A Study on Optimized Mapping Environment for Real-time Spatial Mapping of HoloLens

Leehwan Hwang[1], Jaehyun Lee[1], Jahanzeb Hafeez[1], Jinwook Kang[2],
Seunghyun Lee[1], Soonchul Kwon[1*]

[1]*Graduate School of Information and Contents, Kwangwoon University, Seoul, Korea*
[2]*Geminisoft Co., Ltd., Korea*
hlh0003@hanmail.net, noa6142@gmail.com, jahanzeb@kw.ac.kr, jay@gemini-soft.com
{shlee, [*]ksc0226}@kw.ac.kr

Abstract

Recently, the development of the head mounted display (HMD) device has attracted a great deal of attention to the actual contents. Especially, Augmented Reality (AR), which is a mixture of actual information and virtual world information, is focused on. AR HMD is able to interact by arranging virtual objects in real space through spatial recognition using depth camera. In order to naturally mix virtual space with real space, it is necessary to develop a technology for realizing spatial mapping information with high accuracy. The purpose of this paper is to evaluate the optimal configuration of augmented reality application program by realizing accurate spatial mapping information when mapping a real space and an object placement environment using HoloLens. To do this, we changed the spatial mapping information in real space to three levels, which are the number of meshes used in cubic meters to scan step by step. After that, it was compared with the 3D model obtained by changing the actual space and mesh number. Experimental result shows that the higher the number of meshes used in cubic meters, the higher the accuracy between real space and spatial mapping. This paper is expected to be applied to augmented reality application programs that require scanning of highly mapped spatial mapping information.

Keywords: Augmented Reality, Head Mounted Display, Spatial Mapping, HoloLens

1. Introduction

Recently, as Head Mounted Display (HMD) technology has been developed, realistic contents are actively utilized. Augmented Reality (AR), which combines virtual information with reality is an effective technique for realizing realistic contents. This Augmented Reality overlays the virtual object in real world we see in real coordinates [1]. Therefore, it is important that the augmented reality is arranged by adding real environment and virtual information in real time. In order to place virtual information in reality, the HMD must acquire accurate real space spatial mapping information. Microsoft's HoloLens launched in 2016, is a typical AR HMD device that can show AR environment. HoloLens have the potential to be able to converge with the industry as a diverse form of viewing and enjoying, interacting with virtual objects as well as excelling in 3D representation [2]. HoloLens collects the mapping information by scanning the object precisely and obtains the spatial information of real space by using the TOF depth camera for interaction in

the mixed reality [3]. As soon as the power is turned on, HoloLens starts scanning the surrounding space to obtain spatial information. Thereafter, the update of the spatial mapping information is performed through continuous spatial scanning during execution of the application program. In the process of updating the spatial information, the surface information of an incorrect object is acquired due to an error, and a problem arises about the spatial arrangement of the virtual object.

The objective of this paper is to evaluate the optimal environment configuration for executing the augmented reality application program by realizing accurate spatial information when mapping the actual space and object layout environment using HoloLens. To do this, the surface information of the space is scanned and mapped in triangular mesh form. A virtual environment of the same type as the real space is made into a 3D model. Then, the spatial information obtained from the real environment was compared with the virtual 3D model to analyze the matching.

The composition of this paper is as follows. Chapter 2 describes the structure of HoloLens and the differences between the virtual lens and the virtual lens. Chapter 3 introduces the proposed evaluation system and experimental method, and analyzes the correlation of data from comparative analysis. Chapter 4 concludes the paper.

2. Augmented REALITY (AR)

HoloLens is a device capable of displaying AR technology and uses Augmented Reality (AR) which uses technology to display virtual information on the real world. Unlike Virtual Reality (VR), which previously used to block the real world and create a perfect digital environment, it is called Mixed Reality (MR) because it is a mixed technology based on reality. One of the major features of HoloLens is that HoloLens itself is an independent computer. Since HoloLens is itself a computer, there is no need for a separate device which is a significant advantage when worn and active [4].

