Results in Physics 11 (2018) 96–104

Contents lists available at ScienceDirect

Results in Physics

journal homepage: www.elsevier.com/locate/rinp

Laser stripe image denoising using convolutional autoencoder☆

Zhuoqun Fang[a,1], Tong Jia[a,1], Qiusheng Chen[a], Ming Xu[a], Xi Yuan[a], Chengdong Wu[b,*]

[a] College of Information Science and Engineering, Northeastern University, Shenyang 110004, China
[b] Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110004, China

ARTICLE INFO

Keywords:
Convolutional autoencoder
Deep learning
Depth perception
Image denoising
Laser stripe
Structured light

ABSTRACT

Convolutional autoencoders are making a significant impact on computer vision and signal processing communities. In this work, a convolutional autoencoder denoising method is proposed to restore the corrupted laser stripe images of the depth sensor, which directly reduces the external noise of the depth sensor so as to increase its accuracy. To reduce the amount of training data and avoid overfitting, a patch size of the laser stripe image is determined, on the basis of which a small-scale dataset called Laser Stripe Image Patch (LSIP) is created. Also, a 14-layers convolutional autoencoder is constructed to reduce the noise of the image patches, which can learn the most salient features on the LSIP dataset. Moreover, the trained convolutional autoencoder is applied to an omnidirectional structured light system. Experimental results demonstrate that the proposed method obtains useful features and superior performance both visually and quantitatively on denoising tasks, and significantly improves the accuracy of the structured light system.

Introduction

Depth perception consists of active and passive methods. With the application and development of depth perception technologies in practice, active depth perception, compared with the passive one, attracts more and more attention in numerous fields. The structured light, as a mainstream technology of the active depth perception, uses laser emitters and cameras to capture laser stripe image which is used to calculate the depth. Also, many range-sensing devices based on structured light have emerged. However, in the indoor environment, the structured light systems suffer from noise due to the reflectance property of targets surfaces or disturbances from ambient light [1]. For example, metal surfaces have strong specular reflection, and they cannot be measured by standard laser scanners. In other words, the laser stripe images of these devices have some common drawbacks [2], especially missing data at points where the laser stripe is hidden because of object occlusion, and missing or erroneous data due to specularities. In view of this, laser stripe image denoising can reduce these noise so as to increase the accuracy of range data. In other words, laser stripe image denoising can be an effective method to improve the accuracy of structured light systems.

The algorithms applied to the laser stripe image denoising mainly contains two categories: filtering algorithm and morphology algorithm. Filtering algorithm [3,4] treats the laser stripe image as a grayscale image, so it adopts traditional filters such as the Gaussian filter, the median filter, and the mean filter. However, these filters can only remove some sort of noise caused by electronic devices. Morphological approach [5] first uses thresholding algorithms to convert the laser stripe image to a binary image; then processes the binary image with the closing operation. Although morphological operations can lessen different types of image noise, some of which caused by the object occlusion or specular reflection can be hardly removed. What is worse is that these approaches may produce unexpected errors when the images are processed under complex backgrounds. In conclusion, it is almost impossible to significantly reduce noise in laser stripe image using the traditional denoising algorithms.

The convolutional autoencoder (CAE) [6,7] is a deep learning method, which has a significant impact on image denoising. This provides an opportunity to realize noise reduction of laser stripe images. However, the CAE is rarely used in laser stripe image denoising. Worse still, noise reduction of the whole image needs a large amount of labeled data, but no dataset has been built for this purpose. These difficulties obstruct the application of the CAE in laser stripe image denoising.

☆ This work was supported in part by the National Natural Science Foundation of China under Grants 61701101, U1613214 and U1713216, in part by the National Key R&D Program of China under Grant 2017YFB1300900, and in part by the Shenyang Scientific Research Fund under Grant 17-87-0-00.

* Corresponding author.
E-mail address: wuchengdong@ise.neu.edu.cn (C. Wu).

[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.rinp.2018.08.023
Received 6 June 2018; Received in revised form 9 August 2018; Accepted 10 August 2018
Available online 24 August 2018

In this paper, we propose a denoising method for laser stripe image using the CAE. Instead of creating a dataset of large size images, we establish a dataset that contains smaller image patches of the laser stripe image. Then, we design a CAE and train the neural network based on our created dataset. Additionally, the CAE is applied to an omnidirectional structured light system using an image slicing method. Experimental results show that the CAE outperforms other traditional algorithms in image denoising and provides high-precision range data.

The rest of this paper is organized as follows. Section “Background” discusses the theoretical background of the structured light technology. The content concerning the CAE, including the dataset creation and the network architecture, is detailed in Section “Convolutional autoencoder”. Section “Depth perception” describes the organization of our omnidirectional structured light system and the means to adopt the CAE to it. Experimental results are presented in Section “Experiments and analysis”. Section “Conclusion” concludes the paper with a brief summary.

