



**HAL**  
open science

# Interaction between Humans, NAO Robot and Multiple Cameras for Colored Objects Recognition using Information Fusion

Thanh Long Nguyen, Reda Boukezzoula, Didier Coquin, Eric Benoit,  
Stéphane Perrin

## ► To cite this version:

Thanh Long Nguyen, Reda Boukezzoula, Didier Coquin, Eric Benoit, Stéphane Perrin. Interaction between Humans, NAO Robot and Multiple Cameras for Colored Objects Recognition using Information Fusion. IEEE, Proceedings of the 8th International Conference on Human System Interaction (HSI 2015), Jun 2015, Warsaw, Poland. pp.322-328, 10.1109/HSI.2015.7170687 . hal-01230224

**HAL Id: hal-01230224**

**<https://hal.science/hal-01230224v1>**

Submitted on 29 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Interaction Between Humans, NAO Robot and multiple cameras for Colored Objects Recognition using Information Fusion

T.L.NGUYEN, R. BOUKEZZOULA, D. COQUIN, E. BENOIT, S. PERRIN

LISTIC laboratory, Polytech, University of Savoie Mont Blanc, 74940 Annecy-le-vieux, France

Email: {thanh-long.nguyen, reda.boukezzoula, didier.coquin, eric.benoit, stephane.perrin}@univ-savoie.fr

**Abstract**—Nowadays, the interaction between humans and systems is much better, which allows humans to be able to accomplish a huge range of tasks. Focusing on intelligent communication and human understanding of systems, in this work we develop an application, which allows a humanoid robot to receive oral commands from humans to find a desired colored object. Fuzzy logic is used for color recognition following human perception and reasoning, and Evidence theory is employed to fuse information from multiple cameras in order to improve recognition quality. Experimentation results show that the combination of these methods help the robot to recognize linguistic colors efficiently.

**Keywords**—humans systems interaction, human understanding systems, color recognition, fuzzy logic, evidence theory, robotics.

## I. INTRODUCTION

With the very fast development of technologies, people now can accomplish a huge range of tasks that never can be done before. Admittedly, ones have been proposing and designing a numerous of supporting systems that cater for human life in every domains from industrial production, health care, defense, entertainment to personal purposes. The key of success in working with such kind of systems is the ways they understand correctly what humans want and follow exactly desired actions, it means, the interaction between humans and systems has to be addressed carefully. Given an example in computer systems, few decades ago, they are used only for professionals and specialists, but nowadays most people have their own computers for personal uses. It proves one thing that the interface of communication between humans and computers is better and better, which brings us good usability to obtain efficient performances. However, developing a good interaction strategy is not always an easy task, due to the more and more demanding in having intelligent systems that can follow human perception and reasoning. Such kind of intelligence can help systems intuitively understand what we want, for example in terms of linguistic description, from that improves the quality of communication.

Based on that idea, this work focuses on the problem of understanding colors described in human terms. A humanoid robot receives oral commands from humans to find an object with a color prescribed linguistically like “red”, “blue”, “orange”, “brown”, etc. As the matter of fact, our eyes are able to differentiate a huge range of colors, but we often classify them as into common labels. For example, “dodger blue”, “cobalt” and “sky blue” normally refer to the color “blue”, and

we recognize this through daily experiences in life. However, the interesting point is how the robot – a machine can perceive in the same way as we do, which helps it understand human commands. In addition, the expertise on color perception is transferred from the human to the robot by a learning process based on the Dempster-Shafer theory.

In order to accomplish the proposed scenario, the color recognition problem needs to be carefully addressed. In fact, recognizing colors can be applied in many applications like color indexing, color segmentation, color classification, color naming, and many other situations. This problem has been taken into consideration in several researches, from very simple techniques to more advance ones. For instance, [1][2] considered using Look-up table (LUT) for color recognition in the case of soccer robots. The idea is to firstly construct a table of reference colors through many experiments; this table contains rules of judgment predefined, which allow assigning the most appropriate reference label to the detected color in runtime. Also with soccer robots, [3] proposed to build the distribution of HSL values for the target colors. These distributions are then used to determine the filtering regions for those colors, which specify whether a captured image contains the desired target or not. In [4], the author introduced the application of color recognition for a robotic arm which sorts colored objects. The recognition is done in runtime with a webcam using a blob filter which determines the color of objects through predefined ranges of RGB values. Another interesting case is found in [5] where a humanoid robot is required to find a colored ball using “Hue Blob” detection. It is clear that these methodologies bring us ease of use and may obtain good performance, but they require too many experiments to define fixed parameters for the detection. Moreover, human perception and reasoning are not carefully considered, for example the colors are recognized too strictly, leading to the lack of necessary hesitation and uncertainties.

