

3D Image Processing using Machine Learning based Input Processing for Man-Machine Interaction

Dr. Akey Sungheetha

Data Science SIG member,
Computer Science and Engineering,
School of Electrical Engineering and Computing,
Adama Science and Technology University,
Adama, Nazret,
Ethiopia.

Dr. Rajesh Sharma R

Image Processing SIG member,
Computer Science and Engineering,
School of Electrical Engineering and Computing,
Adama Science and Technology University,
Adama, Nazret,
Ethiopia.

Abstract: In various real time applications, several assisted services are provided by the human-robot interaction (HRI). The concept of convergence of a three-dimensional (3D) image into a plane-based projection is used for object identification via digital visualization in robotic systems. Recognition errors occur as the projections in various planes are misidentified during the convergence process. These misidentifications in recognition of objects can be reduced by input processing scheme dependent on the projection technique. The conjoining indices are identified by projecting the input image in all possible dimensions and visualizing it. Machine learning algorithm is used for improving the processing speed and accuracy of recognition. Labeled analysis is used for segregation of the intersection without conjoined indices. Errors are prevented by identifying the non-correlating indices in the projections of possible dimension. The inputs are correlated with related inputs that are stored with labels thereby preventing matching of the indices and deviations in the planes. Error, complexity, time and recognition ratio metrics are verified for the proposed model.

Keywords: Space projection, digital visualization, human-robot interaction, dimension modeling, 3D images;

1. Introduction

In order to access any object without human involvement within specified time period, human-robot interactions are used. Virtual reality objects are defined using the three-dimensional (3-D) image sequences in Virtual reality modeling language (VRML) [1]. Virtual objects and their scenes may be approached using VRML programming scheme. Based on the object, a virtual robot is generated for supporting the workspace environment. In web settings, a series of virtual images may be generated by the user with the help of VRML that offers connectivity with a seemingly 3D scene by spinning, turning, watching or otherwise [2]. Currently, multipurpose environments or virtual reality headsets are used by the conventional augmented reality devices for creation of compelling emotions, sounds or images by stimulating the actual appearance of a person in a virtual world. Communication between the user and the virtual functions and elements in the simulated environment is made possible using augmented reality devices. From the sensed object, a quantitative data collection may be derived and the location of the object may be observed in this area [3]. The object is acquired from the surrounding in a 3D view for visualization process. Conversion processes are implemented as the object in a 3D view may not be recognized by the robot. Movement and position of the object in space may be detected using a robot configuration space. The changes in position is processed and determined every time by sequential monitoring of objects [4]. With respect to the previous history, the nature of the object is detected by the sensor and the object is visualized. The result of the object in space is derived on implementing matching schemes for finding the object [5].

The visual view of the robot obtained from its surroundings is used for detecting the 3D object. 2D or 3D visualization of an object sensed from the environment is performed based on the boundary box. The labeled objects and their features are used for processing the prediction schemes on identification of the object [6]. The plane and position of the object are represented using its features on detecting the object from the surroundings. The 3D object is obtained by matching the labeled object in the previous state along with the input object. The forthcoming object is deployed using a boundary box derived based on the training dataset during evaluation [7]. Volume, shape and pose of the object are the three aspects based on which the three corresponding planes are represented. Hence pattern identification is performed by the robot based on the extracted features of the acquired object. The surface and structure of the object in the plane corresponds to the patterns identified. Pre-defined knowledge available is used for determining the nature of the object present on the plane. The intensity range and 3D bitmap is used for representing the 3D analysis during visualization process [8]. The region of interest (ROI) is obtained by evaluation of the 3D based image processing and recognition. Based on the quality of image, the necessary features are extracted from the object. Geometric transformation and other deviation angles are used for measuring the 3D image representation.

