Learning Human-Like Color Categorization through Interaction

Rinaldo Christian Tanumara, Ming Xie, and Chi Kit Au

Abstract—Human perceives color in categories, which may be identified using color name such as red, blue, etc. The categorization is unique for each human being. However despite the individual differences, the categorization is shared among members in society. This allows communication among them, especially when using color name. Sociable robot, to live coexist with human and become part of human society, must also have the shared color categorization, which can be achieved through learning. Many works have been done to enable computer, as brain of robot, to learn color categorization. Most of them rely on modeling of human color perception and mathematical complexities. Differently, in this work, the computer learns color categorization through interaction with humans. This work aims at developing the innate ability of the computer to learn the human-like color categorization. It focuses on the representation of color categorization and how it is built and developed without much mathematical complexity.

Keywords—Color categorization, color learning, machine learning, color naming.

I. INTRODUCTION

COLOR is an important element in visual world. It is widely used as an important cue to identify object, to provide additional information (e.g. using color red to represent danger, to give certain emphasize), etc. Although color is actually the result of electromagnetic wave with different wavelength which is continuous in nature, human perceives color in categories as if they are discrete. The category may be, but not always, indicated by name, such as red, blue, etc.

Each human being has his own unique color categorization [1]. This is because of environment, task, and most importantly physical difference. Interestingly, despite such individual differences, the categorization is shared among members in society. This allows a successful communication among them. For example, the color category “red” is shared in the society, hence when one uses the word “red”, the other knows the color or the object with the color he refers to. This does not mean an absolute identical categorization, but sufficient to achieve a successful communication [2].

Sociable robot [3] is aimed to live coexist with human in human society. It is expected to be a new member in human society, which actively participates in social interaction and becomes coworker of human. To be able to interact effectively with human especially on the ones related to color, not only a social interface between human and robot, but the robot must also have the color categorization shared in the society.

Color categorization in one and another society may be different [4]. Human society is also dynamics, in which color categorization may change and evolve over time. That is why simply computerizing color categorization to the robot is not suitable. The computer, as brain of robot, must learn the color categorization from the society it lives. By which, the computer will arrive at the shared color categorization and able to continuously adapt the categorization. This work does not focus on the social interface for interaction, but on developing the innate capabilities of the computer to learn the human like color categorization.

The rest of this paper is organized as follows. Section 2 describes other research works to achieve color categorization by computer or artificial system. Section 3 describes the color learning through interaction, whereas the result and discussion are presented in Section 4. And finally, Section 5 presents the conclusion of this work.

II. RELATED WORKS

In general, the works to achieve color categorization by computer can be grouped into two approaches. The first approach is based on modeling and simulation of human color perception, whereas the other one focuses on developing new clustering algorithm for color categorization.

For computer to achieve not simply a color categorization, but color categorization as of human, modeling human color perception is certainly one possible and quite intuitive approach. Many computer science and engineering solution has also been found by modeling the nature. In addition, this approach may result in both color categorization by computer as well as more understanding of human color perception itself.

However until now, there are still many unresolved issues in human color perception, including human color categorization. It is agreed that each human has his own unique color categorization which has no rigid boundary and cannot be described by mathematical formulation. However it is not well understood how the categorization is shared among members in society. Moreover although the categorization may not be shared across different society, high degree of
Based on empiricism [8][9], the categorization is the result of solely inductive learning from the environment. The genetic basis is only to govern learning mechanism, whereas social interaction is to learn the names of the already known shared categories. Living in the same environment results in the shared color categorization.

Based on culturalism [10]-[12], the categorization is the result of social interaction. Through feedback during social interaction, the color categorization is made and, at the same time, the color names are learnt. The social interaction is highly affected by the environment. The genetic basis is also necessary to govern learning mechanism. As the members in society interact to each other, the color categorization is build and adapted between one and another, which will eventually stabilize as the shared color categorization.

