Depth Propagation with Key-Frame Considering Movement on the Z-Axis

Jin Woo Choi¹, Taeg Keun Whangbo²

¹ Culture Technology Institute, Gachon University, 1342 Seongnamdaero, Sujeong-gu, Seongnam-si, Gyeonggi-do, 461-701, Korea
  cjw49@paran.com
² Department of Computer Engineering, Gachon University
  tkwhangbo@gachon.ac.kr

Abstract. We propose a key-frame-based depth propagation method considering movement on the z-axis. First, a homography matrix is obtained through the feature points between points. Through this homography matrix information, a movement of objects or cameras is inferred in the direction of the z axis, thereby creating a depth map that compensates the movement. Then, bilateral filtering is executed to create a depth map with regard to the non-key frames. Through the experiment, a more accurate depth map can be obtained with the non-key-frame-based method due to the consideration of changes in the z-axis compared to the key-frame-based depth propagation method.

Keywords: Depth propagation, 2D-to-3D conversion

1 Introduction

Although hardware technology for stereoscopic 3D use has advanced, the stereoscopic 3D industry has slowed due to the lack of 3D content. As an alternative technology, much attention has been paid to two-dimensional (2D)-to-3D conversion technology in which existing 2D content can be converted to 3D content. 2D-to-3D conversion can be categorized as full-automatic, semi-automatic, and full-manual conversion depending on quality and cost, objectives, and processing time [1, 2]. Among these, semi-automatic 2D-to-3D conversion has been widely used by most content production companies because it allows automatic processing, if possible, and appropriate intervention of operators, if required, to maintain high quality. A depth map of key-frames is created by an operator because movements between subsequent frames in a video are considered very low while non-key-frames are processed automatically by a depth propagation method [3-5].

However, existing key-frame-based depth propagation methods cannot accurately reflect the information from changes in depth, which is the movement of cameras or objects in the direction of the z-axis. Thus, we propose a novel depth propagation technique considering movements in the z-axis by expanding the study of [5].
2 Proposed Depth Propagation and Experimental Technique

Figure 1 shows a flow chart of the proposed method. First, feature points of segmented objects in the previous key-frame and feature points in the current frame are extracted and matched. A homography matrix is inferred through these matched feature points, thereby allowing calculation of the scale of the objects and perspective information. Then, a normalized value of the above result is reflected in a depth map of the key-frame prior to performing bilateral filtering, thereby compensating the z-axis movement when a depth is propagated to the next frame. This process is iteratively performed if there are a number of objects inside a video. Finally, the depth of the current frame is generated by applying a bilateral filter to the previous frame depth in which depth was compensated by the color difference between two consecutive frames taking into consideration a motion compensation depth.

2.1 Feature points extraction and matching

Figure 2 shows two feature point matching results through SURF: (a) a case where z-axis movement occurred in the overall image due to camera movement and (b) a case where z-axis movement occurred due to object movement. In Fig. 2(b), segmentation information from the previous frame was used as a mask so that features only within the object area were extracted and matched. The Hessian coefficient for feature point extraction was set to 400, while points less than 0.2 of Euclidean distance were used for criteria of similarity between feature points.
2.2 Homography matrix estimation and z-axis movement calculation

Using the relationship between the matched feature points, a homography matrix was calculated by means of the random sample consensus (RANSAC) algorithm [7,8]. The homography matrix is represented by a 3 x 3 matrix \( H \) that expresses a 1:1 mapping relationship between two frames.

\[
H = \begin{bmatrix}
h_{11} & h_{12} & h_{13} \\
h_{21} & h_{22} & h_{23} \\
h_{31} & h_{32} & h_{33}
\end{bmatrix}
\]  

(1)

The degree of changes in the scale and perspective can be inferred by analyzing the homography matrix. Equation (2) is used to calculate a scale factor of the x-axis (Sx), a scale factor of the y-axis (Sy), and a perspective factor (P).

\[
S_x = \sqrt{h_{11}^2 + h_{21}^2}, \quad S_y = \sqrt{h_{12}^2 + h_{22}^2}, \quad P = \sqrt{h_{31}^2 + h_{32}^2}.
\]  

(2)

In the case of Sx and Sy, a value less than 1 means reduction and a value more than 1 means expansion. In the case of P, 0 means no change as perspective information. Figure 1(a) shows calculations using 1.08, 1.12, and 0.0004, respectively, while Fig. 1(b) uses 0.85, 0.50, and 0.0012, respectively. If the P value is too large (normally more than 0.002), it is regarded as abnormal, so P can be used as a value that measures the reliability of homography.

2.3 Compensation of z-axis movement and depth propagation using a bilateral filter

The degree of movement of the camera or object in the direction of the z-axis between two consecutive frames is normalized to a depth value, as shown in Equation (3), using the scale change obtained from Equation (2)

\[
M_z(x, y) = \begin{cases} 
  k \cdot \left(\frac{S_x + S_y}{2}\right), & \text{if } \frac{S_x + S_y}{2} > 1 \\
  0, & \text{if } \frac{S_x + S_y}{2} = 1 \\
  -k \cdot \left(\frac{S_x + S_y}{2}\right), & \text{otherwise}
\end{cases}
\]  

(3)

The \( M_z \) value is added to the depth of the corresponding object area in the previous frame so that the compensated depth can be used in the subsequent bilateral filtering process. The depth conversion coefficient \( k \) value is set to 50 in the cases shown in Fig. 2(a) and (b) (deliberately set to a high value to verify the result) so that a compensated previous frame depth map is shown in Fig. 3(4). Figure 3(5) shows the final result from performed bilateral filtering and motion compensation expressed by Equation (4), which represents a propagated depth of the current frame.

\[
D^t(x, y) = \frac{\sum_{i=-N}^{N} \sum_{j=-N}^{N} e^{-\alpha[\|C^t(x, y) - C^{t-1}(x+i, y+j)\|]} \cdot (D^{t-1}(x, y) + M_z(x, y))}{\sum_{i=-N}^{N} \sum_{j=-N}^{N} e^{-\alpha[\|C^t(x, y) - C^{t-1}(x+i, y+j)\|]}},
\]  

(4)

where \( D^{t-1} \) and \( D^t \) are depth maps of the previous and current frames, respectively, \( C^{t-1} \) and \( C^t \) are color images, \( N \) is a filter size, and \( \alpha \) is a constant that represents the importance of color. Here, a motion compensation method was applied using a pair of bilateral methods [5] to solve the “depth ambiguity” and “new color” problem [4].
which is generated when depth propagation is executed using the bilateral filter.

3 Conclusions

The present study aimed to resolve the problem found in existing depth propagation algorithms of not being able to reflect movement in the direction of the z-axis. Once a homography matrix was obtained and analyzed using the feature point information that is matched between two consecutive frames, a size of scale change is identified; this information is then reflected in the existing bilateral filter. The experiments with images of both camera and object movement showed satisfactory results. In future research, a more advanced algorithm that can accurately reflect even partial movements of unstructured objects will be studied.

Fig. 3. Experimental results: (1) previous color; (2) current color; (3) previous depth; (4) depth after depth compensation of (3); (5) estimated current depth

Acknowledgments. This research was a part of the project titled “3D Scene Analysis and Model Reconstitution Techniques in Stereoscopic 3D Creation and Synthetic Techniques”, supported by National IT Industry Promotion Agency(NIPA) and Korea Creative Content Agency(KOCCA) grant funded by Ministry of Science, ICT and Future Planning(MSIP) and Ministry of Culture, Sports and Tourism(MCST)

References


