Underwater image reconstruction using convolutional auto-encoder

Shinsuke Yasukawa
Dept. of Human Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: s-yasukawa@edu.brain.kyutech.ac.jp

Sreeraman Srinivasa Raghura
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: srinivasa.raghuraman-sreeraman344@mail.kyutech.jp

Yuya Nishida
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: y-nishida@brain.kyutech.ac.jp.

Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp.

Abstract

One of the main tasks of AUVs is to capture deep-sea images like fishes, crabs, other living organisms and resources for information leading to research on deep-sea ecosystems. Acoustic transmission are used to establish wireless underwater communications between the AUV and the ship. However, there are some limitations in the communication channels due to limited bandwidth, multi-path, temperature distribution and change in the direction of transmitting source and receiving sensor which results in losses in data being transmitted. Initially, the captured images are enhanced to reduce the effect of light attenuation and then compressed for transmission through acoustic modems. Only an important part of image is being transmitted through set of data packets. The received data packets in the ship will be reconstructed to predict the presence of living organisms. The loss in data during transmission creates a difficulty for the operators to predict the exact information. In this research, to compensate this transmission loss, an efficient compression and reconstruction technique using convolutional autoencoder with minimal distortion is proposed. Finally, for evaluation of the proposed image compression technique, the quality of reconstruction of images with and without data loss will be compared using the quality metrics signal to noise ratio (PSNR), structural similarity index(SSIM) and perceptual quality of image.

Keywords: sampling-AUV, acoustic communication, image compression, convolutional auto-encoder

1. Introduction

Presently numerous research and development activities have been carried out in the field of underwater robotics in ocean engineering. Research studies, especially in the area of autonomous underwater vehicles (AUV) which are being used as tools for deep sea-floor surveying and observation. AUVs are mainly employed for performing tasks like underwater imaging, sea-floor mapping, exploiting resources and minerals which are of great interest to geologists and marine biologists to understand the sea flora and fauna.

One of the main tasks of AUV is optical imaging of parts of the sea-floor that contains marine habitat. These kinds of images contribute to research studies on deep-sea biology. The further task involves collecting samples using end-effectors. It is difficult for the AUV to perform the above task autonomously as it does not have any prior knowledge on the targets of interest. To overcome this difficulty, it operates semi-autonomously by acquiring assistance from biologists and geologists. Fig. 1 shows the flow of the mission to sampling objects on sea-floor. The acquired image is then transmitted to the vessel. Images are transmitted underwater using acoustic waves through hydrophones. The received image in the ship is examined by the operator and if it shows any existence of marine habitat then behavior commands are sent to the AUV. Finally, AUV is sent back to the specified location for collecting samples. Our group developed the sampling-AUV for the sampling mission[1,2,3] and succeeded three times in a row in sampling experiments in Suruga Bay in March 2018[3].

The process of image transmission involves the following steps:

(i) First, the AUV acquires raw images of seafloor that are of interest.
(ii) These images are then enhanced to reduce the effect of light attenuation and unbalanced illumination.
(iii) Next, images compressed using the required compression technique for transmission through the acoustic medium.
(iv) Finally, the data received in the ship reconstructed and the operator examines the reconstructed image.

During the process of image transmission, there is a loss of data(information). The data loss is mainly due to the following factors: multipath, temperature distribution, Signal attenuation over long ranges, and directional changes with the transmitting source and the receiving sensor. Underwater communication has low data rates when compared to terrestrial communication since it uses acoustic waves instead of electromagnetic waves.

It is known that some loss occurs during the transmission of the compressed data through the acoustic communication. This data loss has to be minimized to obtain a better-quality image reconstruction at the receiving end. Therefore, we propose a lossy based image compression method that utilizes the merits of convolutional autoencoder network. The purpose of this research is developing an efficient image compression model that can reconstruct the image with minimal data loss. To achieve the above, the deep-learning approach is used to simulate the compression/ reconstruction model with the generation of data loss in the hidden codes that would result in improved Peak Signal to Noise Ratio (PSNR) and better Structural Similarity Index (SSIM).

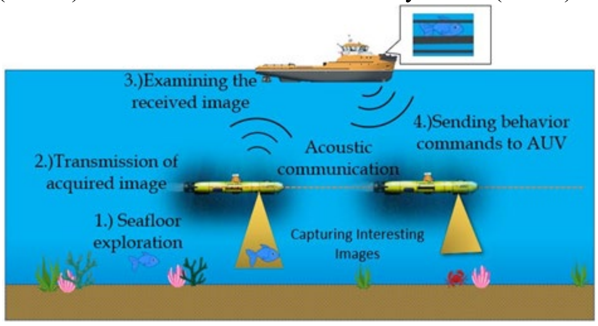

Fig. 1. Underwater Image Transmission.