2.1 Configuration of HoloLens reality

The core device of the computer is the processing unit. In the case of HoloLens, there are a total of three chips, each of which operates by allocating the entire work. There are a central processing unit (CPU), a graphics processing unit (GPU), and a hologram processing unit (HPU) [5]. Fig. 1 shows the processing unit of HoloLens. The display device is equipped with two 2.3 megapixel displays installed separately in a glass-shaped external display as shown in Fig. 1.
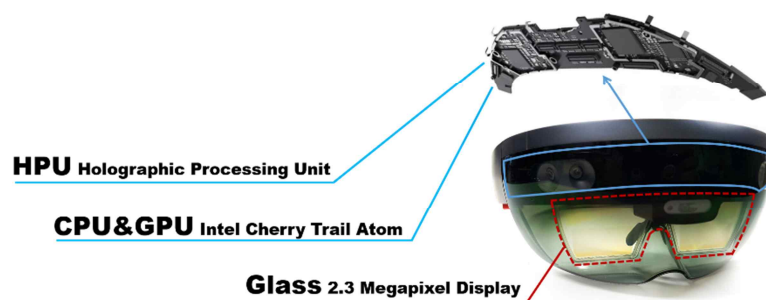

Figure 1. HoloLens system architecture

2.2 Spatial Mapping

Since HoloLens uses mixed reality using virtual objects in the real world, information about the spatial

coordinates of the real world is most important. If HoloLens cannot accurately map the coordinates of the space, it will cause a lot of errors in placing the virtual object in the real space. In addition, the mapping can recognize not only the coordinates of the space, but also the depth information of the object to be viewed through gaze, thereby providing an opportunity to recognize objects such as walls, floors and chairs in HoloLens [1]. Space mapping through HoloLens allows capturing of real-time meshes and verticals through the camera by simply moving the head. Fic. 2 is a scenario that HoloLens actually acquires the mesh information of the object. The model on the left is real object, and HoloLens recognizes vertices and triangles. Green frame on the right is a component that defines the collision area in the physical system. Therefore, if we want perfect interaction between the real world and the virtual world through HoloLens, the exact configuration of this collision area will have a significant impact on placement of hologram.
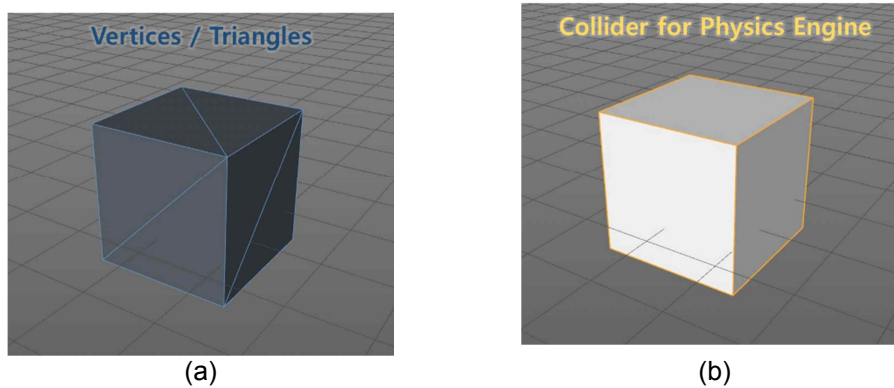

(a)   (b)

