Using ToF camera and two mirrors for 3D reconstruction of dynamic objects

Technical Report Number 1380

Trong-Nguyen Nguyen\textsuperscript{1}, Huu-Hung Huynh\textsuperscript{1,2}, and Jean Meunier\textsuperscript{1,3}

\textsuperscript{1}DIRO, University of Montreal, Montreal, QC, Canada
\textsuperscript{2}University of Science and Technology, Danang, Vietnam

September 8, 2018

Abstract

Vision systems are currently playing an important role in a wide range of applications. Many recent systems attempted to represent the object of interest as a 3D model instead of a sequence of 2D images. The task of 3D reconstruction is usually performed via a collection of color cameras and/or depth sensors. Such approaches thus require a synchronization protocol and a significant budget for device purchase. This report provides an overview of our system that employs only one depth camera together with two mirrors to perform the 3D reconstruction. The device cost is thus reduced and the problem of synchronization can also be avoided. Our approach is appropriate for providing intermediate real-time 3D models of dynamic objects with acceptable quality for practical applications.

1 Context

In order to obtain more useful features for describing an object, vision researchers attempted to perform the task of 3D reconstruction as a preprocessing step. Therefore many 3D reconstruction approaches have been proposed. The principle of 3D reconstruction is the combination of object-related information captured under different view points. There are various systems dealing with that idea such as employing a depth camera together with a turntable [6], shape-from-silhouette, using a collection of depth sensors [1] and/or stereo cameras [2]. These systems have their own limitations such as working with only static

\textsuperscript{*}nguyetn@iro.umontreal.ca
\textsuperscript{1}hhung@dut.udn.vn
\textsuperscript{2}meunier@iro.umontreal.ca
objects, reconstructed models containing redundancy, requiring a synchronization protocol, or high prices. Our approach is thus proposed to overcome these problems.

Our system consists of a depth camera and two mirrors that are placed around the object\(^1\). Each depth map captured by the camera thus contains the object surfaces seen from 3 different view points. There are two common depth estimation schemes: Structured Light (SL) and Time-of-Flight (ToF). Our system using a SL depth camera has been presented in [5]. This report briefly describes the other camera type (ToF) where the reconstructed point cloud has more details compared with the work [5] but requires an additional processing to reduce depth distortion. The depth distortion occurs due to unwanted multiple reflections with mirrors. For further information on this work, the reader can refer to the original paper [4].

2 Algorithms

The reconstruction algorithm attempts to provide a point cloud obtained by combining (1) the object surface that is directly seen by the camera and (2) two reflections of the two other surfaces that are behind the two mirrors. A post-processing is then performed to remove unreliable points, i.e. 3D points having wrong depth measurement due to multiple reflections [3, 4]. Since this algorithm works directly on the pixels of each captured depth map, this can be integrated into real-time systems and is thus appropriate to work with dynamic objects. The details of this algorithm are presented in Algorithm 1, in which \(pts\) is the depth map provided by the camera that contains only object points, \(mir_1\) and \(mir_2\) are two mirror plane equations, \(C\) is the camera position, and \(reg\) is a predefined 3D region containing the object.

3 Experiments

In order to assess our approach, we performed the proposed 3D reconstruction algorithms on easy-to-evaluate geometric objects. We started with an experiment using only one mirror to demonstrate our hypothesis on mirror-related depth distortion as well as our solution (i.e. Algorithm 1) that reduces such distortions. A flat board of size 30\(cm\) \(\times\) 30\(cm\) was employed in this experiment. In the demonstration with two mirrors (our proposed system), we used a cylinder with radius of 15\(cm\) together with the flat board to assess our reconstruction accuracy. The Root Mean Squared Error (RMSE) calculated on the collection of reconstructed 3D points and its fitted shape was used to indicate the reconstruction error. Table 1 and 2 present such errors in our experiments under various settings (in term of geometric relation between an object and mirrors).