Various systems have been developed based on those hypotheses. Dowman [13], following nativism, developed fuzzy Bayesian system to simulate color acquisition and the evolution of color categorization. Lammens [14] developed NPP color space based on neuro-physiological data to relate between visual stimuli (color as electromagnetic wave) and color categorization. Perhaps the most interesting and well rounded research in this approach is done by Steels and Belpaeme [2]. They develop various color categorization systems, modeling each of hypotheses to find out the best way of designing artificial system as well as to give comparison among them. For nativism, the color categorization is encoded into color genes and the genetic evolution is simulated. For empiricism, an inductive learning system using adaptive network is developed. For culturalism, they developed a few artificial systems to interact to each other.

Such modeling and simulation are in fact more towards to gain more understanding on human color perception, instead to develop the artificial one. It is because the model and simulation, in general, must also be concurred with other aspects of human such as: brain development, human’s physiology and neurological processing, etc. which is not necessary for artificial system to follow.

Machine intelligence must be differentiated from computerized human intelligence [15]. In the same notion, artificial system does not necessary follow human color perception. The physical changes affecting color categorization, which may be in form of physical evolution between generations or brain development during the lifetime, etc., are not necessary to be implemented or modeled to develop artificial system. Human’s brain and physical architecture which enables color perception including color categorization, are also not necessary to be implemented in artificial system. For example, it is not necessary for artificial system to have visual sensor the same as human’s eye. The artificial system can be subset or superset of human color perception. The mechanism of human color perception may not be the best solution for artificial system, although it may still be inspired from human color perception. In addition, the current knowledge of neural science may be too low level to be modeled for higher level cognition including color perception.

On the other hand, the other approach focuses on developing a new clustering algorithm for color categorization. They do not aim for color categorization shared with human, but for specific application such as image segmentation, object detection, etc. Artificial neural network (ANN) is generally used. It is mainly because of the learning capability, by which the network adapts the input-to-category connection, so that when the input comes, the correct category can be invoked [16]. ANN is also in favor for modeling human color perception as it possesses hidden units between input and output, which can be interpreted as the internal representation of the input [17][18]. Different types of ANN have been experimented. Yin [19] uses RCE neural network for color categorization for hand segmentation, which can be extended for color based image segmentation in general. Yeo [20] uses combination between SOM and ART neural network. Other works, such as [21]-[23] also use ANN for color categorization.

However, relying on mathematical complexity on clustering algorithm alone does not necessarily result in color categorization shared with human. Although the color categorization may be shared among artificial systems and enable them to do certain task very well, it is not suitable for sociable robot.

III. COLOR LEARNING THROUGH INTERACTION

A. Approach

This work does not model and simulate human color perception, and does not rely on mathematics for color categorization. In this work, a new approach is adopted, i.e. by making computer learn the categorization through interaction with other members in the society. Although similar to culturalism, in which color categorization is the result of social interaction, they are actually different. Here the computer simply learns categorization from human without ability to create or modify the category by its own. The human, on the other hand, does not develop any new categorization due to the interaction with computer.

Learning through interaction is indeed highly related to the work on human computer interaction to develop social interface so that human can interact effectively and efficiently with computer. The innate capability of the computer to learn is mostly overlooked, although it is actually more important
than social interaction. For analogy, it is impossible to teach a
dog to speak human natural language. It is not because of the
social interface between human and dog, but because dog does
not have capability to learn human natural language. On the
other hand, each human has his own personality and way of
interaction. Certain type of human has difficulty to socially
interact with others, but it does not mean that he cannot learn
through interaction with others as he still has the capability.
However indeed the learning process may not be as efficient
as if he has better social interaction ability. In this view,
instead of focusing of the social interface, this work focuses
on developing the capability to learn the human-like color
categorization. By this capability, the computer is able to
learn color categorization, as well as to adapt and follow the
evolution of color categorization through interaction.
Therefore although at the beginning, it does not know any
categorization, eventually through interaction, the color
categorization of the computer will reflect the shared color
categorization with humans in the society.

