \qquad (4)$$

Where M corresponds to the size of $H_\alpha(s_k, \alpha_p, t)$ that is the MHI of the selective neuronal response to a velocity s_i and to an angle α_j. The histograms are used as visual entries to be related to the articulate values of the robotic arm in the visual-motor map.

2.2.3 Imitation block

Meltzof (1997) suggested the existence of an inner visual-motor map. However, the visual-motor map of a human at birth is incomplete due to, among other things, an imprecise knowledge of his/her body. A relevant aspect from the engineering perspective for the construction of an adaptive system is that children acquire the capacity of recognizing and repeating movements through the visual-motor mapping. Therefore, in our research framework, the central role of the macaque's motor visual-mapping is to allow the robot relate a visual stimuli with the motor representations that produces it. This kind of representation have the advantage that it is adjusted to the own visual (system of visual acquisition) and motor (robot's body) capacities of the robot.

A transformation such as the visual-motor map corresponds mathematically to a function that relates a set of variables (visual inputs) in others (motor outputs). Hence the problem of learning a sensory-motor map leads to the problem of function approximation (Koh & Meyer, 1991).

The Generalized Regression Neural Networks (GRNN) belong to the radial based networks and constitute a powerful tool when it comes to learning the function dynamics as it is intended during the construction of a visual-motor map.

The visual-motor map was built through the generalized regression neural networks with a value of one in the smoothing parameter. Although for the "training" 48 of the 140 available videos were used, the number of examples is much greater according to the duration of the gesture performance. The "real" angular values of the articulations during the movements needed to obtain the visual-motor map during the training are described in (Nope, Loaiza & Caicedo, 2010). Figure 2 shows an example of the demonstration performed by a human and the imitation performed by the simulated arm. The result of the imitation demonstrates that visuo-motor map can be used to perform novel gestures out of the trained ones.

Figure 2. Demonstration and imitation of a random gesture.

2.3 Evaluation method

To evaluate the gesture imitation performed through the simulation of a simulated robotic arm, a quantitative and qualitative analysis was made. The first one accounts for the subjective character of the judgment of different observers and is obtained through a poll. The second involves metrics for each gesture that provide a percentage measure of the success of the imitation performance.

2.3.1 Poll

To evaluate subjectively the gesture imitation performed through the simulation of the robotic arm, a two-question poll was made to a group of eight people: 4 of them who were familiar with computers and 4 who were practically not familiar with them.

Initially, the objective of the poll was explained to them, and the gestures to be evaluated were performed associating those with the following names: gesture 1: greeting; gesture 2: fanning; gesture 3: rotating; gesture 4: getting close and away. After that two videos with different visual/performance perspectives were presented to the participants in the poll, and they were asked to identify the gesture. Among the response options, the option "other" was included to respond in case of not finding an appropriate association with any of the pre-established gestures. Subsequently, the demonstrator and the robot performance were presented to the participants simultaneously and they were asked: "Do you consider that the robot imitated the gesture performed by the demonstrator?" The possible answers are YES, NO. This process was repeated for each of the 92 test videos, 23 for each gesture.

2.3.2 Metrics

In spite of the subjective nature of the judgment on the imitation, quantitative measures were used to describe, in this case, the quality of the imitation. The metrics compare the demonstrator's trajectory with those of the imitator (Thobbi, Anad

& Sheng, Weihua, 2010) (Billard et al., 2004), or determining acceptable variance in execution and generalization over initial conditions and perturbations for deriving autonomous controllers from observations of human performance (Dong, S. & Williams, B, 2011). This is little appropriate when it is intended for the imitator to give his/her own interpretation.

Mataric and Pomplum (1998) suggested that people fixed their attention especially in the trajectory followed by the final effector for having the greatest quantity of information on the movement task. The exact definition of final effector depends on the task itself and could vary, for instance, it could be a finger, the whole hand, the arm, etc. In this case, a following of the middle finger trajectory was made and the percentile adjustment to certain geometric shapes of the trajectory was used as an indicator as follows: Gesture 1 (greeting) it is related with two parables (two way movements) where the dependent variable is the Y axes. Although the parables are different, they can coincide. Gesture 2 (moving the hand up and down: fanning) It is related with two parables, but in this case the dependent variable is the X axes. Gesture 3 (rotate the hand) it is related with an ellipse. Gesture 4 does not have a particular form in the trajectory of the final effector that it is observed during the different performances. In fact, it seems to vary much with little changes in the demonstrator's orientation towards the camera. However, it was found that the change in time of the hand size could approximate to a Gaussian curve.