In order to improve the intuitive recognition of the colors as humans do, we choose to use a Fuzzy logic approach. Indeed, this can describe perfectly well the link between input image and color label in terms of human reasoning, the latter is more widely used. In [6][7][8], Fuzzy Logic was used to classify the pixels into color labels. Firstly, they defined the set of membership functions for each dimension of a color space; those functions are used to fuzzify the input pixels into linguistic terms. After that, a set of predefined inference rules are used to infer the possible output colors. Finally, a technique of defuzzification helps to decide the final color label

associated with the input crisp pixel values. Another interesting case of using Fuzzy logic to recognize the color of wood can be found in [9]. Obviously, Fuzzy logic is easy to use, easy to understand and it follows the way humans perceive and reason, especially in the work of color recognition.

From the above point of view, in this work we take the advantages inspired from the previously quoted papers to develop a fuzzy system for color recognition. In detail, a Sugeno-type fuzzy system [10] is constructed, it takes the average HSV values of a required object as inputs, from which it determines the color of the object in human terms. The choice of Sugeno type comes from the need for calculative simplicity for the system performance in comparison to the nonlinear MIN/MAX operator of Mamdani type. Moreover, it also presents structural properties that allow exact piecewise multi-linear representation which enables the integration of this fuzzy system design in some adaptive or learning strategies to specify the rule base parameters [11]. The selection of HSV color space is due to the fact that it represents colors to human eyes much more intuitively than traditional RGB space. HSV stands for Hue, Saturation and Value; the first represents the tone of color and the second describes the amount of color while the last denotes the lightness.

There is no doubt that symbolic techniques are very capable of taking human perception in to account which provides more intuitive recognition. However, this characteristic sometimes introduces uncertainties and hesitation which are reasonable but should be considered to be resolved. In addition, imprecise measurements that come from working under non-ideal conditions may lead to a reduction in recognition quality, for example the changing of lighting conditions, unreliability of sensors, the similarity between some near colors, viewing angles of the cameras, or the other limitations. In this context, more sensors must be employed to overcome these difficulties, and our idea is to add multiple cameras to provide more exact information about the color to be recognized. From that, Evidence theory (or Dempster-Shafer theory) [12][13][14][15] is considered as a solution for the information fusion between cameras in order to reduce conflicts and uncertainties. In this paper, without losing generality and for simplicity, we add another displaced IP camera as the second information source.

The paper is organized as follows. Firstly, section 2 introduces our scenario of application in NAO robot. After that, section 3 and section 4 explains the use of Fuzzy logic for color recognition and Evidence theory for multiple cameras fusing, respectively. Section 5 shows experimentation results and finally section 6 gives conclusion and some perspectives.

## II. APPLICATION IN ROBOTICS SCENARIO

In our application, the robot receives oral commands from humans which require it to find an object with a specific color described linguistically. For the concentration on the color recognition, the target objects chosen are ball-shaped. The mission of the robot is to walk to find the ball with correct color, and respond its results to humans.

### A. Development Platform

The robot platform we chose for the implementation is a humanoid NAO robot which was developed by the Aldebaran company. Interestingly, this robot is equipped with several sensors which allow it to sense the surrounding environment. In this application, the robot's microphones and a camera on its

head are activated to be used for human voice recognition and images processing, respectively; the tactile sensor on its head and a bumper on its foot are used for humans to start/stop the application. Additionally, 25 degrees of freedom (DOF) helps NAO perform many motion tasks, and we use its two legs to move in search of the ball, its hands are used to point to the target ball it has found.

Furthermore, as mentioned in the previous section, another IP camera is used as support for the NAO robot in color recognition. In fact, the robot has its own two cameras attached to its head, but they do not have a common point in their viewing angles, so we cannot capture the same object at the same time, therefore in this work, we attached another IP camera which is connected to one camera of NAO through a local wireless network and use these two. This additional camera is located close to the robot such that the object is seen by all cameras.

Thanks to the advice from manufacturer, we chose C++ for the application development in order to gain better performance, because it is the native language of NAO. For vision processing, OpenCV 2.4.9 is employed to process images captured from the cameras. The speech recognition and movement of the robot are done using libraries provided by the manufacturer.

Fig. 1 shows the schema of the application where NAO's camera and the IP camera use Fuzzy logic and Evidence theory to search for the colored ball demanded. The process of the scenario is described in the next subsection.

