Detection of Static Body Poses using Skin Detection Techniques to Assess the Accessibility Information of Human Beings During HRI

Surekha Dongre[1], Rahul Nawkhare[2]

[1]*Student, Dept. Of Electronics Engineering, Wainganga college of Engineering and Management, Dongargaon, Maharashtra, India*
[2]*Professor, Dept. Of Electronics Engineering, Wainganga college of Engineering and Management, Dongargaon, Maharashtra, India*

---***---

Abstract -*The future of robotics lies in Human-robot interactions (HRI), wherein, a robot should be able to interpret the human gestures and body language and respond appropriately. This ability of robots will help them to easily get assimilated in the society. For this end, the challenges faced are to perfect the ability of the robot to distinguish between various non-verbal cues exhibited by human beings in their interactions with them. In this paper, the focus is on body postures of a person, especially the static and natural body language exhibited by him/her. Five body postures are detected automatically and classified into categories of accessibility of human being, namely, not – accessible, accessible and neutral. The detection is done using skin colour segmentation using superpixels and Hidden Markov Model (HMM) technique is used to automatically classify images of the static poses.*

Key Words: HRI, static body postures, HMM, superpixels, skin colour segmentation.

1. INTRODUCTION

HRI is a sub-domain of Artificial Intelligence (AI), which focuses on the social aspect of the robot's design. The goal of HRI research is mainly to improve the communication methods of the robot with a human by training the robot to identify the emotional state of the humans through their gestures (both dynamic and static), speech, facial expressions and body language [1].

Most of the research in the training phase is originated from centuries of psychological research of human behaviour. The robots have to simulate the human – to – human interaction as accurately as possible. A Human mind takes in millions of sensory cues every time it interacts with another human and the reciprocating behaviour is decided by the cues the mind processes based on the inputs it receives from visual, tactile and auditory senses. This requirement is more critical in areas where close proximity with humans is necessary, for example, in areas such as rehabilitation, search and rescue operations, hospitality, assistive robots in schools, homes and hospitals.

1.1 Body language and static poses

There are many forms of poses and gestures identified in humans in social situations. They fall under two broad categories, namely; dynamic and static. There are many computer vision techniques proposed to detect these human gestures and poses. Dynamic gesture involves movement of upper limbs. While segmenting the human gesture information, the speed and position information has to be taken into consideration [2]. Static poses, on the other hand, do not require speed and position information to be tracked by a camera sensor. The other advantage of using static body poses is that the person displaying the pose isn't usually aware of it, and thus it is considered to be natural.

In this paper, we focus on static poses involving the head and lower arms. Five static pose orientations have been chosen with combination of head and lower arms. The static poses are known to convey various human emotions or state of mind during conversation or interaction with them in social situations, namely, stress levels, openness to interact, mood, feelings and emotions [3] [4].

1.2 Skin Detection

The static poses are detected and segmented based on the skin color information. Many HRI based gesture and pose detection papers include skin detection as one of the steps to segment the pose or gesture area. This method is done before the classification stage. The skin areas are identified to be in the head and arms region. Skin detection is a challenging concept because of the different types of colors present in the human population. While detecting skin region, the color space has to be appropriately chosen.

There are many color spaces available in computer vision, namely; YCbCr , CMY, UVW, RGB, LSLM, xyY, HSI, L*a*b*, L*u*v*, XYZ, LHC, HSV, YIQ, LHS, YUV [5]. In [6], the skin features are extracted based on the number of pixels in the Region of Interest (ROI), the centroid location obtained from a 3D skeleton model acquired by kinect sensor, the pixel amount in the region's perimeter, the expansivity and eccentricity of the skin region. The color space chosen here is YcbCr because of its non – linear characteristics and close representation of how human color perception actually works and also because it works in images with illumination differences in different regions. Moreover, skin segmentation can be done using pixel – based detection and region – based detection. In pixel based detection, each pixel is checked for whether it belongs to skin or non-skin region. In the region based detection, the spatial information is gathered and arranged for better performance. But this technique requires good illumination, texture and intensity information.

2. PROPOSED METHOD

The paper proposes a novel method of detecting pixels in the skin region. This section details the various processes used to segment the skin region so as to ultimately train the algorithm to be able to classify the accessibility rate of a human according to the static body pose displayed.

2.1Preprocessing

The image acquired through a 2D camera might contain noise pixels, due to the resolution and sensor of the camera. The noise information contained in the image is removed using a combination of Gaussian and median filters. Median filters are commonly used for removing noise in low – frequency pixel intensity regions. It is also known as a smoothing filter. Gaussian filter retain the details and remove noise from high frequency pixel intensity regions. This filter is also used to remove blur in images. Existence of noise in images can lead to an image being segmented and classified incorrectly.

2.2 Segmentation and Classification

The skin region is detected first using superpixels based on Simple Linear Iterative Clustering (SLIC) algorithm. It is a color based segmentation technique and hence is preferred for skin detection. SLIC algorithm clusters pixels of similar intensities, hence when segmentation or feature extraction is done, computation time is significantly reduced. Here clustering of pixels is done in YCbCr color space. YCbCr color space is efficient in terms of extracting luminance values as well.

Once the superpixels cluster and segment the skin region, the different static poses are classified according to their accessibility degree. For example, psychologists study the body language of their patients to assess their willingness to be evaluated by them. It is a common knowledge that a person with hands folded is said to be not open to discussion or agreement with the opposite person. Similarly, in this project, we categorize the degree of openness to communicate or what is called the degree of accessibility with regard to communicating with a robot. There are three differentcategories into which the five images are classified into, namely, accessible, not – accessible, neutral. . Each skin region can be identified as one of five different lower arm and/or head configurations: 1) lower arm; 2) head 3) crossed arms; 4) both the lower arms are touching and 5) two arms touching or one arm touching the head.

