Multi-Approach Integrated To Lumen Robot Friend On A Mission In For Emergency Respond

A Sukoco[1], Marzuki[2]

1. Computer Science Faculty, Bandar Lampung University, Lampung, Indonesia, agussukoco16@gmailcom
2. Computer Science Faculty, Bandar Lampung University, Lampung, Indonesia, marzuki@ubl.ac.id

Abstract— As a social robot, a mission to find and locate objects in for evacuated of disaseter is the ability to be possessed by Lumen Robot Friend (LRF). The robot can easily become a friend and help human's activity especially for emergency respond. Complexity in an uncertain environment from emergency, LRF is difficult to find the object that requires a scenario that involves a lot of approaches. LRF's mission is the process of completion of the mission for helping human's activity, through the incorporation the capabilities of Face Recognition, Sign and Text recognition, gesture recognition, and object recognition of LRF in an uncertain environment, then the goals can be achieved. The focus of this research is to develop the ability to recognize the LRF, understand commands through gesture and text, search for the path based on directions, and determine a rational action and finally the object can be founded and given to the right person. Keywords— NAO, Intellegence Agent, Machine Vission

1. Introduction

Human-Robot Interaction is believed in capabling of producing a high impact when a robot is able to communicate a positive and interact well with humans [1], interactions which may be done for example talk, play, walk, or as an assistant. Lumen Robot Friend (LRF) is a development project "Robot Friend" by LSKK-ITB Research that can move (act like a human being) using NAO platform. Therefore, LRF must have a social interface that includes all social attributes then the robot easily become friends and can help human's activities. In this project, one of main points of this research is the ability LRF in finding objects and give them to someone who addressed through the exploration of environment by vision.

How LRF can complete its mission to emergency respond in helping human activity, through the ability Face Recognition [2][3][4][5], Sign and Text recognition [7][8][9][10][11][12] as well as object detection is a challenging activity. The approach that is taken in the development of the LRF is implementing agent delibertarive coherence driven Architecture [13][14]. This paper is presented in several sections include: Introduction in section I, Related Research in section II, Scenario and Experiment we explain in section III and Conclusion in section IV.

2. RELATED RESEARCH

2.1 Face Recognition With Viola Jones and PCA

Image source originating from various types, one comes from the video, especially face recognition, thus becoming one of its own science [12], [13] with regard to identify or verify a face [12]. Over the last few years, it has been facial recognition receiving significant attention as one of the important applications of image understanding and analysis. Many algorithms have been implemented in a variety of static and non-static conditions. Non static conditions includes positioning, hair etc. All these factors are impact to the process of facial recognition. Various things determine the quality of a visual face recogniton one thing that is determined by its environmental performance [12]. This paper uses Viola Jones to Face Detection and Principal Components Analysis (PCA) [5], [4], [3], [14], [15]. The methodology that is used in this Principle Component Analysis (PCA) is Eigenfaces, obtained citri reduce the dimension of the object so that it can take any important characteristics of an object. Then eigenfaces are using PCA in reducing its dimensions with the goal of finding vectors to obtain

the best value for the distribution of face images in the input image space [13]. It means centered training images is calculated by subtracting the mean normalized image is computed. If W is the matrix of mean-centered training image Wi (i = 1, 2, ... L) and L is the number of training images, the covariance matrix D is computed from W as in Equation 1 and 2 [13], In equation 2, zi is represented as a vector with the new features that have been reduced and acquired new sub-space. The result can be obtained intra or inter class scatttering scattering. for classification are considered either intra scattering class.

2.2 Text Recognition

One current trend of the text detection and recognition, with the increasing development is appropriate from the results of the survey paper [16], meanwhile some examples of the scope of research [17]: (1) image understanding, (2) image seach, (3) the target geolocation, (4) and robot navigation of the four topics is closely related to the detection and the introduction of text found in the area. The methodology of a study area consists of Text Detection and character recognition, the first stage is (1) to extract an object from an image based features such as character alignment of the object (2) Extraction by the edge of the effective techniques to localize and extract characters from the image by finding the vertical edges (3) Since rectangular shape has an aspect ratio of unknown property, extract the object by finding all the possible rectangles on a given input image, the second stage is to recognize the character extracted by using contour analysis after finding the contours of the segmented character. An introduction by contour analysis system consists of two phases, namely phase detection (detection) and introduction phases (Recognize). Contour analysis which consists of the extraction of contour (contour extraction) and the vector representation (vector representation) are applied to the segmented character. Then the character can be recognized as matching the character (character matching).