The position of the object in the plane is measured through quantitative analysis. Objects are acquired using image processing schemes and their dimensions are detected and categorized into 2D or 3D [9]. Any specific object visualized by the robot can be processed with the inference from this identification. A large percentage of processing time is essential for identification of inner and outer contours. Tracking of contours takes more time based on the number of contours identified as the contours of the picture in a recurrence loop are controlled by contour monitoring [5]. The robot vision system is largely dependent on the image recognition component. The acquired image is preprocessed by the robot. Further, picture recognition is performed and the location is identified for reducing the signal to noise ratio and enhancing the image quality. Error and process complexity are avoided while arranging the image with the help of efficient classifiers. The error and processing while increasing the recognition ratio is the major contribution of this paper. Three dimensional objects are converted into two dimension while addressing the correlation factor and detecting the conjoint index of the object acquired.

2. Related Work

Robotics and computer vision is greatly benefited by the application human-computer interaction (HCI) and human-robot interaction (HRI) [10]. A HCI system based on vision is used for understanding the sensory motion and manipulating the robotic arm while reducing the system complexity. The efficiency of this technique is validated by computational analysis. The challenges in conditional generative adversarial network (CGAN) [11] is optimized by a crucial emotion recognition system termed as facial expression recognition (FER). The FER challenge represent substantial differences between machine and human [12]. The generative discriminatory representations and facial expression factor is validated simultaneously with the help of CGAN. Quantitative and qualitative test results are obtained for recognizing the expression using FER system. Required information from a 3D image is extracted using a 3D integral image display in [13]. The distance between input images are compared for reducing the error with the help of Axially Distributed Sensing (ADS).

The 3D max array is used for extracting depth map. 3D localization is introduced in mobile robot while deploying point cloud treatment and Monte Carlo localization in [14]. A 3D point cloud is deliberated in an unstructured and uneven platform during evaluation. 2D conjunction and occupancy map is obtained by feature extraction. Prior orientation and vanishing point are fused in 3D object detector for multiple processing in [15]. Mean average precision score is improved while plotting the pixel coordinates based on vertices. The image accuracy range is enhanced using Euler angle principle. High frame rate is improved while estimating the depth and intensity of an image using Time-of-flight (ToF). The object's reliable visual information is enhanced using geometric and intensity features. In real time environment, the object movements can be detected.

3. Proposed Methodology

The objects identified from the surroundings are shared between humans and robots by means of human robot interaction. 3D based images are more commonly obtained from the environment. However, detection of these images is difficult for robots. For this purpose, we deploy a 3D to 2D conversion process for accurate identification of the object and to provide results to the humans. Figure 1 represents the architecture of the proposed HRI model using machine learning and input processing.

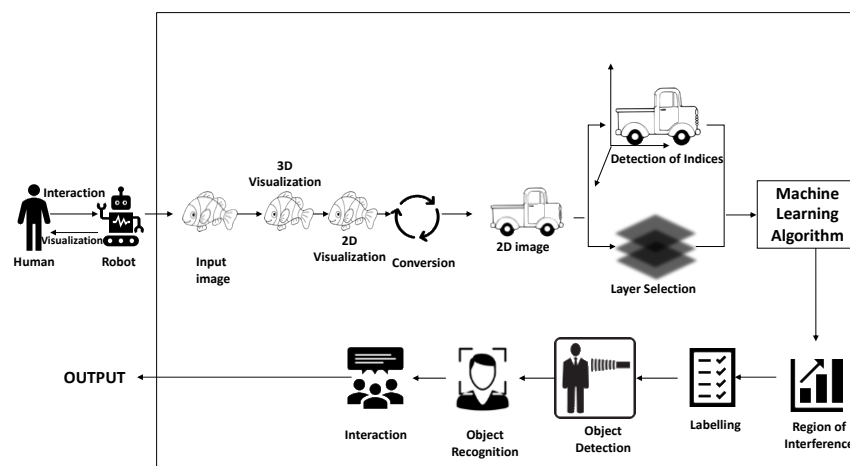


Figure 1: Proposed HRI model using Machine Learning and Input processing

Machine learning algorithm and projection dependent input processing features are used for image recognition using specific entity and prevention of incorrect identification. On the basis of joint indexes, recurrent analysis is performed for deletion of detected components. Timeline recurrent analysis is used for describing and denoting the identification based on the prior analysis. The direction of the object is acquired by the robot from a specific region from which the sense of the object is used for validation of the input. Plane calculation is used for detection of position and indices of the object. These factors are detected based on the 3D to 2D object that is obtained. Two types of optimization is performed for the indices. First type of optimization is on the principle that multiple positions may be available for objects that reside on a plane. Second optimization technique is based on the principle that the object relies on a different plane if the object position does not vary. Optimization of indices is performed for detection and calculation of these constraints. The plane in which the object resides and its position are the two major constraints based on which the object is selected from the plane and output is produced.