\(^{1}\)A different number of mirrors can be used if needed.
Algorithm 1: Reconstructing raw point cloud (appropriate for dynamic objects)

1: procedure GetCloud(pts, mir1, mir2, C, reg)
2:     cloud ← null
3:     P(3D) ← reproject(pts) \(\rightarrow\) 2D to 3D reprojection
4:     for each point P in P(3D) do
5:         if P inside reg then
6:             cloud ← push(P)
7:         else if P behind mir1 then
8:             Pr ← reflect(P, mir1)
9:             if Pr not inside reg then
10:                continue \(\rightarrow\) check another point
11:         end if
12:         P2 ← reflect(Pr, mir2)
13:         if CP < CP2 then
14:             cloud ← push(Pr) \(\rightarrow\) reliable point
15:         end if
16:     else if P behind mir2 then
17:         Pr ← reflect(P, mir2)
18:         if Pr not inside reg then
19:                continue \(\rightarrow\) check another point
20:         end if
21:         P1 ← reflect(Pr, mir1)
22:         if CP < CP1 then
23:             cloud ← push(Pr) \(\rightarrow\) reliable point
24:         end if
25:     end if
26: end for
27: return cloud \(\rightarrow\) Return object point cloud
28: end procedure

Table 1: Reconstruction errors (in centimeter) when using 1 mirror. The term angle indicates the geometric angle (in degree) between the board and the mirror surface, measured and enhanced correspond to distance between ground truth points (determined directly by the camera SDK) and reconstructed points before and after performing our solution, respectively.

<table>
<thead>
<tr>
<th>Angle</th>
<th>91.13°</th>
<th>89.22°</th>
<th>87.82°</th>
<th>86.26</th>
<th>84.70</th>
<th>83.46°</th>
<th>82.50°</th>
<th>81.50°</th>
<th>77.44°</th>
<th>73.74°</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>91.29</td>
<td>93.03</td>
<td>93.22</td>
<td>93.80</td>
<td>94.63</td>
<td>94.71</td>
<td>96.22</td>
<td>98.45</td>
<td>96.42</td>
<td>98.49</td>
</tr>
<tr>
<td>Enhanced</td>
<td>02.22</td>
<td>01.94</td>
<td>02.03</td>
<td>01.71</td>
<td>01.77</td>
<td>01.87</td>
<td>01.73</td>
<td>01.78</td>
<td>01.83</td>
<td>01.59</td>
</tr>
</tbody>
</table>
Table 2: Reconstruction errors (in centimeter) when using a system of 2 mirrors. The term distance indicates the average distance between the object and the two mirrors.

<table>
<thead>
<tr>
<th>Board</th>
<th>Distance</th>
<th>60.70</th>
<th>67.06</th>
<th>76.12</th>
<th>82.09</th>
<th>86.74</th>
<th>91.78</th>
<th>97.27</th>
<th>101.50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Naive reflection</td>
<td>0.209</td>
<td>0.242</td>
<td>0.250</td>
<td>0.243</td>
<td>0.216</td>
<td>0.212</td>
<td>0.202</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>Algorithm 1</td>
<td>0.196</td>
<td>0.215</td>
<td>0.197</td>
<td>0.187</td>
<td>0.193</td>
<td>0.180</td>
<td>0.180</td>
<td>0.158</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cylinder</th>
<th>Distance</th>
<th>53.96</th>
<th>62.46</th>
<th>65.79</th>
<th>70.82</th>
<th>74.66</th>
<th>78.50</th>
<th>84.26</th>
<th>90.57</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Naive reflection</td>
<td>0.446</td>
<td>0.430</td>
<td>0.374</td>
<td>0.352</td>
<td>0.404</td>
<td>0.342</td>
<td>0.399</td>
<td>0.541</td>
</tr>
<tr>
<td></td>
<td>Algorithm 1</td>
<td>0.359</td>
<td>0.384</td>
<td>0.285</td>
<td>0.324</td>
<td>0.371</td>
<td>0.333</td>
<td>0.387</td>
<td>0.499</td>
</tr>
</tbody>
</table>

According to Table 1, the depth measured on distorted points were significantly incorrect since the average distance between them and the ground truth was nearly 1 meter. By applying our Algorithm 1 that performs unreliable point removal, the measurement error was reduced about 53 times (95.9 cm compared with 1.8 cm).

Similarly, Table 2 shows that our procedure of unreliable point removal enhanced reconstructed point clouds since it decreased the reconstruction errors.

As indicated before, our 3D reconstruction approach is to provide intermediate 3D models with an acceptable quality in practical applications. An example of point cloud representing a 3D human body using the Algorithm 1 is presented in Fig. 1. This method is currently applied in our works on gait analysis.

Figure 1: Example of a point cloud obtained in our setup for an application of gait analysis.

If the researchers use our reported information in their works, please refer this document as

References