2.3 Sign Recognition

Methods of detection and recognition of traffic signs are used by features colors and shapes of traffic signs. This method consists of image segmentation based on HSV color space (Hue, Saturation, Value), HSV, the color space which is suitable to identify basic colors, where the base color is used in research as a color identification robot. In addition, HSV tolerances to change in light intensity. This is one of the advantages of HSV color space rather than the color segmentation that is done by using the HSV color space together with the form by filtering through template matching candidate detected color is used to detect signs of the backgrounds. Many references can be used to do this research as is deemed to be able to represent the same color as the human eye concept in view of a color [11],[12],[21]. Phase Detection and normalization consists of: (a) The transformation of the image from RGB color space HSV color chamber can be done by the following equation:

$$S = \frac{Max(R, G, B) - Min(R, G, B)}{Max(R, G, B)},$$

$$V = Max(R, G, B),$$

$$H' = \cos^{-1}\left\{\frac{0,5[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}}\right\}$$

$$H = \begin{cases} \dfrac{60(G - B)}{V - min(R, G, B)} & jika\ V = R \\ 120 + \dfrac{60(B - R)}{V - min(R, G, B)} & jika\ V = G \\ 240 + \dfrac{60(R - G)}{V - min(R, G, B)} & jika\ V = B \end{cases}$$

From the results of the HSV color space to the next step (b) shape sign detection based filtering approach features, cross-correlation in this case using equation 1 as follows:

$$d_{f,t}^2 = \sum (f(x,y) - t(x-u, y-v))^2 \dots\dots\dots\dots(1)$$

$$C(u,v) = \sum (f(x,y) - t(x-u, y-v))\dots\dots\dots(2)$$

At the moment, the cementation of color is associated with a different template, if the similarity factor is larger than the threshold value is detected as a sign. (c) Recognition stage approach Support Vector Machine (SVM).

2.4 Object Recognition

The introduction of the object that is used in this study is a speeded Up Robust Feature (SURF), it is an algorithm used by linking interest interconnected point between two images. SURF process begins with the detection of interest points consist of an integral process of counting the number of image intensity in the rectangular region, and then find the interest point using the Hessian matrix and represented by small or large scale. Once the interest points found in the image object and then did the localization interest point. The interesting extraction point that requires in an estimation of the relative motion of the object. The second process is to find an interest point descriptor and matching between two pieces of image objects. This process determines the orientation of the object image rotations, interest point descriptor is obtained by summing the Haar wavelet Responses. Having obtained the interest point descriptor then the Indexing and matching process can be done quickly [22].

2.5 Gesture Recognition

Skin detector is a classifier on the color space that is representative in defining the binary decision ang in determining the classification of skin pixel pixel or not. Where the result is a binary image then do the segmentation of the binary image. If it found non-significant regions of elimination, it means it is carried out on the section using Gabour Filters Cascade. After the elimination process is done then discovered the location of the area of skin significantly[23].

2.6 Adaptive Decision Making Process

How does a machine able to see and identify and act [24], this is called the agent. By detecting and recognizing objects other thing that should be done by considering various things in accordance with its goal [14], in another problem is how a machine capable of making decisions by considering a neat / deliberative coherence in accordance with a dynamic environment to act via actuators, taking into account the existing neatly by evaluating their missions. These things made from this paper is able to determine an implementation using objects and represents its form, read the instructions and identify the person that is going to be given things. Agent decision making is a process that is selected from the existing face recognition, the conditions around, set of several alternatives, a strategy that has been made in a criteria [13] . It can be described in F.
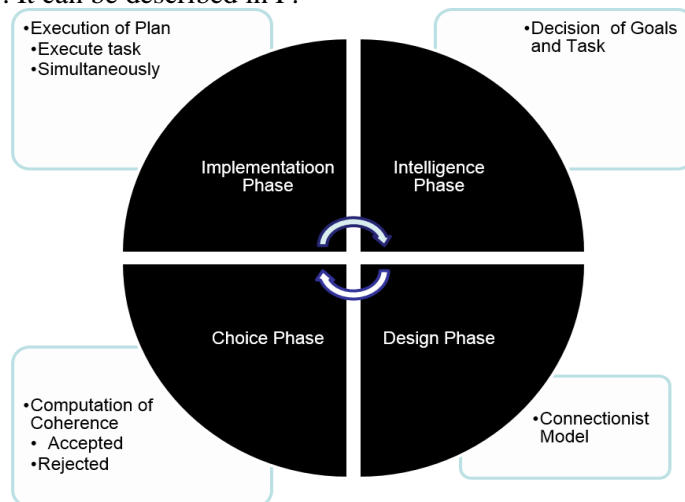