To design this capability, the following are considered: how
color and its categorization are represented in memory and
how this representation is built and developed. Without
accommodating representation, it is impossible for computer
to reach human-like color categorization. For analogy,
although monkey can learn mathematics, it is only to a very
limited extent. This indicates that monkey understands the
concept of number, otherwise monkey will not be able to
perform even a simple number comparison (such as: smaller
number, etc.). However human and monkey understands
different number differently. Or in other way, number is represented
differently in monkey’s brain. And it is impossible to develop
the representation to reach human like knowledge [24]. For
another analogy, human is capable to use natural language
because of special brain mechanism, which other animal does
not have [25]. This does not mean that robot must follow
human’s innate mechanism. Robot may have its own
mechanism which is different from human. The
accommodating representation is crucial, but the mechanism
to build and develop the representation must also be
considered.

The rest of this section is organized as follows. Section B
and C describe representation of color and its categorization
respectively, whereas section D describes how the color
categorization is built through interaction.

B. Perception Space

Perception space defines the colors which can be perceived
and how they are organized. Based on which, the color
categorization is made. That is why perception space is an
important factor to achieve color categorization shared with
human. In this work, CIELAB color space, defined by CIE
(International Lighting Committee), is used. CIELAB color
space is currently the best color space for categorization. Its
performance for categorization is found to be better than NPP
color space developed based on neuro-physiological data
[14]. In addition, CIELAB color space can also handle color
constancy to some extent [2][26].

CIELAB color space has three parameters: L reflecting
human lightness perception, a* and b* reflecting opponent
channel red-green and yellow-blue in human physiology
respectively.

C. Categorization

As for human, the color categorization is evolved and
developed depends on the culture and environment. More
advanced society knows more categorization compared to the
primitive ones [4]. Also, for certain task and environment,
more detail color categorization is necessary. For example, an
expert on soil will know more detail categorization of brown.
Another example, Eskimo knows more categorization of blue.
To handle this issue, the color categorization for the system is
made as hierarchical categorization of the perception space
into color basic representation (CBR), color identity (CI) and
color category (CC).

Color category (CC) refers to categorization in general,
which is commonly identified by color name such as red, blue,
etc. Color identity (CI) refers to a more specific color
categorization. It may be identified by color name such as
crimson, azure, etc. However it is also common that color
identity does not have any name. For example, a certain color
of soil may not have any name, but it is certainly a distinct
category. The cluster(s) of CI will form a color category
(CC).

Color basic representation (CBR) is the lowest level of
categorization. It does not reflect any color categorization in
human, but it reflects the color resolution or perceptual
discrimination, i.e. the minimum color difference that can be
detected. Technically, CBR is also used to reduce the effect
of noise and computational load. CBR are made during the
design stage and uniformly distributed across the perception
space. On the other hand, CI and CC are built during learning
process through interaction with human. CI and CC are non
uniform and may be overlapped between one and another.
The cluster(s) of CBR will form a color identity (CI). In this
work, size of each CBR is 5 × 5 × 5 in L, a* and b* directions.

The hierarchical categorization can be represented by the
neural network shown in Fig. 1.

![Fig. 1 Neural Network for Color Categorization](image-url)

The neural network consists of four layers. The first layer
is the input layer, which receives the L, a* and b* parameters.
The second layer is called representation layer. Each cell in
this layer represent a CBR and identified by a certain range of
The third layer is called identity layer. Each cell in this layer represents a cluster of CBR. This cluster of CBR is called color identity element (CIe). Each CIe may have a conceptual label or name. A group of CIe having the same conceptual label will form a color identity. The conceptual label itself becomes the name for the color identity. If a CIe does not have conceptual label, it will form a CI by its own. The connection between cells in 2nd and 3rd layer represents whether the CBR belong to the particular CIe. There are two parameters in each connector: Y (yes) and N (no) counters, to measure the confidence level (difference between Y and N counters) as well as total number of adjustment of the connection (total of both Y and N counters). A CBR may connect to more than one CIe. At the beginning there is no cell at all in this layer. The cells and connections are built through interaction.

The fourth layer is called category layer. Each cell in this layer represents a cluster of CIe, which is called color category element (CCe). Most of CCe will have a conceptual label or name. A group of CCe having the same conceptual label will form a color category. The conceptual label itself becomes the name for the color category. The connection between cells in 3rd and 4th layer represents whether the CIe belong to the particular CCe. Two parameters in each connector, Y (yes) and N (no) counters are used to measure the confidence level as well as number of adjustment of the connection as in identity layer. A CIe may connect to more than one CCe. Similar to the identity layer, at the beginning there is no cell at all in this layer. The cells and connections are built through interaction.