The imitation is evaluated through this indicators allowing for the imitator to give his/her own interpretation of the gesture including variations in speed, differences in physical characteristics of between the demonstrator and the imitator, the distance between the camera and the demonstrator.

3. Results and discussion

3.1 Metrics results

Table1 presents the minimum and maximum values of the entire set of metrics used to validate each gesture, as well as the deviation standard and the average. The results obtained with the proposed metrics leads to the conclusion that good performances of the different gestures were generally obtained, since the shapes were approximate and hardly ever there was an adjustment percent of 100%, even when the real data obtained during the demonstrator's performance was used. The best performances occurred as follows: Gesture 1 (86%); 2 (79.5%); 4 (77.2%) and 3 (71.8%).

3.2 Poll results

The results of the polls on the identification of the gestures performed by the imitator were presented through a confusion matrix in Table 2. It is desired for the diagonal to have values of 100% and 0% outside it. Different values from 0% outside the diagonal are interpreted as a wrong identification of the gesture in row for the gesture in the column. The average successful in recognition was 60%. It is observed that the best identified gesture was gesture 3 followed by 2 and 1. On the other hand, the worst identified gesture was 4 and the majority of times were associated with the option "other". The results of the second question (Do you think that the robot imitated the gesture performed by the demonstrator?) subsequent to the visualization of the gestures performed by the demonstrator

Table 1. *Statistical results obtained for the four metrics for each of the 23 gesture executions.*

Gesture	Average	Standard deviation	Maximum	Minimum
Gesture 1 – greeting	86.0	9.6	100	62
Gesture 2 – fanning	79.5	12.4	100	58.1
Gesture 3 – rotating	71.8	15.5	98.0	41.2
Gesture 4 – getting close/away	77.2	17.1	100	47.4

***Table 2.** Average confusion matrix in the recognition of imitation.*

		Identified Gesture				
		Gesture 1	Gesture 2	Gesture 3	Gesture 4	Unidentified gesture
Executed gesture	**Gesture 1**	62,0	17,9	1,6	14,1	4,4
	Gesture 2	18,5	63,0	1,6	11,4	5,4
	Gesture 3	7,6	1,6	83,2	3,3	4,4
	Gesture 4	8,7	4,4	7,1	<u>39,7</u>	40,2

***Table 3.** Statistics of the results for the second question of the poll.*

	Average	Standard Deviation	Maximum	Minimum
Gesture 1 – greeting	75,0	21.4	91,3	30,4
Gesture 2 – fanning	79,4	13,5	100	65,2
Gesture 3 – rotating	88,0	8,3	100	78,3
Gesture 4 – getting closer/away	44,0	33,9	82,6	4,4

and the imitator are summarized in Table 3. The participants considered that there was a good imitation in the case of gesture 3 (rotating), 2 (fanning) and 1 (greeting) with average percentage superior to 75%. It means than the average percentage reached in the first part of the poll was increased in 13%.

This result indicated a possible difficulty with the visualization that is used to show to the participants the imitation performance which uses "sticks" as approximation to the robotic arm. By the other hand, in an image is lost part of depth perception which is especially important to identify the motion performed in gesture 4. Because of that, after few days the participants were asked the same questions on the 20 videos randomly selected out of the 92 initial ones that included the most deficient gesture imitations. In this manner, it is assumed through the first poll that the observer is used to the visualization.

The participants tended to improve the gesture average identification percentage (from 62% to 68.8%). This improvement was remarkable in 6 of the 8 cases and only for participants 2 and 6 the identification worsen. However, it must be taken into account that the random selection of the twenty videos did not guaranty that the same number of performances of the same gesture was employed,

and the majority of participants evaluated the worse performances in the second poll. Therefore it can be concluded that the evaluation was done in a highly pessimistic scenario.