Figure 2. Physical objects using vertices and triangles

In order to correctly recognize the information of vertices or triangles of such an accurate object, the hole lens correctly recognize the spatial volume and start to scan the object depending on the accurate depth information. For this purpose, HoloLens has a fixed scan density (resolution) as the number of triangles per cubic meter. Fig. 3 is a result of adding a white frame to scan the space by dividing the resolution by 100, 750, and 2000 and to check the scan density.
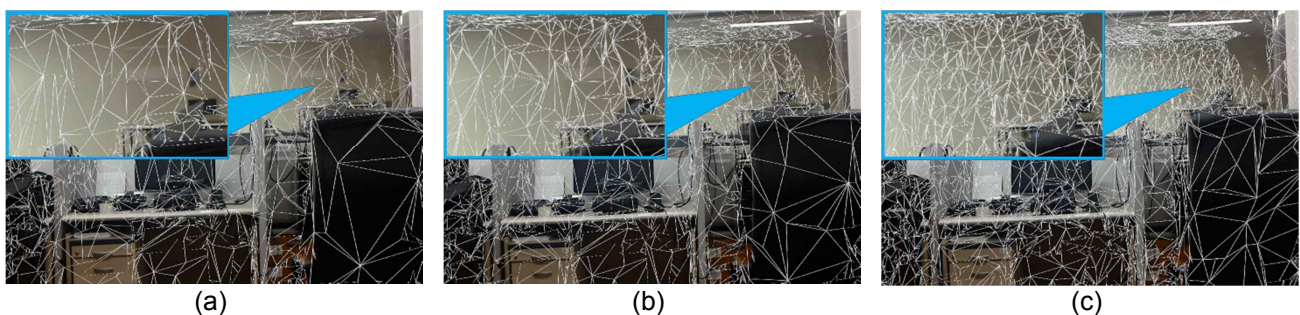

(a)   (b)   (c)

Figure 3. **Scanning with HoloLens while changing the number of triangular meshes per cubic meter. Fig. (a) – (c) represents triangle scan density of 100, 750, 2000**

A resolution lower than 100 will speed up scanning and rendering, but it's easy to lose surface detail. In addition, scanning at higher resolutions can bring in accurate spatial information that can interact with the surrounding world, but it also allows you to invest in longer rendering times and there is a drawback to use a lot of memory. Therefore, HoloLens have a density value standard value of 500, and changing this value cannot be changed in HoloLens own setting without a separate application.

3. Experimental Method

The methodology involves two steps. First, we create a reference model that can compare the actual space to be compared with virtual 3D data extracted using HoloLens and call it ground truth model. We placed objects in real space as shown in Fig. 4 (a) and created a virtual space as shown in Fig. 4 (b) by placing additional objects in the space, we can model space and objects in real size using Cinema4D which is 3D graphics program. Objects with the same specifications in Table 1 are arranged as shown in Fig. 5, and the process and characteristics of modeling are described in Section 3.1.
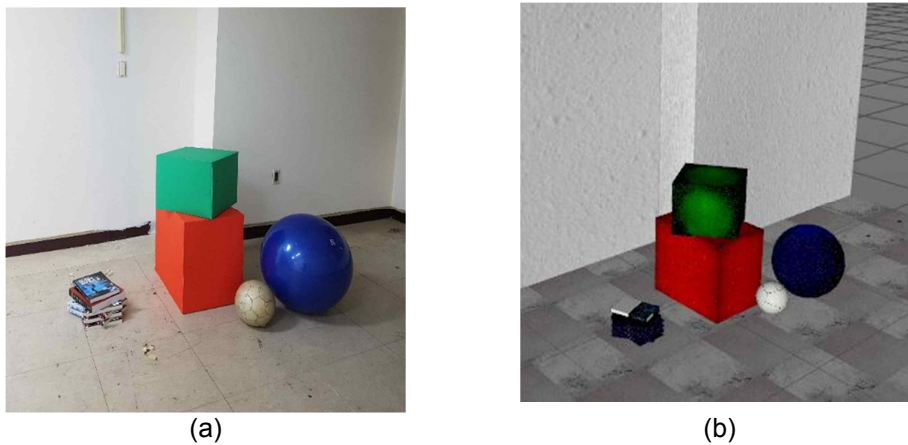

(a)   (b)