Both CIe and CCe are the intermediate categorization to achieve CI and CC. In summary, the complete hierarchical categorization can be seen in Fig. 2.

![Fig. 2 Complete Color Hierarchical Categorization](Image)

**D. Interaction**

Through interaction, the computer learns color identity and color category, i.e. by building the identity and category layer of neural network. The interaction is done with a human master at one time using the Graphical User Interface shown in Fig. 3. The mechanism for each interaction is as follows:

1) Human master will show an image to computer. This can be acquired from webcam or image file. From the image, the human master will deliberately select a color to be learnt by computer. The computer will then determine the CBR of the selected color.

2) The computer will show the CC of the selected color having the highest confidence level. The CC is shown by showing the CBR having highest confidence level for each CCe belonging to the particular CC. This category is to be confirmed or denied by human master.

3) If the human master confirms, the particular CC is optimized. Its Y counter is increased, while for other CC, the N counter is increased. This lateral inhibition mechanism helps to minimize wrong or multiple categorization. The computer will then show all CCe belonging to the CC (in similar way as in #2, i.e. by showing the CBR having the highest confidence level for each CCe) as well as all adjacent CCe which do not belong to the particular CC. For each CCe, the human master will confirm or deny whether it belongs to the CC. Through this confirmation or denial, the category can be reduced or expanded incrementally and the counters, which reflect confidence level, are updated.

4) If the human master denies, the computer will show CC having the next highest confidence level. If it is confirmed, it is then optimized in the same way described in #3.

5) If all CC known by computer are denied, a new CCe is made. It consists of only the CIe where the selected color belongs to. If the CIe is also not known, then a new CIe is created as described in #9. The new CIe is then optimized in the same way described in #3. The human master may also give the conceptual label for this CCe.

6) The computer will then show the CI of the selected color having highest confidence level. The CI is shown by showing the CBR having highest confidence level for each CIe belonging to the particular CI. This identity is to be confirmed or denied by human master.

7) If the human master confirms, the particular CI is optimized. Its Y counter is increased, while for other CI, the N counter is increased. The computer will show all CIe belonging to the CI as well as all adjacent CIe which do not belong to the CI. For each CIe, the human master will confirm or deny. Based on which the identity can be reduced or expanded incrementally and the counters on each connection are updated.
8) If the human master denies, the computer will show CI having the next highest confidence level. If it is confirmed, it is then optimized in the same way described in #6.

9) If all CI known by computer are denied, a new CIe is made. It consists of only the CBR where the selected color belongs to. The new CIe is then optimized in the same way described in #6. The human master may also give the conceptual label for this CIe.

Using the above mechanism, the color categorization are built and developed. The color identity is formed of CIe having the same conceptual label, whereas color category is formed of CCe having the same conceptual label. The conceptual space is also developed through this mechanism. Conceptual space is the space in which the computer knows the categorization. Ideally, perceptual space (the space in which the computer is able to perceive) and conceptual space are the same. However, as also happened for human, they can be different. For example, Eskimos have perceptual space like other human being, but their conceptual space may only be developed in blue region. In this system, the integration of all CIe forms the conceptual space for computer.

There is no merging or deletion of cluster, but they can be reduced or expanded. The one with negative confidence level (when N counter is larger than Y counter), will simply be ignored, but it can be revived through much more confirmation from human master. Each category is built incrementally to ensure the appropriate clustering. It may also help human master during interaction in optimizing the CI and CC.

IV. RESULT AND DISCUSSION

After weeks of interaction, the color categorization by computer is compared with human color categorization as shown below.

Fig. 4 shows Munsell chips (as also used in [14]) overlaid by both human and computer color categorization. The data of color categorization of computer is obtained after a few weeks of interaction. The categorization is indicated by a rectangular on top of the Munsell chips. The size of the rectangular reflects the confidence level within the same category, but not across categories. The category with a very small or negative confidence level is not shown. Also, if there is more than one category for the color chip, only the most dominant ones are shown. There are 11 color categories, which consists of 64 CIe (not shown in the figure).