Figure 1. Decision Making Deliberative Coherence Model

2.7 Rational Agent

An agent is always trying to optimize a performance it is measured and called as agents who have rational, since it has been chosen by selecting a reason to act, under consideration of the planning, logical, and simple reflex. They are called as Rational Agent [24], [25]. Implementation of a planning of the agency that has been programmed to follow the format of[25] :

$$(!) \; Event : \{Context\} \leftarrow$$
$$\{Plan\};! \; \#$$

An agent can be said rational if he/she can choose the best possibility to act at any moments, according to what he know about the environment at the time. Thus, a rational agent is expected to perform or provide the correct action. The correct action is the action that caused the agent reaches the level of the most successful [24].

Knowledge of agents in the environment of the whole situation is not always enough. Thus, a particular agent must be given information about the purpose of which is to be achieved by state agents. Thus, the agent will work to achieve its objectives. Searching and planning are two rows of work that are being done to achieve the goal of an agent. The goal-based reflex agent adds information about these goals (Fig. 3). If rational agents able to act rationally it means they have the best ability and they able to remain elaborating objectives, then the action is called rational action [26].

2.8 Rational Action

If the agent is rational action by starting from an ideal situation, it would follow $\forall n.H$ (σ, n, s). but if the agent acted with uncertain situations, it will enable keep track goals, with $\forall s.K$ (ss) $\supset \forall n.H$ (σ, n, s) It so rational if the agent pursuing a strategy of coherence by considering the priority goal[14][26].

According to [26] the final state, will elaborate on some estimates of the situation they are better than $\sigma 2$ $\sigma 1$ for priority with the situation (ie, P ($\sigma 1$, $\sigma 2$, n, s)), if: (i) better than $\sigma 2$ $\sigma 1$ then a top priority and (ii) $\sigma 1$ $\sigma 2$ dominate at the level of n (ie, n _s $\sigma 1$ $\sigma 2$), it is to achieve an outcome of the agent who acted according to their priority.

3. Scenario And Experiment

Lumen Robot Friend (LRF) is a development project "Robot Friend" by LSKK-ITB Research that can move (act like a human being) using NAO platform. NAO is a humanoid robot that has been developed since 2004 by a robotics company in Paris called Aldebaran Robotics. NAO was first launched on August 15, 2007 were used in the RoboCup Platform League (SPL), an international robot soccer competition. NAO has on-board Computer Atom Z530 1.6 GHz, has a wide range of sensors including cameras, microphones, sonar sensors, tactile sensors, gyrometer and accelerometer. For actuators, NAO brush DC coreless equipped with 25 DOF in all parts and speakers located on the side of the head.

Table 1. Spesifikasi LRF

Category	Spesification
High	23 in (58 cm)
Weight	9,5 lb (4,3 kg)
Otonomi	60 minute (Active Condition), 90 minute (normal condition)
degrees of freedom	21-25
CPU	Intel Atom @ 1,6 GHz
OS	Linux
OS Supported	Windows, Mac OS, Linux
Language Programming	C++, Phyton, Java, MATLAB, Urbi, C, Net
Vision	Two HD 1280*960 cameras
Conektivity	Ethernet, Wi-Fi, Infrared
Spesification	Description