4. Results and Discussion

3D images from a source are used for analysis of the performance of the proposed machine learning algorithm based input processing model. 1000 sample images from various classes are considered as 3D models. Verification of correlation and matching is performed using a testing and training dataset with 500x2 images. MATLAB software is used for analyzing these images. For every matching instance, 50 images are trained. Input correlation and verification of indices is performed for every instance. Error, processing complexity, processing time and recognition ratio are analyzed for estimating the efficiency of the proposed method. In all tests, the association factor and number of items are verified. Existing schemes like multi-scale convolutional neural network, Context-Assisted 3D, Conventional voxel-based occupancy grid, radio-frequency identification, support vector regression and Dynamic Statistical Parametric Mapping are compared. Figure 2 provides few examples of input and output image samples.

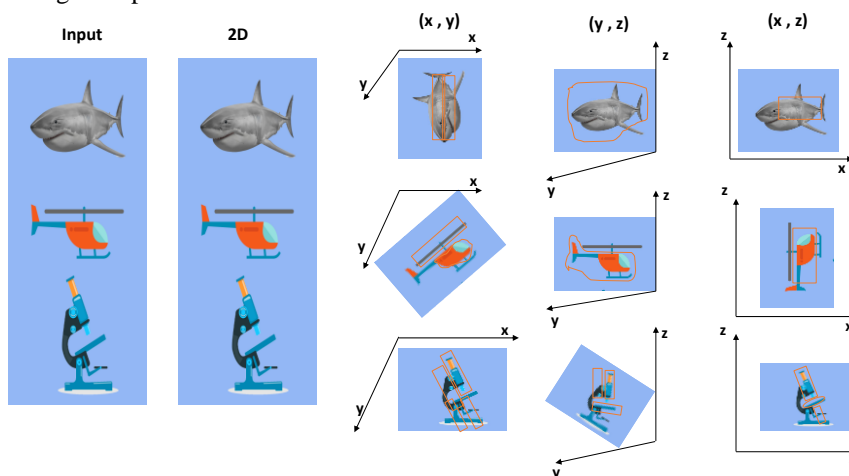


Figure 2: Input and Output Image Samples

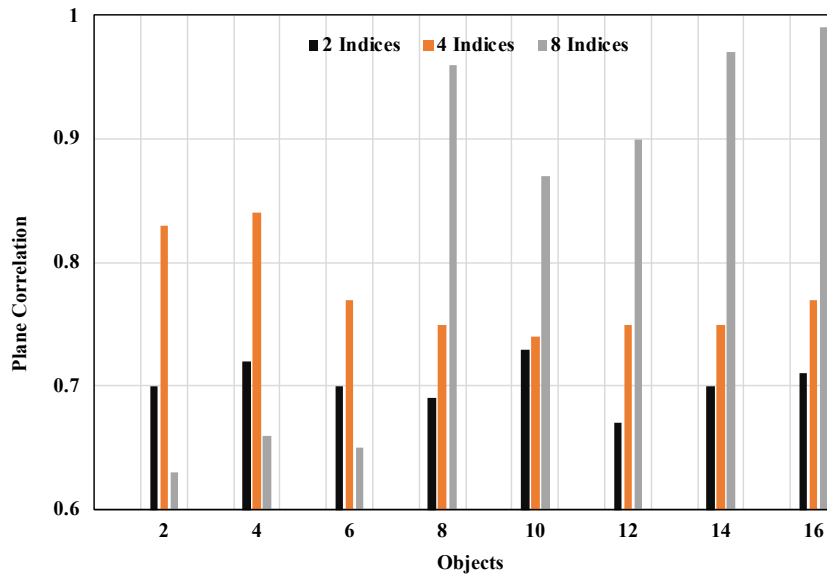


Figure 3: Analysis of Plane Correlation

Figure 3 provides the graphical representation of analysis of plane correlation. Every forthcoming object is assigned with a label based on the varying indices detected for the object. The previous object is used as a reference for matching the planes and positions during analysis. The correlation increases with the increase in the plane. The indices of the object also varies simultaneously.

