IMAGE DEPTH EVALUATION SYSTEM BY STREAM VIDEO

Sych O. S.
National Aviation University, Kyiv
Scientific supervisor - Vasylenko M.P., senior lecturer

One of the data processing applications is stereo vision, in which obtaining a three-dimensional scene is based on models for determining the depths of key points of images from a video sequence or several images. If it is considered an example with a person, then a two-dimensional image is formed on the retina, but despite this, a person perceives the depth of space, that is, has three-dimensional, stereoscopic vision. As a result, in the presence of data on the size of an object, it can be estimated the distance to it or understand which of the objects is closer. When one object is in front of the other and partially obscures it, the person perceives the front object at a closer distance. Because of this, the need arose to teach machine devices to do this for various tasks. Based on the processing results, you can have spatial information for assessing the relief, obstacles while driving, etc.

This algorithm is based on combining images of the same object, photographed or filmed on video with constant camera parameters and in the same focal plane from different angles, allows to obtain information about the distance to the object by perspective distortions (discrepancies).

There are different classifications of models in the matching of two images. Models are divided, depending on the algorithms used in them, into local (in which the discrepancy is calculated at the control points based on the “similarity” of the window around this point and the window around the points in another image) and global (find the discrepancy for all points at once).

One of the fastest local mathematical methods is the depth map method.

To solve this problem, image data from two cameras were used, which in turn significantly increased the accuracy of finding points, after which the distance to a certain pixel, that is, the main idea of the method is that similar scenes should have approximately the same depth distributions. The use of such preprocessing significantly reduces the probability of an error in calculating the pixel depth, especially if there are homogeneous, low-contrast areas on a stereo pair. In other words, scene images have similar depth values in areas with similar appearance. Since not all obtained estimates will be correct, several candidates for the image (maps from two cameras) are found to refine and interpolate these estimates using global optimization.

Thus, the algorithm of image depth estimation (Fig. 1) is presented:

1. Start - preparation of technical equipment. According to this method, two webcams were used to shoot RGB images with resolution of 1280x720, a Raspberry Pi single-board computer, an Arduino Nano microcontroller, and a Matlab application package for solving the recognition problem.

2. Loading an image from a video and a corresponding depth map taken from two cameras. In this case, on the second camera, you can remove the so-called Bayer

filter (it ensures that the light falling on each pixel has one of the three primary colors) from the matrix, i.e. one camera will be black and white and the other color. This increases the sensitivity for black and white by about 3 times. Those. the sensitivity of 2 cameras grows conventionally not in 2 times, but in 4.

3. Filters an image with a Gaussian filter to reduce noise and improve clarity, which uses a normal distribution (also called a Gaussian distribution) to compute a transform applied to each pixel in an image.

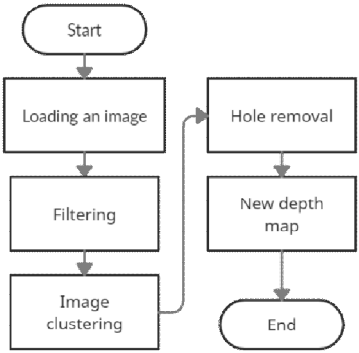

Fig. 1. Algorithm of image depth estimation

4. Image clustering for grouping objects. The main point is that all the original objects (in this case, pixels) are divided into several non-intersecting groups in such a way that the objects that fall into the same group have similar characteristics, while for objects from different groups these characteristics should be significantly different.

5. Removing holes in the depth map (closed areas). Holes cause serious problems in applications that use depth sensors and often cover the background and subject at the same time. Accordingly, the hole must be filled taking into account the boundary of the object.

6. Creation of a new depth map for use of an orientation purposes, namely, it allows to capture control points of a moving object to obtain three-dimensional coordinates from the MATLAB environment.

Based on the above, it can be concluded that the main algorithm is suitable for the tasks of recognizing objects and calculating the distance to them, it is worth noting that this method can also be configured for calculating the size of a plane that is of direct interest to the user, but you also need to understand that this is possible only with a frontal image, since for structuring volumetric models, it is forced to shoot the entire area at 360 degrees.

References:

1. Depth map generation for 2d-to-3d conversion by short-term motion assisted color segmentation/Yu-Lin Chang, Chih-Ying Fang, Li-Fu Ding, Shao-Yi Chen, and Liang-Gee Chen - DSP/IC Design Lab, Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan

2. Scharstein D., Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms // Int. Journal of Computer Vision 47. April-June 2002. PP. 7–42.

3. Разработка и исследование алгоритма вычисления карты глубины стереоизображения/ В.В. Воронин.

4. Метод оценки глубины сцены и текстуры невидимых частей изображения URL: https://neurohive.io/ru/papers/pokazat-to-chto-skryto-metod-ocenki-glubiny-i-nevidimyh-chastej-izobrazhenij/ (Last accessed: 11.01.2021).