Five normalized geometric features are identified for each skin region in order to autonomously classify the region into the aforementioned configurations via a learning technique. These features include: 1) the number of pixels within the region; 2) the location of the centroid of the region; 3) the number of pixels along the perimeter of the region; 4) the eccentricity of the skin region; and 5) the expansiveness of the region [6].

After these values are calculated, they are used to classify the images using HMM due to its advantages when compared to SVM classifier [18] into the aforementioned categories.

3. RESULTS

The skin area is detected and segmented as shown below. The Fig 1. Is categorized as not accessible, Fig 2 and 5 are categorized as neutral accessibility and fig, 3 and 4 are grouped into accessible category. This grouping into classes is done using HMM classifier.

Fig 1. Head and crossed arms.

Fig 2. Lower arm and another arm touching head

Fig 3. Both lower arms touching and head

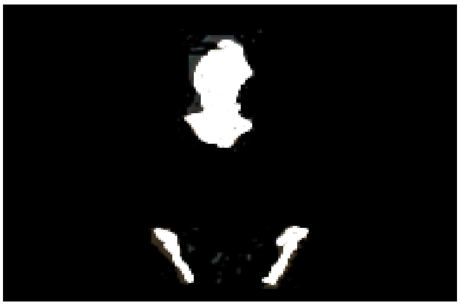

Fig 4. Head and lower arms (not - touching)

Fig 5. One of the arms touching and another folded

3. CONCLUSION

The proposed method is quite robust and new combination of techniques have been used. Superpixels help improve the speed and robustness of the skin detection stage. HMM classifies the images into three different categories based on the parameters discussed. The future scope of this project is to include more categories of accessibility.

REFERENCES

[1] M. A. Goodrich and A. C. Schultz, "Human–robot interaction," Found.Trends Human Comput. Interact., vol. 1, no. 3, pp. 203–275, 2007.

[2] A dynamic gesture recognition and prediction system using the convexity approach, Pablo Barrosa Nestor T . Maciel-Junior Bruno J. T. Fernandes Byron L. D. Bezerra Sergio M. M. Fernandes, Computer Vision and Image Understanding Volume 155, February 2017, Pages 139-149.

[3] M. Davis and D. Hadiks, "Nonverbal aspects of therapist attunement,"J. Clin. Psychol., vol. 50, no. 3, pp. 393–405, 1994.

[4] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," J. Nonverbal Behav., vol. 28, no. 2, pp. 117–139, 2004.

[5] G. Gomez, On selecting colour components for skin detection, in: Proceedings of International Conference on Pattern Recognition, vol.2, Quebec, 11–15 August 2002, pp. 961–964.

[6] Classifying a Person's Degree of Accessibility fromNatural Body Language During Social Human–Robot Interactions, Derek McColl,Chuan Jiang, and Goldie Nejat, IEEE Transactions on Cybernetics (Volume: 47 , Issue: 2 , Feb. 2017) 524 – 538

[7] A skin tone detection algorithm for an adaptive approach to steganography Abbas Cheddad, Joan Condell, Kevin Curran, Paul Mc Kevitt , Signal Processing 89 (2009) 2465–2478, elsevier.

[8] A Survey on Pixel-Based Skin Color Detection Techniques, Vladimir Vezhnevets , Vassili Sazonov, Alla Andreeva

[9] P. Ekman and V. Friesen, "The repertoire of nonverbal behavior:Categories, origins, and coding," Semiotica, vol. 1, no. 1, pp. 49–98, 1969.

[10] J. Sanghvi et al., "Automatic analysis of affective postures and bodymotion to detect engagement with a game companion," in Proc.ACM/IEEE Int. Conf. Human Robot Interact., Lausanne, Switzerland,2011, pp. 305–311.

[11] T. Lourens, R. van Berkel, and E. Barakova, "Communicating emotionsand mental states to robots in a real time parallel framework using Labanmovement analysis," Robot. Auton. Syst., vol. 58, no. 12, pp. 1256–1265, 2010

[12] Comparative Study of Skin Color Detection and Segmentation inHSV and YCbCr Color Space, Khamar Basha Shaik Ganesan P ,V.Kalist , B.S.Sathish , J.Merlin Mary Jenitha, ScienceDirect Procedia Computer Science 57 (2015) 41 – 48

[13] HMM-Based Gesture Recognition for Robot Control, Hye Sun Park , Eun Yi Kim , Sang Su Jang , Se Hyun Park , Min Ho Park, and Hang Joon Kim, pp. 607–614, 2005.Springer-Verlag

[14] Human–Robot Interaction: A Survey, Michael A. Goodrich and Alan C. Schultz, Foundations and Trends #inHuman–Computer InteractionVol. 1, No. 3 (2007) 203–275.

[15] Human Skin Detection Using Histogram Processing and Gaussian Mixture Model Based on Color Spaces, Satishkumar L, Varma Vandana Behera, Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017)

[16] Skin Color Detection Based Super Pixel, Rong Hua, Yang Wang, 2017 3rd IEEE International Conference on Computer and Communications.

[17] Ahmed Elgammal, Crystal Muang and Dunxu Hu, "Skin Detection -a Short Tutorial" Encyclopedia of Biometrics by Springer-Verlag Berlin Heidelberg 2009

[18] Comparative Study on Hidden Markov Model Versus Support Vector Machine, A Component-Based Method for Better Face Recognition. Mukundhan Srinivasan and Sriram Raghu., 2013 UK Sim 15th International Conference on Computer Modelling and Simulation. IEEE computer science society.