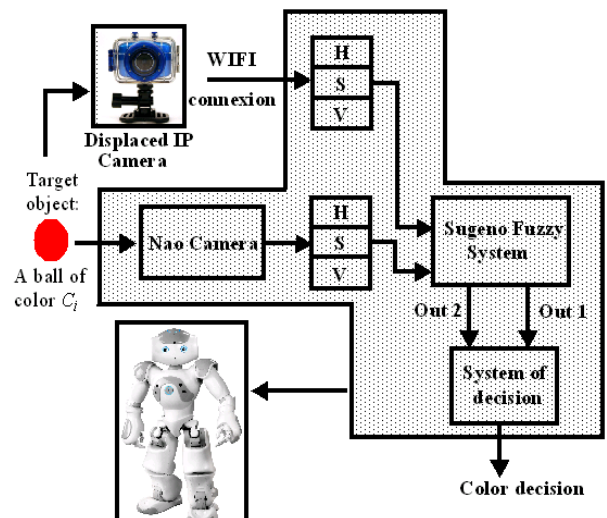


Fig. 1. Color recognition schema for NAO robot.

### B. Scenario Description

Firstly, an interactive color learning is performed. Starting from an initial knowledge, i.e. the threshold of membership function, the robot interacts with the human to acquire the human expertise of color perception.

Then the process of the scenario described below is performed (see the associated states transition diagram in Fig. 2.). the NAO robot comes to the "idle" state when it starts; this is because we may have several applications installed in the robot, so there is a need to have a way of starting for each application. In our case, we touch the tactile sensor on its head

or the bumper on the right foot of NAO to start/stop the application of colored object searching. When it is touched, it will wait for the oral commands from a human. The commands are in the following form:

“NAO, find the  $C_i$  ball”

In this case,  $C_i$  is the demanded color in the set of linguistic labels  $C = \{Blue, Purple, Pink, Red, Brown, Orange, Yellow, Green, Cyan\}$ ,  $i = 1, 2, \dots, 9$ .

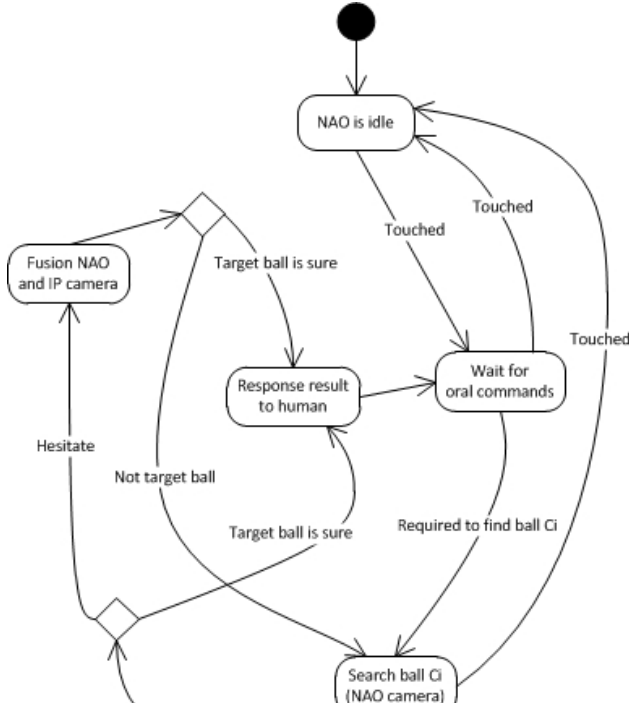


Fig. 2. States transition diagram of colored objects searching.

When NAO recognizes an oral command from a human, it starts searching for the desired colored ball by walking around and uses its camera to scan the surrounding environment. Hough transformation technique [16] is used to detect the balls from captured images and NAO will decide on the detected ones. When it is sure that a ball has the desired color, it will stop moving and says that it has found the target ball. In the case where NAO hesitates about the color of a detected ball, for example it is not sure that the detected ball is yellow or green (because these two colors are sometimes quite similar due to the change of lighting conditions), it will call the connected IP camera to make the fusion between the two cameras. After the fusion, if NAO decides that it is the target ball, it will respond to the human, otherwise, it continues its search. During the finding process, we can stop NAO by touching the tactile sensor or its bumper.

