# Depth video enhancement for 3D displays

**Ljubomir Jovanov***, **Hiêp Quang Luong***, **Wilfried Philips***

*Ghent University-TELIN-IPI-iMinds, Belgium

**ABSTRACT**

*At the current stage of technology, depth maps acquired using cameras based on a time-of-flight principle have much lower spatial resolution compared to images that are captured by conventional color cameras. The main idea of our work is to use high resolution color images to improve the spatial resolution and image quality of the depth maps.*

## 1. INTRODUCTION

The introduction of the third dimension into the scientific visualization, computer aided design (CAD) and electronic entertainment was a great leap forward, compared to the earlier two-dimensional systems. While in CAD and scientific visualization user creates 3D content (e.g. 3D drawing in CAD software), in electronic entertainment systems 3D geometry of the scene often has to be captured using specialized sensors.

At the current stage of development 3D video capture techniques are still not mature enough to produce broadcast quality of the content. Most often, 3D video is captured using two cameras or a multi-camera setup. However, these systems are not flexible enough since their physical size and transmission bandwidth increases with the required number of views. Another, more recent possibility are 3D sensors based on the time-of-flight principle (ToF). ToF cameras are based on measuring time needed for modulated infrared light to travel to the scene and back by estimating phase. These sensors capture texture-like information simultaneously with depth images, up to 50 frames per second. The main drawbacks of ToF sensors are their inherently low spatial resolution and high level of signal dependent noise, as shown in Figure 1.

It is beneficial to use hybrid setups, which contain both classic color cameras and ToF sensors and take the best characteristics of both. In this paper we present a method for denoising and upsampling of depth video sequences using the high-resolution color image as a guidance map. Depth images are then merged with color images from a classical video camera and shown on a auto-stereoscopic 3D display.

## 2. METHOD

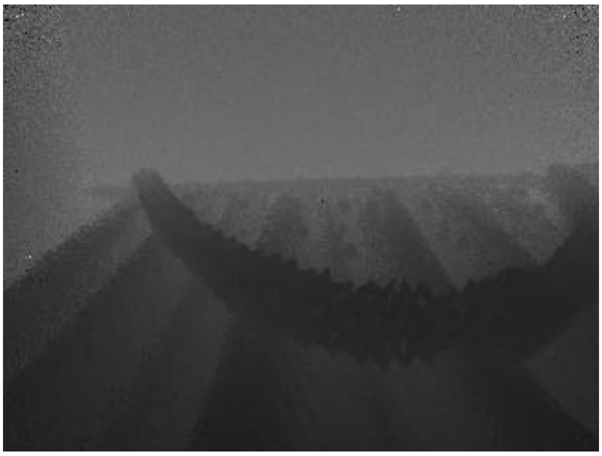

**Figure 1: Low resolution depth map**

Fusion of images acquired using two or more cameras with different characteristics is a challenging problem. In practice cameras have different resolutions, viewing angles, intrinsic parameters and certainly different viewpoints. To perform registration of these images it is necessary to determine the parameters of both cameras and use these to compensate the differences in parameters of the camera pair. First we perform stereo calibration using the algorithms from [1] and [2]. For the calibration of the depth camera we use reflection intensity images, since these contain texture information. Since the resolution of the depth camera is quite low we have used a larger calibration pattern than in normal circumstances and placed it closer to the depth camera to ensure that the squares in the calibration pattern containing sufficient number of pixels to prevent large calibration error. As a result of the calibration, we obtain intrinsic parameters of both cameras, i.e. focal lengths, principal point coordinates and distortion coefficients of camera lenses. Additionally, the calibration estimates rotation matrix and translation vector, which determine mutual positions between the two cameras (i.e. extrinsic parameters).