Background

Depth perception or range-sensing is the visual ability to perceive the world in 3D and measure the distance of an object. It can be divided into two categories: the active method and the passive method. Contrary to the passive range-sensing, the active one adopts a source of controlled illumination to scan the surface so as to acquire the depth. The active depth perception contains time-of-flight approach, structured light approach, etc. According to the type of light pattern that it projects, structured light technologies can be classified into spot-, line- and 2D-structured light. In this section, we focus on the line structured light [8] and discuss the fundamental principle of acquiring depth maps.

The line structured light system projects a laser stripe on the scene and captures the corresponding image to construct dense and accurate depth maps as shown in Fig. 1. The laser beam can be transformed by a cylindrical lens into a light plane, on which a stripe will appear if it strikes a surface. In this case, the point where the laser beam strikes the surface, the bright spot P, is found as the intersection of the beam with the projection ray joining the spot to its image [2]. Therefore, the depth of the spot P can be calculated by triangulation [9,10]. Similarly, a whole image column of the depth map can be acquired at each frame by using the laser stripe. Therefore, we can use a laser and a rotating mirror to sequentially scan the surface. It should be noted that this setup will not cause matching ambiguities since the laser spot associated with an image pixel can be retrieved as the unique intersection of the corresponding projection ray with the plane of light [2].

However, the main drawbacks of the structured light technology are missing data at the point where the laser stripe is hidden from the camera by the object itself and missing or erroneous data due to specularities. The latter is actually a common difficulty of all active ranging techniques: a purely specular surface will not reflect any light in the direction of the camera unless it happens to lie in the corresponding mirror direction. Worse, the reflected beam may induce secondary reflections, which will give false depth measurements.

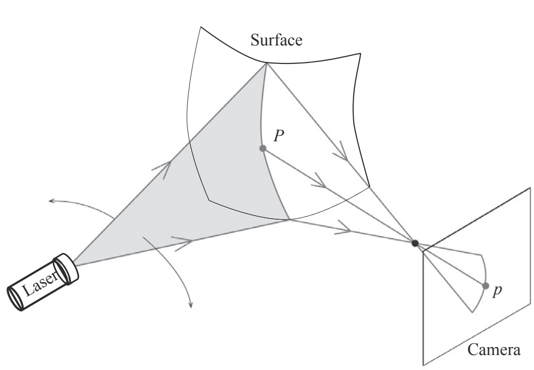

Fig. 1. Schematic diagram of line structured light.

Convolutional autoencoder

The autoencoder (AE) [11] model is based on an encoder-decoder paradigm. The encoder first transforms the input into a typically lower-dimensional representation, and the decoder is tuned to reconstruct the initial input from this representation through the minimization of a loss function [12]. A CAE contains convolutional layers in the encoder part and transposed convolutional layers in the decoder part. CAEs can better meet the need of the image processing tasks because they fully utilize the properties of convolutional neural networks (CNNs), which have been proved to provide better results on noisy image data. To improve the accuracy of the line structured light system, we introduce the CAE to denoise the laser stripe image. First, we define a patch of laser stripe image which is used as input and output of the autoencoder. Second, we create a dataset based on it to train the CAE. Third, we design the network architecture of the CAE. Finally, we update the weights of the neural networks using our dataset.

Patch size

There are a variety of noise sources in a laser stripe image, which makes the image denoising complicated and difficult, even for the CAE. Although the CAE has a huge potential to reduce the noise of images, taking a large size image as the network input amplifies the computation complexity. In this case, it increases the risk of overfitting if we do not use a mass of data. Worse still, the image size may be different, which may prevent the CAE from generating the correct results. Therefore, instead of running the CAE on the whole image, we use image patches as the input and output data of the CAE.

To simplify the image denoising process, we decide to start from the perspective of image patches, where the noise is generally reduced to three categories, namely, specularity, intermittent stripe, and noisy background as shown in Fig. 2a, b, and c. Therefore, the focus of the CAE will be on these types of noise. In other words, the image patch simplifies the capacity of the CAE, which makes it possible to train the CAE networks based on a smaller dataset. Furthermore, more data is provided by segmenting a large laser image into several image patches, which satisfies the demand for data volume in neural network training. Also, we can determine the proportion of different types of data to prevent generating skewed data and improve the ratio of valid data, which makes the denoising algorithm more robust. On the other hand, as discussed above, because the local information is more important than the global information in our case, the patch-wise denoising algorithm does not break the continuity of the whole image, which is proven by the experimental results in Section “Experiments and analysis”.

In our case, the original laser stripe image is segmented into square patches without overlap. In other words, we slice up the original image and extract contiguous but non-overlapping patches of size 100×100 from it. As shown in Fig. 2, there is at most one stripe in each patch. The patch is designed to facilitate the denoising task and make the goal more clear, through which the CAE predicts a blank or a laser stripe with a blank background in the patch as the denoising result.

Dataset

To train a CAE, we design and create a small-scale dataset called Laser Stripe Image Patch (LSIP), with examples shown in Fig. 2. In the image denoising task, the CAE is trained to receive a corrupted or noisy image as input and generate an uncorrupted image as output so that the original image is predicted. Therefore, LSIP contains 520 pairs of image patches, each of which is composed of a corrupted 8-bit grayscale image