### III. FUZZY LOGIC FOR COLOR RECOGNITION

#### C. Fuzzy System Structure

As introduced previously, in this work we propose to use Fuzzy logic to recognize colors based on human perception. A Sugeno-type fuzzy system which takes average HSV values of the detected ball as input is constructed. Indeed, due to the difficulty of obtain an exact input-output mathematical model between the triplet  $\{H, S, V\}$  and the output color, the rules of

fuzzy system are based on human observation and assessment. An expert describes the crisp values of H, S and V using human terms and he defines the output color label associated with each combination of  $\{H, S, V\}$ , this association can also be called inference rule. Since we have 7 linguistic labels for component H, and 3 for each of S and V, so we get the total number of inference rules is 63. Fig. [3-5] illustrate the partitioning of the input values into linguistic terms and table 1 shows some examples of inference rules which generate the output color “pink”.

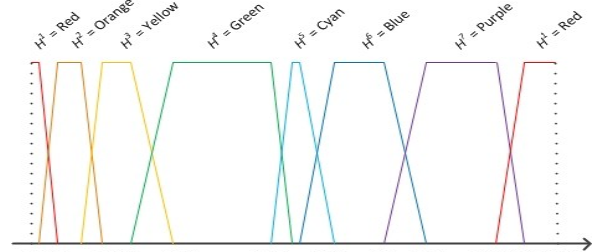


Fig. 3. Partitions / membership functions for component Hue (H).

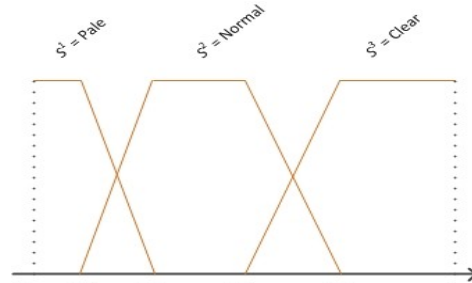


Fig. 4. Partitions / membership functions for component Saturation (S).

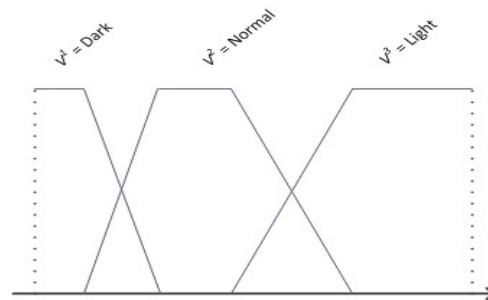


Fig. 5. Partitions / membership functions for component Value (V).

Fig. 6.

Fig. 7.

TABLE I. INFERENCE RULES FOR PINK COLOR

H	S	V	Color ( $C_i$ )
Red	Pale	Light	Pink
Red	Normal	Light	
Purple	Pale	Light	
Purple	Normal	Light	

Our fuzzy system is formalized as zero-order Sugeno-type with the set of inference rules in the following form:

$$R^{(i_1, i_2, i_3)} : \text{if } H \text{ is } H^{i_1} \wedge S \text{ is } S^{i_2} \wedge V \text{ is } V^{i_3} \\ \text{then } C = C_i(1)$$

In that,  $H^{i_1}$ ,  $S^{i_2}$  and  $V^{i_3}$  are the linguistic values in the set of predefined labels for the input  $H$ ,  $S$  and  $V$ , respectively.  $C_i$  is a numerical constant assigned to the output  $C$  of the rule indexed by  $(i_1, i_2, i_3)$  as shown in the table 2. In this work, we use nine chromatic colors for the system outputs, i.e.  $\{C_i\}_{i=1, \dots, 9} = \{\text{Blue, Purple, Pink, Red, Brown, Orange, Yellow, Green, Cyan}\}$ .

Finally, the output which decides the detected color is generated by the following formula:

$$C = \frac{\sum_{(i_1, i_2, i_3) \in I} \zeta^{(i_1, i_2, i_3)}(H, S, V) \cdot C_i}{\sum_{(i_1, i_2, i_3) \in I} \zeta^{(i_1, i_2, i_3)}(H, S, V)} \quad (2)$$

$\zeta^{(i_1, i_2, i_3)}$  represents the truth value of the premises of the rules and  $I = I_1 \times I_2 \times I_3$  indicates the set of labels representing the rules base, with:

$$i_1 \in I_1 = \{1 \dots 7\},$$

$$i_2 \in I_2 = \{1, 2, 3\},$$

$$\text{and } i_3 \in I_3 = \{1, 2, 3\}.$$

In fact, the chromatic output colors in this work were chosen and assigned the constant numbers according mostly to their order appearing intuitively to human perception, which was not addressed carefully in most of the Fuzzy-based previously quoted works. This ways helps our fuzzy output more reasonable because it avoids unexpected colors as outputs after activating some inference rules that are not related. For example, the combination of only Purple color and Red color cannot be a Green one, so the Green color should not be assigned a number between the ones of Red and Purple.