On the other hand, the data of color categorization of human is obtained from the survey of Berlin and Kay [4]. In this figure, the boundary of each category is represented by a bold border, while the approximate location of color foci for each category is represented by an oval shape. As can be clearly seen, although categorization by machine looks wider, the overall categorization is sufficiently good. There is only a little mismatch categorization between human and computer. The comparison is done only in color category level, as there is no much data on color identity of human. However as the mechanism is quite similar in developing color category and color identity, the analysis can also be valid for color identity.

Color foci, which are also known as color prototypes, are defined as the best representation for each category. As for human, it is argued that color foci are innate due to physical characteristic of humans and does not much affected by environment and social interaction [5][27]. However large variation of color foci [28][29] is also found. In this work, color foci are not considered during the design of the system. Interestingly, from this result it is found that the color with highest confidence level by computer is almost the same as the color foci from Berlin and Kay’s data. This indicates that it is not necessary to consider whole aspects of human color perception as they may also be obtained as the result of interaction.
This is happened because for color foci, the number of
denial is very little and only occurred when the number of
interaction is very small. As the number of interaction
increases and the correct categorization is known, human
master will always confirm it, which results in very high
confidence level. It is different with the color in boundary of
categorization. In the boundary, each human master may have
his own opinion. Hence the number of confirmation and
denial may be both increased as the number of interaction
increases. However color foci do not always have the highest
confidence level. If the system rarely exposed to the color
foci, but mostly to another unambiguous color in the category,
confidence level for the color foci is not the highest, such as in
orange.

The color categorization result can also be seen from the
following image segmentation result. The image
segmentation is done without any pre-processing and post-
processing. The computer simply determines the category of
color of each pixel in the image. In those figures, the left part
shows the original image, whereas the right part shows the
segmentation result.

![Fig. 5 Image Segmentation Result #1](image1)

![Fig. 6 Image Segmentation Result #2](image2)
As can be seen in Fig. 5, Fig. 6, and Fig. 7, the segmentation is performed quite well despite of overlapping in color boundary and quite number of small regions especially in Fig. 6 and Fig. 7. There is no apparent wrong categorization such as red being categorized as green. The overlapping occurs only in the color boundary where each human master may have his own opinion. In those segmentation results, all color categories as long as it has positive confidence level are shown, not only the one with highest confidence level. Post processing may be necessary to improve the segmentation result.

Through confirmation and denial by human, the computer learns and, at the same time, mimics the color categorization from human. This makes computer have the shared color categorization. However the performance of the system highly depends on variation of exposed color, where the system may only be exposed to certain color all the time, and error made by human either intentionally or not. One possible way to improve is by implementing social status role during interaction. The one with higher social status will have more influence compared to the lower one. This may help to reduce the effect of human error.

The learning process is quite slow, however after the learning process all the data can be copied to another computer. Hence for the new machine, it is not necessary to learn from scratch. It can use the already learnt categorization by other machine as the basis of further learning. Not only for color categorization, but this may also be valid for machine intelligence in general.

Although the computer is able to have the shared color categorization, it is still far to be implemented for color vision in sociable robot. Shared color categorization is not sufficient. Sociable robot, to live coexist with human in real environment, will face various lighting condition and surface context which affects the color dramatically [30]. Preprocessing color for the basis of categorization is necessary. In addition, as sociable robot will not only handle color, the modality with other features and sensors must also be taken into consideration for the innate capability of sociable robot.

V. CONCLUSION

In this paper, the work to enable computer to learn color and have the shared human color categorization through interaction is presented. This work focuses on the learning capability of robot, instead of the social interface between human and robot. This learning capability is designed by considering how the color and its categorization are represented and how this representation is being developed through interaction. The result of this work shows that, without following human color perception and much mathematical complexity, the computer is able to learn color and eventually arrive at the shared human color categorization.
Color categorization by computer, even for sociable robot to have human like color categorization, is not necessary to be constrained by human color perception. Machine color perception is different with human color perception. And interaction may serve as bridge to achieve human-like knowledge and intelligence by machine. The overall system can still be improved further, including the social interface.

REFERENCES