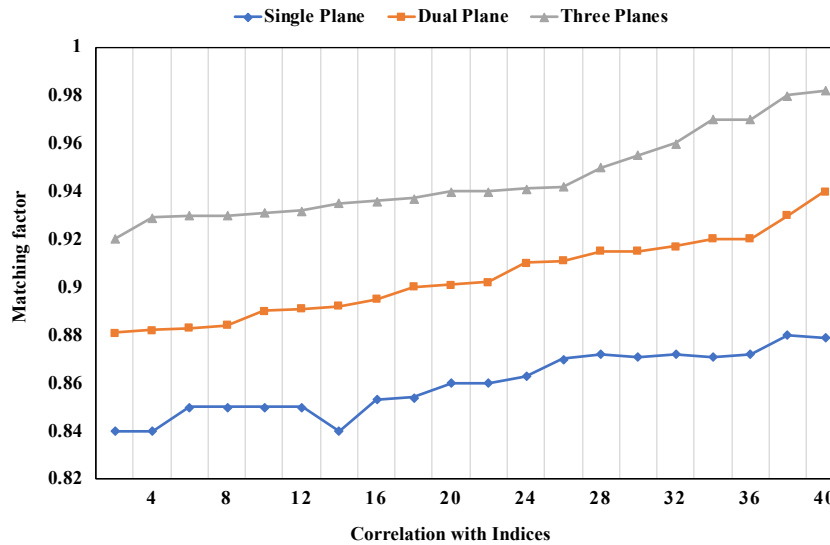


Figure 4: Analysis of Matching Factor

Figure 4 represents the analysis of matching factors on multiple planes. The sensor provides a series of objects on which data indices are performed. The object features are defined using labels. Reliable matching is performed on detection of object despite variation in indices. The data indices increase with the increase in the plane and matching.



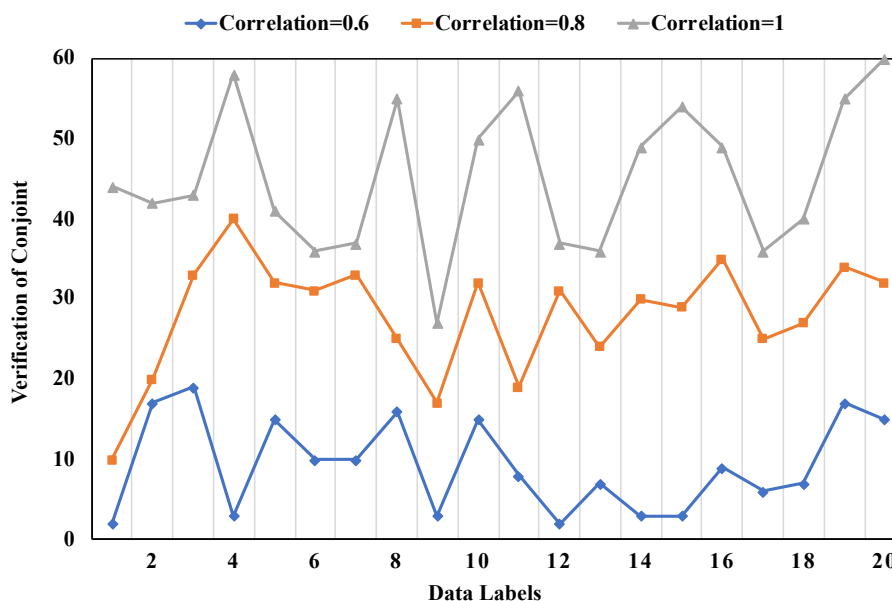


Figure 5: Verification of Conjoint

Figure 5 represents the verification of conjoint with the proposed architecture. Easy analysis may be performed as the plane and position of all forthcoming objects are depicted by the labels assigned to them. The label of the object increases with better correlation matching that helps in evaluation of verification of conjoint. The correlation factor is improved when there is an increase in conjoint verification caused by increase in the object label. Improved correlation is obtained with a decrease in error rate using the label. The object verification is processed and error is reduced for every varying object. The conversion process helps in determining layer selection. Further, error in the labels decrease with increase in correlation.