TABLE II. CONSTANT NUMBER FOR COLORS

Color $C_i$	Color Number
Blue	1
Purple	2
Pink	3
Red	4
Brown	5
Orange	6
Yellow	7
Green	8
Cyan	9

In addition, an exceptional case in this work is the first and the last color, they are the Blue and the Cyan, and as human perception perceives, they are the two colors that near together (like in the Hue colors circle). In the normal process of Sugeno fuzzy system, if there are only conflicts between these two colors (with their code is 1 and 9, respectively), the fuzzy system will result a number between 1 and 9, for instance 3, 4 or 5... and it is not good in the human perception point of view. In order to solve this problem, we consider the list of possible output colors as a circular manner as illustrated in Fig. 6. We add a virtual number "10" representing for the Blue color if there are conflicts between it and the Cyan color, after that the Sugeno output will be between these two colors only. By this way, we can guarantee the right input-output behavior of the Sugeno fuzzy system according to the human evaluation. Generally, more virtual numbers can be added if there are conflicts between the two first and the two last colors, but from the rule base in this work, we found that using only one virtual number is sufficient.

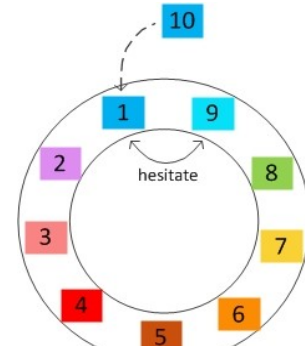


Fig. 8. Illustration for the problem of the first and the last color.

#### D. Certainty and Uncertainty of Fuzzy-based Recognition

The output colors in the work are masked by integer values (table 2), but the Sugeno results of the fuzzy system are real numbers (Equation 2), therefore we need to synthesize these real values to the "nearest" constant integer values that specify linguistic colors, so we define a threshold  $\varepsilon$  for doing that, as illustrated in the Fig. 7

$$\text{If } C \in [C_i - \varepsilon, C_i + \varepsilon] \text{ then the color for } C \text{ is } C_i. \quad (3)$$

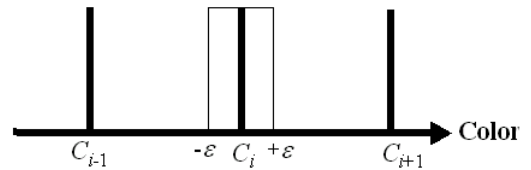


Fig. 9. Threshold for the output of recognition.

Interestingly, the interval  $[C_i - \varepsilon, C_i + \varepsilon]$  can be considered as the sureness interval in which the robot believes that the detected color is  $C_i$ . The value of threshold can affect to the rate of detection and a compromise between lower and higher values of its need to be addressed carefully. In fact, a low value of threshold can lead to reduction in rate of detection, but it increases the uncertainty interval in which the robot hesitates between neighbor colors. On the contrary, a large value can increase detection rate, but the uncertainty

interval is reduced. However, the uncertainty is necessary because it reflexes correctly what humans normally do in reality. Table 3 shows the influence of threshold values on the detection rate with the Red color.

TABLE III. INFLUENCE OF THRESHOLD VALUES ON DETECTION RATE OF RED COLOR

Threshold $\epsilon$	Detection rate	Conflict and other situations
0.1	61%	39%
0.15	64%	36%
0.2	65%	35%
0.3	74%	26%
0.4	80%	20%

In this work, human perception and reasoning is important, so a low value of threshold is preferred, and we chose a value of 0.15 for this application. In order to solve the problem of uncertainties coming from imprecise measurement, operating conditions and affection of threshold value, we use Evidence theory for multiple cameras, which is explained in the next section.

#### IV. EVIDENCE THEORY FOR RECOGNITION IMPROVEMENT

##### E. Fusion Context

Particularly, after a lot of experiments, we have found that in most of the tests, the imprecision of recognition using Fuzzy logic is  $\pm 1$  color. In another word, if  $C_i$  is the required color to be recognized, then the results are usually in the interval  $[C_{i-1}, C_{i+1}]$ . For the rarely cases outside of this interval, we suppose not considering in the interests. Therefore, given a demanded color  $C_i$ , we can propose the following space of discernment:

$$\Omega_i = \{C_{i-1}, C_i, C_{i+1}\} \quad (4)$$

So the power set which contains all possible hypotheses can be constructed:

$$P_i = 2^{\Omega_i} = \{ \emptyset, \{C_{i-1}\}, \{C_i\}, \{C_{i-1}, C_i\}, \{C_{i+1}\}, \{C_{i-1}, C_{i+1}\}, \{C_i, C_{i+1}\}, \{\Omega_i\} \} \quad (5)$$

The three main stages used for the decision system are illustrated in Fig. 8. We firstly construct the mass vectors from the output of Sugeno system, then these masses are combined using a specific combination operator, finally decision output is provided by using a decision method. The construction of masses is explained in the next subsection.