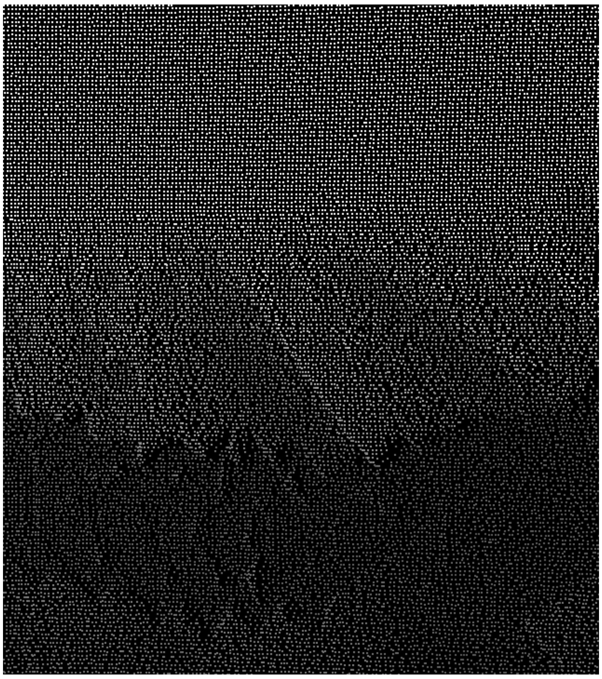

**Figure 2: Upsampled depth points mapped to the high-resolution color image**

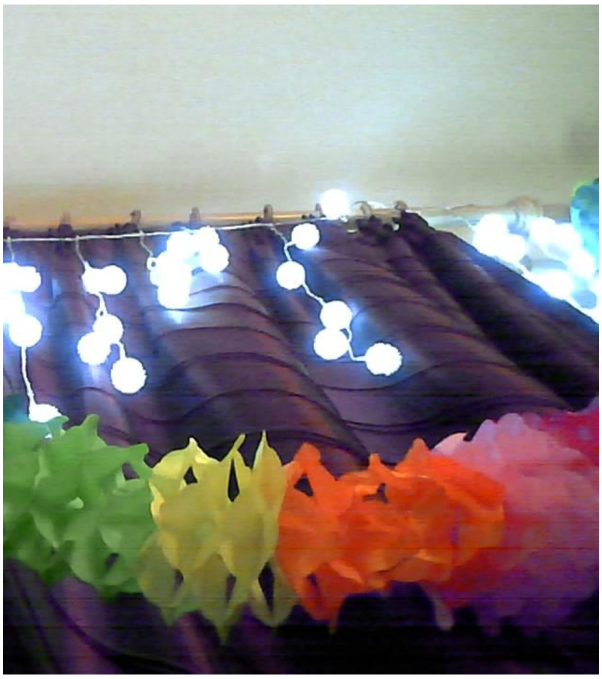

**Figure 3: High-resolution color image used as guidance for upsampling**

Calibration parameters are then used to map 3D points obtained using the low-resolution depth camera to the high-resolution color image. In the first step, we apply rotation, defined by the matrix R, and translation, defined by the vector t, to the 3D points defined by the coordinates (X,Y,Z) to map it to the coordinate system of the color camera, resulting in the coordinates (x,y,z). In the following we apply distortion coefficients to the coordinates (x,y,z), and finally, we project it into the image plane of the color camera.

Since depth image in our case contains twelve times less pixels, and has larger viewing angle, after the registration, depth points are sparsely distributed over the color image. Moreover depth points are not uniformly distributed over the color, due to the different depths which create different levels of disparity in the re-projected depth image. Due to the stereo effect some of the depth points would be associated to the pixels that are normally not visible in the color image.

Upsampling based only on co-registered depth points, would produce depth maps of lower quality, due to non-uniformity of points and occlusions due to the stereo effect.

To improve the quality of the upsampled depth image we propose an adaptive method, which relies on a higher resolution color image to improve the quality of the upsampled depth. The proposed algorithm traverses all locations in the sparse depth image, where depth values are not defined (dark areas in Figure 1), and calculates new, interpolated depth values based on the existing depth samples and color values which exist at the positions where depth points are not defined.