2. Related works

Ahn et. al.[2] have developed an image enhancement and a compression technique for underwater acoustic communication with limited data information density. Their proposed method enhances the brightness of the sea floor images and reduces the image depth from 24-bit to 4-bit. They have used a data compression method that is based on indexed color technique which expresses images with a fixed number of color indices using a suitable color palette. The color palette consists of set of indexed colors to express the enhanced images depending on the targeted creatures. As shown in Fig. 2, the problem of data loss that occurred during transmission could not be minimized which affected the quality of image reconstruction at the receiving end. The image is transmitted in the form of packets of information.

Hoag et.al.[4] have developed an underwater image compression technique using wavelet transforms (WT). They have employed a wavelet decomposition method in combination with vector quantization for compressing still underwater images. Compared to JPEG compression their proposed method tends to perform well at low bit rates. However, over a wide range of directionality the wavelet transforms are generally inefficient in representing geometric structures. The texture and details of reconstructed images are not efficiently retained using wavelet transforms.

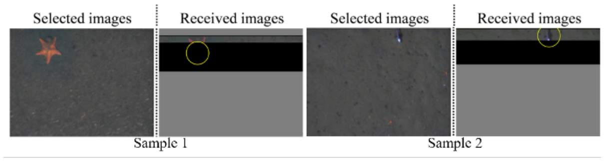

Fig. 2. Selected and received images with data loss during transmission[2].

In the research work carried out by Cheng et.al.[5] a lossy image compression method using deep convolutional autoencoders (CAE) was proposed. They have designed a convolutional autoencoder network that replaces conventional transforms and performed training using rate-distortion loss function. Compared to JPEG2000 and other traditional algorithms their method performs well on the images from kodak database.

3. Proposed Method

The Fig. 3 shows an example of overview of the proposed image compression method utilized for underwater image transmission process. Initially the acquired image data is compressed to codes (assuming the encoder in the AUV). In the compressed representation we generate a data loss by discarding a limited amount data from the codes (i.e. eliminating few rows or columns). This corrupted data will be sent to the decoder for reconstruction. The decoder extracts relevant information from this data and attempts to reconstruct the image similar to that of the original. The reconstruction depends on the robustness of the decoder.

Fig. 4 shows CAE Network with data loss. As an initial step, the dataset consisting of interesting images was created for training and testing from a database consisting of enhanced underwater images. The selected images from the database were resized from a dimension of 320×240×3(width, height) to 224×224×3(width, height) to match the overall receptive field of the standard convolutional layers. As a preprocessing step, the input images of dimension 224×224×3 are initially normalized to [0,1]. The dataset consists of a total 100 images with 70 images for training and 30 images for testing. Using convolution and deconvolution filters a symmetric convolutional auto-encoder network is designed. Then, training is performed using mean squared error loss function and Adam optimizer.

Here, based on the standard CAE, an additional pseudo layer was introduced which takes the input data from the encoded layer. This layer attempts to slice the compressed data (i.e. from a dimension of 14×14×16 to 8×8×16, where 16 refers to the number of feature maps) there by creating some loss of data randomly. Therefore, the decoder attempts to reconstruct the image from this representation. Depending on the size of data loss, the decoder needs to be modified to match the size of the original image. Due to the above sperate decoder networks are required for reconstruction.

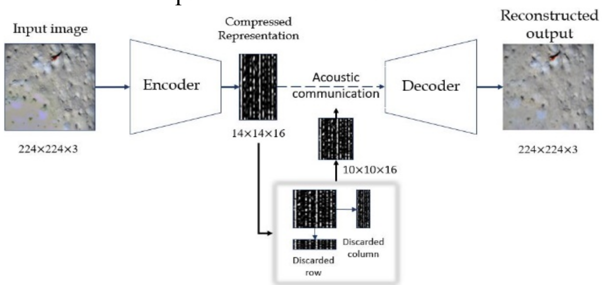

Fig. 3. Overview of the proposed image compression method.

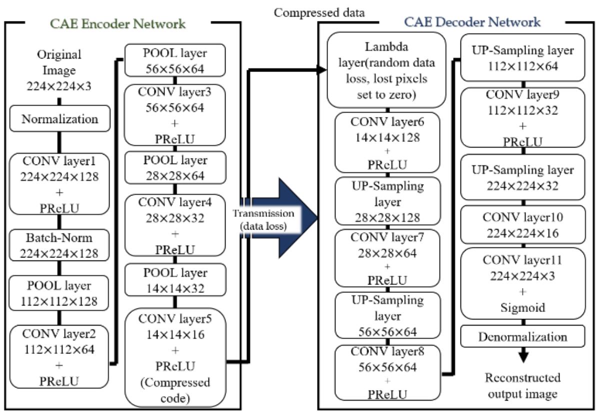

Fig. 4. CAE Network with data loss.