Application of multi-approach (Fig. 4) is a combination of visual processing LRF through the camera, do the analysis and decide what they should do. Gesture Recognition is the implementation of the introduction of Hand Motion through language that uses hand movements to communicate in conveying information in the form of an order to carry out state wake of the condition of rest. In this process, skin detector uses an algorithm developed by[27] are understood and executed by the agent as a wake-up motion. As for understanding the target object to be searched is the application [28] which is associated with the use of datasets for object detection [22].
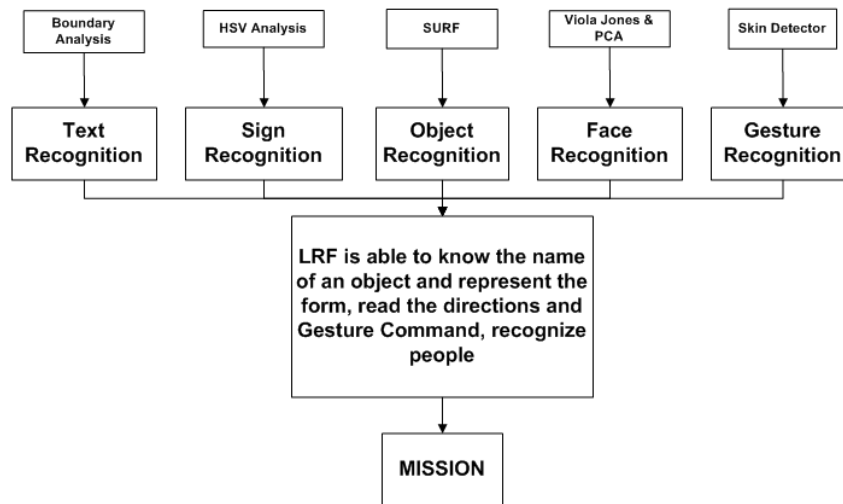

Figure 2. Integration Multi Approach Diagram

Exploration conducted by LRF is to explore their environment through movement forward, turn left, turn right, turn. As a guide the direction of motion is through LRF understanding of the Sign (Left, Right, Forbidden), which is implemented through the algorithm [11]. Furthermore, to give the object on the right is the use of algorithms [5]. To develop the agent at LRF, then described in detail regarding the task environment using PEAS (Performance, Environment, Actuators, and Sensors) [24]. According to

Goal agents LRF's mission is (1) capable of searching for objects and give it to someone, (2) capable of exploring the environment through vision. Task LRF is (1) see a task that is determined via text (2) understand and interpret the gesture person becomes an order, (3) stand seeing calls (4) detects a command based text (4) take the objects by selecting (5) runs to understand the signs around it (6) looking for someone (7) determining the provision object to someone according to his decision. (8) Ask them to go to the same place. LRF scenario that will do the search for objects through Pseudocode mission described in Fig 4.

As the implementation that has been done by LRF ability, it recognizes a command gesture, face recognition, object recognition, sign and text recognition can be used as an approach that robots can help humans. The using of LRF camera also has important participation, if the level of accuracy better so that it will be influenced other such as the lighting, sign color clarity, and range. Introducing this robot to public is affected by light intensity and amount of training data so that the accuracy of facial recognition is not maximized.

However, further research is needed to improve the ability of the LRF so that the robot will be able to automatically in performing many tasks, it is necessary to use deliberative approach coherence that LRF is able to decide what action to do next are prioritized in a complex and dynamic environment.

References

[1] L. Ismail, S. Shamsuddin, H. Hashim, S. Bahari, and A. Jaafar, "*Face Detection Technique of Humanoid Robot NAO for Application in Robotic Assistive Therapy,*" 2011, pp. 517–521.

[2] Q. Li, U. Niaz, and B. Merialdo, "*An improved algorithm on Viola-Jones object detector*," in Proceedings - International Workshop on Content-Based Multimedia Indexing, 2012, pp. 55–60.

[3] S. Agrawal and P. Khatri, "*Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis,*" in 2015 Fifth International Conference on Advanced Computing & Communication Technologies, 2015, pp. 108–112.

[4] P. Viola and M. Jones, "*Robust real-time face detection,*" Int. J. Comput. Vis., **vol. 57**, no. 2, pp. 137–154, 2004.

[5] P. Viola and M. Jones, "*Rapid object detection using a boosted cascade of simple features,*" Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition. CVPR 2001, **vol. 1**, 2001.

[6] R. Kolar, Akshay Thakar, and M. Shabad, "*Image Segmentation for Text Recognition using Boundary,*" Int. J. Emerg. Technol. Adv. Eng., **vol. 4**, no. 2, pp. 294–298, 2014.

[7] Y. Zhu, C. Yao, and X. Bai, "*Scene Text Detection and Recognition : Recent Advances and Future Trends,*" Front. Comput. Sci., 2014.