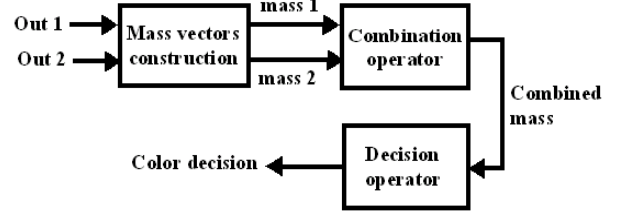


Fig. 10. Decision strategy principle.

##### F. Fuzzification-based Masses Construction

In order to describe the belief about some hypotheses of each information source, we need to build a mass vector which is described by the following function:

$$m_{camera}^i: 2^{\Omega_i} \rightarrow [0, 1],$$

$$\sum_{X \in 2^{\Omega_i}} m_{camera}^i(X) = 1,$$

$$m_{camera}^i(\emptyset) = 0 \quad (6)$$

In the detail of this work, each numerical output of Sugeno fuzzy system is fuzzified as illustrated in Fig. 9. The range of each hypothesis is described by a membership function, and the membership degree of each Sugeno output to each label describes the degree of belief of the source about that hypothesis.

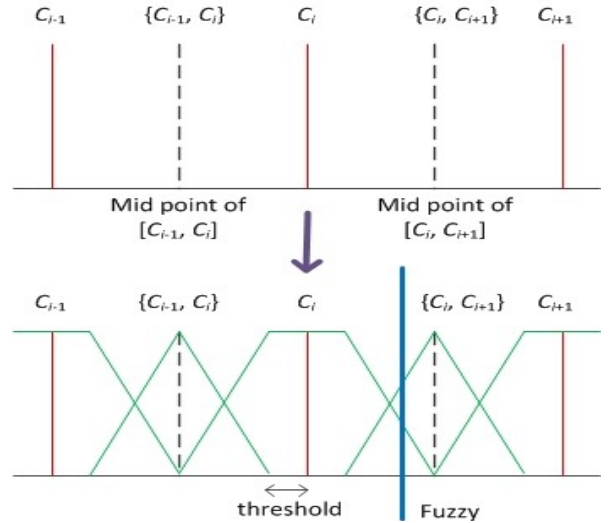


Fig. 11. Fuzzification-based masses construction strategy.

Obviously, due to the natural order of the colors' constants numbers proposed in this work, we always have:

$$m_{camera}^i(\{C_{i-1}, C_{i+1}\}) = 0,$$

$$m_{camera}^i(\{\Omega_i\}) = 0 \quad (7)$$

##### G. Selection Of Combination Method

After constructing the mass vector for each camera from Sugeno fuzzy outputs, the next step is to combine those masses in some ways that integrate all the information provided. There

have been several works proposing combination operators. Initially, this kind of rule was introduced by Dempster, and then recovered by Shafer as the normalized conjunctive rule [17]. This normalization considers the total conflict and is thus interesting only in closed world for the combination of relatively reliable sources. In addition, Florea [14] proposes a rule based on a weighted sum between the conjunctive and disjunctive rule according to the total conflict. On the other hand, Yager [15] distributes the total conflict on total ignorance in order to remain in closed world and suppose that we do not know anything in conflict situations. Additionally, Dubois and Prade [13] proposed a powerful rule considering the cases where information from the source is stated on singletons. Indeed, this method distributes the partial conflicts into partial ignorance, and is more suitable with unreliable sources.

In this work, we have implemented all of the above combination strategies, and found that the one proposed by Dempster-Shafer (DS) adapts well the recognition quality requirement. The related results are not detailed in this paper. For each demanded color  $C_i$  and for each camera, we have the combined mass defined by the following:

$$(8)$$

The parameter  $k$  represents the conflict between the  $s$  sources, and  $Y_j$  is the information given by the source  $j$ .

#### H. Selection of Decision Method

The final step before giving a final output is to choose the method of decision, this is the way we consider all the combined information and choose an output that is the most appropriate. Since the maximum of belief is too pessimistic, and the maximum of plausibility is too optimistic, we consider using the maximum of probability pignistic, which was proposed by Smets [18], as the decision method.

For each color  $C_i$  and for each camera, the decision is the singleton that maximizes the following function:

$$(9)$$

$|A|$  represents the cardinality of a subset  $A \subset \Omega_i$ .