4. Experiments and Result

4.1 Experimental setup

A dataset consisting of 100 images was used for training and testing the proposed model. These images were obtained from previously conducted seafloor exploration experiments using Sampling-AUV: TUNA-SAND2. The number of filters in the convolutional layers were set as {128,64,64,32,16} and the decoder part mirrors the encoder. Mean square error loss function was used during training process to compute the distortion between original and reconstructed images. The parametric ReLU activation was employed in the network to improve the quality of reconstruction at high bit rate. For measuring the efficiency of the proposed image compression technique, the quality of reconstruction of images with and without data loss were compared using the quality metrics signal to noise ratio (PSNR) and structural similarity index(SSIM).

The simulations were performed on a PC with the following specifications Intel Core i7-8750H CPU at 2.40 GHz, 8GB RAM and GeForce GTX 1060Ti GPU. The CAE model was trained for 150 epochs i.e. 150 iterations over all the samples in mini-batches of 2. The running time i.e. the time taken for one complete encoding and decoding process for RGB images with resolution of 224 × 224 was approximately 3.60s/step.

4.2 Results

The Fig. 5 shows the comparison of PSNR at various compression ratios of original and reconstructed images. The proposed model was able to achieve a high signal to noise ratio which results in better quality of image. At higher compression ratios there is slight reduction in quality along with decrease in PSNR value.

Since the PSNR evaluation considers only numerical comparison of images, SSIM is also taken as a measure for evaluating the quality of the image. From the graph shown in Fig. 6 we can see that a maximum SSIM value of 0.78 for sample image 1 at a compression ratio of 18:1, the decoder is able to reconstruct the image with satisfactory image quality even with data loss in the compressed codes. The quality of reconstruction decreases at higher compression ratios. This is because when trying to reconstruct the discarded portion the autoencoder tries to add black and white color pixel intensities in place of lost pixels.

While evaluating the quality metrics in terms of luminance, contrast and structure of the reconstructed image with the original image considering the data loss, the results indicate that the proposed CAE model is able to provide better reconstruction quality and achieve an accuracy of 55 to 70%.

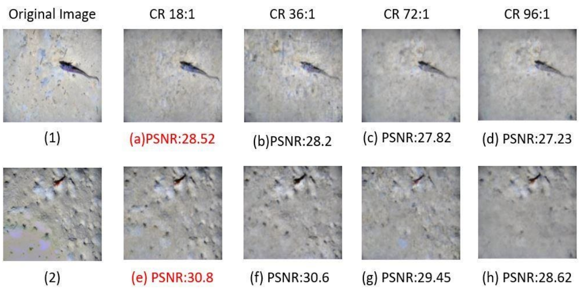

Fig. 5. Comparison of PSNR at different compression ratios with data loss.

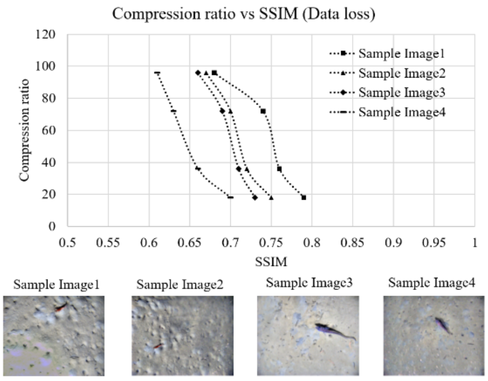

Fig. 6. Comparison of SSIM with data loss.

5. Conclusion

In this research, an efficient lossy image compression/reconstruction architecture using convolutional autoencoder was developed. The proposed CAE model tends to achieve better image reconstruction and perceptual quality with minimal distortion. The performance of the image compression algorithm was evaluated using quality metrics like PSNR and SSIM. It was observed from the results that the proposed autoencoder model was able to achieve fairly high signal to noise ratio with data loss providing a good reconstruction quality image.

References

1. Nishida, Y., Sonoda, T., Yasukawa, S., Nagano, K., Minami, M., Ishii, K., & Ura, T., Underwater Platform for Intelligent Robotics and its Application in Two Visual Tracking Systems, *Journal of Robotics and Mechatronics,* 30(2), 238-247, 2018.
2. Ahn, J., Yasukawa, S., Sonoda, T., Nishida, Y., Ishii, K., & Ura, T. An optical image transmission system for deep sea creature sampling missions using autonomous underwater vehicle, *IEEE Journal of Oceanic Engineering*, 2018.
3. Yasukawa, S., Nishida, Y., Ahn, J., Sonoda, T., Watanabe, K., & Ishii, K., Development and Sea Trials of Vision-Based Control for Sampling-AUV, *In 2019 IEEE Underwater Technology*, pp. 1-4, 2019.
4. Hoag, D. F., & Ingle, V. K., Underwater image compression using the wavelet transform. *In Proceedings of OCEANS'94*, Vol. 2, pp. II-533, 1994.
5. Cheng, Z., Sun, H., Takeuchi, M., & Katto, J., Deep convolutional autoencoder-based lossy image compression. *In 2018 Picture Coding Symposium (PCS)* , pp. 253-257, 2018.