Summarizing, it is observed that if metrics are used, the best imitation is for gesture 1, while with the poll this gesture comes in third place. Gesture 2 occupies the second place in both analysis cases. Although gesture 3 obtained the worst score in the metrics evaluation, the poll participants considered that gesture 3 was the best executed during the imitation by the robot. This probably happens because for the participants, a good imitation of gesture 3 includes the most varied range of trajectories than that of a rigorous ellipse. The opposite occurs with gesture 4 that based on the 4th indicator seem to have a good performance, however, it was subjectively given a very deficient score (44.02%). That indicates that there is a better correspondence between the metric results for gestures 1, 2 and 3 with the results of the first poll.

In scientific literature, the work that keeps greater similarity with the one exposed in this document corresponds to the one done by Kuniyoshi et al. (2003) in which 3 of the 4 gestures were used in the system evaluation are employed and it is also bio-inspired. However in the system

proposed here includes the emulations of the most involved areas of movement perception which makes possible to identify the types of movement (expansions, squeezes, or rotations) including the direction and the speed of such movements. The imitations performed are improved compared with the mentioned work in the sense that the robot gives its own interpretation and it does not reproduce any pre-established execution.

On the other hand, in this work it was achieved the imitation of a gesture such as getting the hand closer and away which implies the movement of all the robotic arm articulations at the same time. Similarly, the focus of the visual-motor map makes possible for the learned movements to be employed during the imitation of unknown gestures which is impossible to achieve through the work proposed by Kuniyoshi and his colleagues.

Finally, the evaluation on the success of the imitation was improved as the robot performances were scored for quality and quantity. In the last evaluation it was presented a set of indicators that take into account the goal to be achieved and intent to reject in a lesser or greater degree the irrelevant details of the performance.

4. Conclusions

It was presented a system that uses the paradigm of learning by demonstration for a robot to learn how to imitate a set of gestures through the computing algorithms that integrate aspects of neuro-sciences, neuro-physiology and psychology.

The visual-motor map makes possible for the initially learned movements to be employed during the imitation of unknown gestures which is impossible to achieve through the previous work done by Kuniyoshi and his colleagues. Furthermore, it was achieved the imitation of a gesture such as getting the hand closer and away which implies the movement of all the robotic arm articulations at the same time. The visual entries present non-variation at scale.

In order to evaluate the imitation produced by the system, it was proposed a qualitative and quantitative evaluation. The first one was made through the use of polls and it explores the role of the observer who determines in a subjective manner if he/she finds an association between the observed action and the action imitated by the robot. The second determines the percentage of similarity or the success in the imitation based on the geometric shape of the trajectory followed by the final effector.

The simultaneous analysis of the results obtained through the metrics and the poll indicates a good imitation of gesture 1 (greeting); gesture 2 (fanning); and gesture 3 (rotating). Gesture 4 (getting the hand closer and away) obtained a good score according to the indicator, but it obtained a poor score according to the poll participants (44%). There are many reasons for these results. The first one is that this is a gesture for which depth information is important and this is lost due to the visualization used. The second one is that during the making of the visual-motor map there are less examples if one is to take into account the number of articulations that intervene in the gesture.

According to the results, the percentage for imitation seems to be promising in the programming of robots that could have physical differences with the demonstrator; the programming of many robots simultaneously while a task that is going to be taught is performed; and the possibility of reusing the early motor skills learned for the performance of unknown gestures. However the developments made in this area up until now, correspond to relatively simple tasks. The limitations of the proposed system are not related with the principles of learning by imitation, but with the selected algorithms own limitations.

5. Acknowledgements

6. References

Bakker, P. & Kuniyoshi, Y. (1996). *Robot See, Robot Do: An overview of robot imitation*. presented at Workshop on Learning in Robots and Animals, p. 3-11.

Billard, A., Epars, Y., Calinon, S., Schaal, S. & Cheng, G. (2004). Discovering optimal imitation strategies. *Robotics and Autonomous Systems, Special Issue: Robot Learning from Demonstration*, 47, 69-77.

Bobick, A.F. & Davis, J.W. (2001). The Recognition of Human Movement using Temporal Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 257-267.

Bruce, V. & Green, P.R. (1990). *Visual Perception: Physiology, Psychology and Ecology*. Nottingham: Lawrence Erlbaum Associates.