#### I. Mass Improvement by Considering Reliabilities

Considering the low reliability of each camera: the embedded Nao camera and a low cost IP camera it is interesting to take into account this reliability factor. These reliabilities characterize the fuzzy-based recognition quality of each camera with each color. In order to obtain those reliabilities, hundred tests for each color were carried under different lighting conditions; finally the detection rates are chosen as the reliabilities applied in the application.

These reliabilities can be used as “discounting factors” which are used to decrease the degree of belief from sources and consider moving a part of belief into the set of ignorance, from that a more reasonable decision can be obtained.

Let  $\alpha_i$  be the discounting factor for the color  $C_i$  with a camera. Now we define the new mass vectors as the following:

$$(10)$$

In that,  $m_{\alpha_i}^i(X)$  is the mass value of the color for the hypothesis  $X$  after discounted by the factor  $\alpha_i$ .

#### J. Illustrative Example

In order to illustrate the methodology discussed above, we will take an example in which the NAO robot is demanded to find a ball having yellow color (so we have  $C_i = 7$ ). During the searching process, both NAO robot and IP camera find the same ball, but give the two different Sugeno results:

$$\begin{cases} \text{Sugeno}_{NAO} = 6.480 \\ \text{Sugeno}_{IP} = 7.567 \end{cases}$$

From these outputs, NAO robot perceives the color of the detected ball as more orange than yellow ( $C_{i-1} = 6$ ), and the IP camera considers more green than yellow ( $C_{i+1} = 8$ ). In this case, it is difficult to take a decision with each camera separately. By using the Fuzzification-based masses construction, we are able to build mass vectors from those Sugeno outputs. The reliability of NAO and IP camera in this case are respectively 0.32 and 0.43. The conflict  $k$  between sources before considering reliabilities is measured as 0.238. The threshold value chosen is 0.15.

TABLE IV. EXAMPLE OF MASSES CONSTRUCTION AND FUSION

	$m_{NAO}$	$m_{IP}$	$(m_{NAO})_R$	$(m_{IP})_R$	$m_{DS}$	PP
$\{\emptyset\}$	0.00	0.00	0.00	0.00	0.00	
$\{6\}$	0.06	0.00	0.02	0.00	0.01	0.233
$\{7\}$	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.11</b>	<b>0.453</b>
$\{6, 7\}$	0.94	0.00	0.30	0.00	0.18	
$\{8\}$	0.00	0.19	0.00	0.08	0.06	0.314
$\{6, 8\}$	0.00	0.00	0.00	0.00	0.00	
$\{7, 8\}$	0.00	0.81	0.00	0.35	0.24	
$\{6, 7, 8\}$	0.00	0.00	0.68	0.57	0.40	
<b>Decision</b>	<b>Yellow (7)</b>					

In the table 4, we can see:

- The mass vector of each camera without reliability  $m_{NAO}$  and  $m_{IP}$ .
- The mass vectors after considering reliabilities  $(m_{NAO})_R$  and  $(m_{IP})_R$ .
- The combined mass vector using DS operator  $m_{DS}$ .
- The decision given by maximum of probability pignistic PP.

According to the table, the robot knows that he should consider the detected ball as a yellow one, and says out his result to the human.

## V. EXPERIMENTATION RESULTS

The table 5 shows our results of experimentation with the threshold value of 0.15. The expertise of the human perception of each color is acquired as follow. 100 images are acquired by NAO’s camera and the IP camera. Reminding that with such

low threshold, the rate of detection will be decreased; however it increases the cases of uncertainties and conflicts, which are then resolved well by evidence theory to give more reasonable results. In detail, the first column lists all the possible colors considered in this work, the second and the third column shows the detection rate of NAO and IP camera, respectively, with using only Fuzzy logic separately. Finally, the last column is the result of using Evidence theory to fuse information of the two cameras. The last row is the average values of the upper ones, which draws an important improved quality by using fusion technique (>30% improvement in comparison to the use of fuzzy logic for a single source).

It is also worthy to note here that the cases of unsuccessful detection are due to essentially the presence of total conflicts between sources, or there are some fuzzy outputs that fall outside the space of discernment assumed previously. For these cases, we can handle by using other combination operators, or considering other masses construction method, for example from a learning base, which will be taken into account in the future work.