[8] Q. Ye, D. Doermann, and S. Member, "*Text Detection and Recognition in Images and Video : a Survey*," **vol. 8828,** no. c, pp. 1–20, 2014.

[9] T. A. Pham, M. Delalandre, and S. Barrat, "*A contour-based method for logo detection,*" in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2011, pp. 718–722.

[10] K. R. Soumya, A. Babu, and L. Therattil, "*License Plate Detection and Character Recognition Using Contour Analysis,*" **vol. 3**, no. 1, pp. 15–18, 2014.

[11] C.- Gubbi, "*Automatic Tracking of Traffic Signs Based on HSV,*" Int. J. Eng. Res. Technol., **vol. 3**, no. 5, pp. 914–917, 2014.

[12] Y. Chen, "*Detection and Recognition of Traffic Signs Based on HSV Vision Model and Shape features,*" J. Comput., **vol. 8**, no. 5, pp. 1366–1370, 2013.

[13] O. Topcu, "*Adaptive decision making in agent-based simulation,*" Simul. Trans. Soc. Model. Simul. Int., **vol. 90,** no. 7, pp. 815–832, 2014.

[14] S. Isci, O. Topcu, and L. Yilmas, "*Extending the Jadex Framework with Coherence- Driven Adaptive Agent Decision-Making Model,*" in 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligence Agent Technologies (IAT), 2014, pp. 48–55.

[15] H. Zhou, A. Mian, L. Wei, D. Creighton, and M. Hossny, "*Recent Advances on Singlemodal and Multimodal Face Recognition : A Survey*," IEEE Trans. Human-Machine Syst**., vol. 44**, no. 6, pp. 701–716, 2014.

[16] A. Özdil and metin mete Obilen, "*A Survey on Comparison of Face Recognition Algorithms,*" in 8th International Conference on Application of Information and Communication Technologies, 2014, pp. 249–251.

[17] M. N. M. Yawale, "*Face Detection and Recognition Using Eigen Faces by Using PCA,*" Int. J. Technol. Appl., **vol. 4,** no. 1, pp. 84–87, 2013.

[18] L. C. Paul and A. Al Sumam, "*Face Recognition Using Principal Component Analysis Method,*" **vol. 1**, no. 9, pp. 135–139, 2012.

[19] Q. Ye and D. Doermann, "*Text Detection and Recognition in Images and Video : a Survey,*" IEEE Trans. Pattern Anal. Mach. Intell., **vol. 8828**, no. c, pp. 1–20, 2014.

[20] C. Yao, X. Bai, and W. Liu, "*A Unified Framework for Multioriented Text Detection and Recognition,*" IEEE Trans. Image Process., **vol. 23**, no. 11, pp. 4737–4749, 2014.

[21] F. Ren, J. Huang, R. Jiang, and R. Klette, "*General traffic sign recognition by feature matching,*" 2009 24th Int. Conf. Image Vis. Comput. New Zeal., no. Ivcnz, pp. 409–414, 2009.

[22] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "*Speeded-Up Robust Features (SURF),*" no. September, 2008.

[23] A. Hajraoui and M. Sabri, "*Face Detection Algorithm based on Skin Detection, Watershed Method and Gabor Filters*.," Int. J. Comput. Appl., **vol. 94,** no. 6, pp. 33–39, 2014.

[24] S. Russel and P. Norvig, *Artificial Intelligence A Modern Approach,* Third Edit. Pearson Education Inc., 2003.

[25] N. K. Lincoln, S. M. Veres, L. A. Dennis, and M. Fisher, "*Autonomous Asteroid Exploration by Rational Agents*," IEEE Computational Intelligence Magazine, no. November, pp. 25–38, 2013.

[26] S. Sardi and S. Shapiro, "*Rational Action in Agent Programs with Prioritized Goals*," in AAMAS, 2003, pp. 417–424.

[27] B. Jalilian and A. Chalechale, *"Face and Hand Shape Segmentation Using Statistical Skin Detection for Sign Language Recognition,"* **vol. 1**, no. 3, pp. 196–201, 2013.

[28] S. Mandvikar, "*Augmented Reality USing Countour Analysis In e-Learning,*" Internatioanl J. Innov. Res. Sci. Eng. Technol., **vol. 2**, no. 5, pp. 1591–1595, 2013.