Interpolated depth value is obtained as a weighted sum of the existing depth samples within some spatial neighborhood of the currently interpolated depth sample. Interpolation weights are derived from the sums of pixel differences within the patch around the currently upsampled location and the location in the neighborhood.

Since depth points are spatially non-uniformly distributed, using a neighborhood of a constant size would result in different number of neighboring samples used for interpolation, which would produce in large deviations from the actual value of depth at the current location. To prevent this from happening, we propose the use of variable size neighborhood. We start with some minimal size of the neighborhood and increase it successively until some predefined number of nonzero depth samples is reached. Moreover, we calculate interpolation weights only for those locations in the color image corresponding to locations with defined depth values. This results in a large complexity reduction, since we calculate interpolation weights only for 10% of pixel in the color image.

### 3. RESULTS

We have experimentally evaluated the proposed algorithm on a sequences recorded using a DS 325 sensor [3], which consists of the color camera with 1280x720 resolution and the depth sensor with 320x240 resolution.

The result of the co-registration of the depth to the color image is shown in Figure 2. Due to the limited space we show only the right half of both images. As one can notice, depth values are correctly mapped to the color image, since they follow the shapes of the objects in the scene. Such co-registration would not be possible using classical 2D registration methods, since the mapping function changes locally due to the different values of depth. Classical 2D methods would produce a large mismatch for objects, which are close to the camera due to the larger local disparity.

The final result of our upsampling method is shown in Figure 4. Resulting depth map follows the borders of the object in the scene, providing each color pixel with the depth value. As one can notice, the proposed method does not introduce any depth artifact originating from the color guidance. This image is then merged with the color image, and used as an input for Philips auto-stereoscopic lenticular screen. Viewing of upsampled images on an auto-stereoscopic display confirmed again that the registration was done correctly.

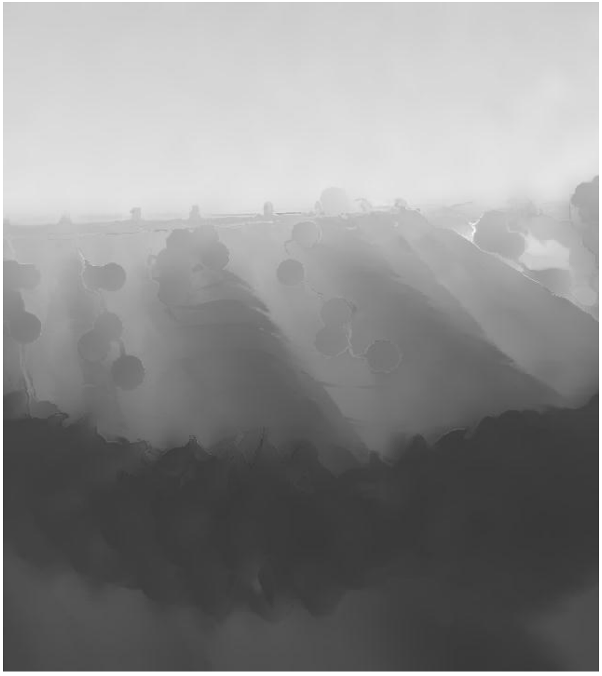

**Figure 4: Resulting high-resolution depth map**

## 4. CONCLUSION

In this paper we have presented a method for upsampling of low resolution depth map using higher resolution color images. The proposed algorithm involves calibration of the camera pair and co-registration using these parameters. Finally, we perform color guided depth image interpolation and display the resulting textured depth map on an auto-stereoscopic 3D display.

The proposed method significantly improves the visual quality of the depth map, and perfectly matches the color image.

## ACKNOWLEDGEMENT

This work was funded by iMinds and IWT through the project 3DTV 2.0.

## REFERENCES

[1] Z. Zhang. “A flexible new technique for camera calibration”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.

[2] Heikkilä, J, "Geometric Camera Calibration Using Circular Control Points", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, pp. 1066-1077, Oct 2000.

[3] Soft Kinetic DS 325 depth camera http://www.softkinetic.com/solutions/depthsensecameras.aspx