Dong S, Williams B. (2011). *Motion learning in variable environments using probabilistic flow tubes*. International Conference on Robotics and Automation, Shanghai, p. 1976 – 1981.

Duffy, C. & Wurtz, R. (1997). MTS neurons respond to speed patterns in optical flow. *Journal of Neuroscience* 17 (8), 2839-2851.

Galef. (1988). Imitation in animals: History, definition and interpretation of data from the psychological laboratory. In*: Social learning: Psychological and Biological Perspectives*, T. R. Z. B. G. G. (Eds.), Ed.: Hillsdale, NJ: Lawrence Erlbaum. 3-28.

Graziano, M., Andersen, R. & Snowden R. (1994). Tuning of MTS neurons to spiral motions. *Journal of Neuroscience* 14 (1), 54-67.

Grollman, Daniel & Billard, Aude. (2012). Robot Learning from Failed Demonstrations. *International Journal of Social Robotics* 4 (4), 331-342.

Grollman, Daniel & Billard, Aude. (2011). *Donuts as I do: Learning from failed demostrations*. International Conference on Robotics and Automation, Shanghai, p. 3804 – 3809.

Koh, K. & Meyer, D. (1991). Function Learning: Induction of continuous stimulus-response relations. *Journal of Experimental Psychology: Learning Memory and Cognition* 17 (5), 811-836.

Kuniyoshi, Y., Yorozu, Y., Inaba, M. & Inoue, H. (2003). *From visuo-motor self learning to early imitation - a neural architecture for humanoid learning*. presented at International Conference on Robotics & Automation, Taipei, Taiwan.

Nakayama, K & Loomis, J.M. (1974). *Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis*. Perception, 3, 63-80.

Nope, S., Loaiza, H. & Caicedo, E. (2008). Modelo Bio-inspirado para el Reconocimiento de Gestos Usando Primitivas de Movimiento Visual. *Revista Iberoamericana de Automática e Informática Industrial* (RIAI) 5 (4), 69-76.

Nope, S., Loaiza, H. & Caicedo, E. (2010). Reconstrucción 3D-2D de Gestos usando Información de Vídeo Monocular Aplicada a un Brazo Robótico. *Revista Facultad de Ingeniería Universidad de Antioquia* 53, 145-154.

Mataric, M.J. & Pomplun, M. (1998). Fixation behavior in observation and imitation of human movement. *Cognitive Brain Research* 7, 191-202. Meltzoff, A.N. & Moore, M.K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting* 6, 179-192.

Meltzoff, A.N. (1999). Born to Learn: What infants learn from watching us. In: *The Role of Early Experience in Infant Development,* L. A. L.

N. A. Fox, and J. G. Warhol (Eds.), Ed.: Skillman, N.J: Pediatric Institute Publications, 145-164.

Orban, G., Kennedy, H. & Bullier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas Vl and V2 of the monkey: influence of eccentricity. *Journal of Neurophysiology* 56, 462-480.

Pastor, Peter, Kalakrishnam, Mrinal, Chitta, Sachin, Theodourou, Evangelos, & Schaal, Stefan (2011). *Skill learning and task outcome prediction for manipulation*. International Conference on Robotics and Automation, Shanghai, p. 1976-1981.

Pomplun, M., Liu, Y., Martinez, J., Simine, E. & Tsotsos, J. (2002). *A neurally-inspired model for detecting and localizing simple motion patterns in image sequences*. In: Proceedings of the 4th Workshop Dynamic Perception, Bochum, Germany.

Santos-Victor, J. & Sandini, G. (1996). Uncalibrated obstacle detection using normal flow. *Matching Vision and Applications* 9, 130-137.

Thobbi, Anand & Sheng Weihua. (2010). *Imitation Learning of Hand Gestures and its Evaluation for Human Robots*. In Proceedings of the 2010 IEEE Conference on Information and Automation, Harbin, China, p. 60-65.

Thorndike. (1988). *Animal Intelligence*. Macmillan, Ed. New York: Cited in (Galef, 1988). Treue, S. & Andersen, R.A. (1996). Neural responses to velocity gradients in macaque cortical area MT. *Visual Neurocience* 13, 797-804.