5. Conclusion

In human-robot interaction settings, 3D objects are identified using the proposed projection based input processing model. Based on layer selection and vector representation, a 3D object is converted into various planes of 2D. Two vectors are defined and their cross product is determined based on the element wise vector product in a 3D space. The external product with an n dimensional abstract vector product, the name vector is combined with wedge notation. With respect to the concurrent and prior detection process, labeling of indices is performed during the conversion process. The errors are reduced and identification and mitigation of non-correlating points from different planes is also done based on the indices matching instances. Detection of region of interest in the indices data projection helps in identifying the variations in the plane. Periodic matching updates are performed along with correlation and matching process concurrently. The error, complexity and processing time is reduced while improving the recognition ratio using the proposed technique based on experimental analysis. Future work is directed towards improving these values further and enhancing the performance of the model during real time implementation.

References

- [1] Luo, R. C., & Wu, X. (2014, March). Real-time gender recognition based on 3d human body shape for human-robot interaction. In Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction (pp. 236-237).
- [2] Waldherr, S., Romero, R., & Thrun, S. (2000). A gesture based interface for human-robot interaction. *Autonomous Robots*, 9(2), 151-173.
- [3] Liu, Z., Wu, M., Cao, W., Chen, L., Xu, J., Zhang, R., ... & Mao, J. (2017). A facial expression emotion recognition based human-robot interaction system.
- [4] Li, X. (2020). Human-robot interaction based on gesture and movement recognition. *Signal Processing: Image Communication*, 81, 115686.

- [5] Mazhar, O., Ramdani, S., Navarro, B., Passama, R., & Cherubini, A. (2018, October). Towards real-time physical human-robot interaction using skeleton information and hand gestures. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 1-6). IEEE.
- [6] Li, J., Mi, Y., Li, G., & Ju, Z. (2019). Cnn-based facial expression recognition from annotated rgb-d images for human-robot interaction. *International Journal of Humanoid Robotics*, 16(04), 1941002.
- [7] Filippini, C., Perpetuini, D., Cardone, D., Chiarelli, A. M., & Merla, A. (2020). Thermal infrared imaging-based affective computing and its application to facilitate human robot interaction: a review. *Applied Sciences*, 10(8), 2924.
- [8] Chen, L., Zhou, M., Su, W., Wu, M., She, J., & Hirota, K. (2018). Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction. *Information Sciences*, 428, 49-61.
- [9] Du, G., Chen, M., Liu, C., Zhang, B., & Zhang, P. (2018). Online robot teaching with natural human-robot interaction. *IEEE Transactions on Industrial Electronics*, 65(12), 9571-9581.
- [10] Deng, J., Pang, G., Zhang, Z., Pang, Z., Yang, H., & Yang, G. (2019). cGAN based facial expression recognition for human-robot interaction. *IEEE Access*, 7, 9848-9859.
- [11] Fang, B., Sun, F., Liu, H., & Liu, C. (2018). 3D human gesture capturing and recognition by the IMMU-based data glove. *Neurocomputing*, 277, 198-207.
- [12] Shridhar, M., & Hsu, D. (2018). Interactive visual grounding of referring expressions for human-robot interaction. *arXiv preprint arXiv:1806.03831*.
- [13] Shakya, S. (2019). Virtual restoration of damaged archeological artifacts obtained from expeditions using 3D visualization. *Journal of Innovative Image Processing (JIIP)*, 1(02), 102-110.
- [14] Dhaya, R. (2020). Improved Image Processing Techniques for User Immersion Problem Alleviation in Virtual Reality Environments. *Journal of Innovative Image Processing (JIIP)*, 2(02), 77-84.
- [15] Ranganathan, G. (2020). Real Life Human Movement Realization in Multimodal Group Communication Using Depth Map Information and Machine Learning. *Journal of Innovative Image Processing (JIIP)*, 2(02), 93-101.