TABLE V. EXPERIMENTATION RESULT

Color	Sugeno Fuzzy System (NAO)	Sugeno Fuzzy System (IP)	Fusion of two cameras
Blue	98%	38%	100%
Purple	90%	94%	100%
Pink	34%	38%	72%
Red	64%	48%	90%
Brown	66%	48%	88%
Orange	62%	72%	100%
Yellow	32%	43%	86%
Green	59%	58%	77%
Cyan	26%	23%	98%
<b>Avg.</b>	<b>59.00%</b>	<b>51.33%</b>	<b>90.11%</b>

## VI. CONCLUSION

This paper discusses about the importance of human understanding in humans system interaction. An illustrated scenario was implemented in the case of a humanoid NAO robot which is required to find a target colored object by oral commands from humans. A Sugeno-typed fuzzy system is constructed to help for the color recognition more intuitively, after that Evidence theory is employed to fuse information from multiple cameras in order to improve the recognition quality. Experimentation results shows that the combination of the two methods gives a good color recognition based on human perception and reasoning.

The future scenario of the work will consider the case of 3D objects recognition for NAO robot following human

commands. We also consider using fuzzy and gradual intervals arithmetic [19][20]. Moreover, Choquet integral may be used to determine relevant cameras reliabilities according to the interaction between colors.

## REFERENCES

- [1] F. Liu, H. Lu, and Z. Zheng. "A modified color look-up table segmentation method for robot soccer." Proceedings of the 4th IEEE LARS/COMRob 7 (2007).
- [2] T. Kikuchi et al. "Improvement of color recognition using colored objects." RoboCup 2005: Robot Soccer World Cup IX. Springer Berlin Heidelberg, 2006. 537-544.
- [3] S.H. Tsai, and Y.H. Tseng. "A novel color detection method based on HSL color space for robotic soccer competition." Computers & Mathematics with Applications 64.5 (2012): 1291-1300.
- [4] R. Szabo and I. Lie. "Automated colored object sorting application for robotic arms." Electronics and Telecommunications (ISETC), 2012 10th International Symposium on. IEEE, 2012.
- [5] E. Yoshida et al. "'Give me the purple ball'-he said to HRP-2 N. 14." Humanoid Robots, 2007 7th IEEE-RAS International Conference on. IEEE, 2007.
- [6] L. Shamir, Human Perception-based Color Segmentation Using Fuzzy Logic. IPCV 2, pp. 96-502, 2006.
- [7] O. Kucuktunc and D. Zamilieva, Fuzzy color histogram-based CBIR system. Proceedings of 1st International Fuzzy Systems Symposium, 2009.
- [8] S.G. Lee, K.B. Kim, E.Y. Cha. Color Inference Using an Enhanced Fuzzy Method. Red (R), Vol. 330, 2013.
- [9] V. Bombardier and E. Schmitt. "Fuzzy rule classifier: Capability for generalization in wood color recognition." Engineering Applications of Artificial Intelligence 23.6 (2010): 978-988.
- [10] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man and Cybernetics, Vol. 1, pp. 116-132, 1985.
- [11] S. Galichet, R. Boukezzoula and L. Foulloy, Explicit analytical formulation and exact inversion of decomposable fuzzy systems with singleton consequents. Fuzzy Sets and Systems, Vol. 146, N° 3 pp. 421-436, 2004.
- [12] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping. The annals of mathematical statistics, pp. 325-339, 1967.
- [13] D. Dubois and H. Prade, Representation and combination of uncertainty with belief functions and possibility measures. Computational Intelligence Vol. 4, N° 3, pp. 244-264, 1988.
- [14] M.C. Florea, Combinaison d'informations hétérogènes dans le cadre unificateur des ensembles aléatoires: approximations et robustesse. Thèse de doctorat. Université Laval, 2007
- [15] R.R. Yager, On the Dempster-Shafer framework and new combination rules. Information sciences, Vol. 41, N° 2, pp. 93-137, 1987.
- [16] R.O Duda, and P.E. Hart. Use of the Hough transformation to detect lines and curves in pictures. Communications of the ACM 15.1 (1972): 11-15.
- [17] G. Shafer, A mathematical theory of evidence., Vol. 1. Princeton: Princeton university press, 1976.
- [18] Ph. Smets, Constructing the Pignistic Probability Function in a Context of Uncertainty. UAI. Vol. 89. 1989.
- [19] R. Boukezzoula, S. Galichet, and A. Bissierier, A Midpoint-Radius approach to regression with interval data. International Journal of Approximate Reasoning Vol. 52, N° 9, pp. 1257-1271, 2011.
- [20] R. Boukezzoula, S. Galichet, L. Foulloy and M. Elmasry, Extended gradual interval (EGI) arithmetic and its application to gradual weighted averages. Fuzzy Sets and Systems Vol. 257, pp. 67-84, 2014.