Figure 4. (a) Actual space set for experiment (b) 3D Model for actual space based on actual size using Cinema4D

Table 1. Object specification

Object	Quantity	Volume Size (cm^3)
Big Box	1	90.95×10^3
Small Box	1	41.29×10^3
Big Ball	1	52.62×10^3
Small Ball	1	59.51×10^2
Book1	4	78.13×10
Book2	1	16.15×10^2
Book3	1	72.01×10

In the second step, the number of triangular meshes per cubic meter used for scanning through HoloLens is changed based on Table 2. We modify the 3D scan models to a size that can be compared with the model of Fic. 4. We simplify the model and explain this procedure in Section 3.2.

Table 2. Number of triangular meshes for each compared model

Object	Number of meshes	Cubic meter
Model1	100	1 m^3
Model2	750	1 m^3
Model3	2000	1 m^3

3.1 Methods

The spatial mapping of HoloLens is performed by using a built-in TOF camera and two environment cameras on the left and right sides using triangles mesh [1]. The three models each have a triangles mesh number per cubic meter in Table 2. The space consisting of walls and floors excluding the ball, box, and book is shown in Fig. 5, and the 3D model obtained by HoloLens scanning is shown in Section 3.2.
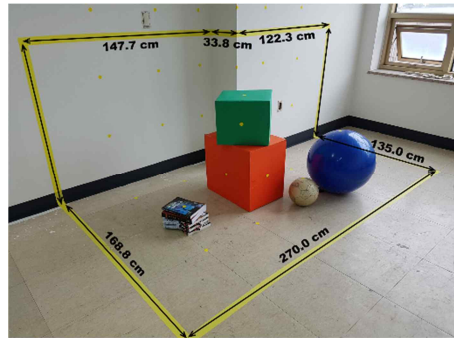

Figure 5. Actual space with objects for mapping

Various errors may occur depending on the scanning time and viewing point of HoloLens. Therefore, we put a total of 20 yellow markers on the wall and floor. And we put one on the book, and five on the small box except for the bottom. In the case of the big box, four markers were attached to four sides except the bottom and the top. And we put one marker on the small ball and two big balls on the top and side. A total of 33 markers have been marked on walls, floors, and various objects on the floor. The scan of HoloLens is shown in Fig. 7 and the experimental space, scan distance, number of test points, and time are shown in Table 3. The test moves between the markers and the marker from right to left, from top to bottom with a delay of 5 seconds to observe constant head tracking. And we had the same experiment at five different angles. The number of visible markers and the total time of experiment at different angles are shown in Table 3 and angle changes are shown in Fig. 6. This experiment was performed three times in total, and each experiment was carried out while changing the values as shown in Table 2.

Table 3. Measurement information from (a) to (e) mentioned in Fig. 5

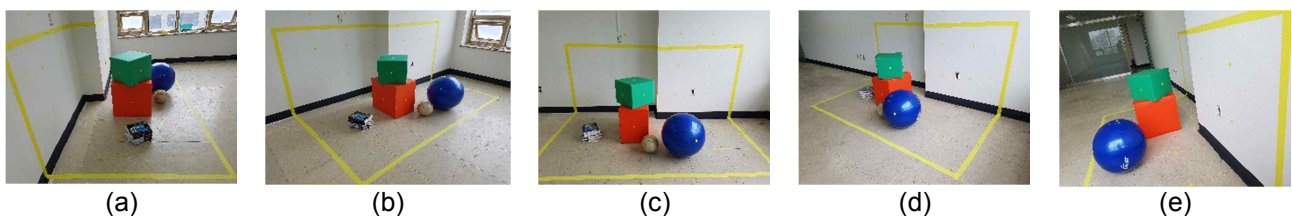| Angle	Number of markers	Total experiment time
(a)	25	125 sec
(b)	24	120 sec
(c)	31	155 sec
(d)	22	110 sec
(e)	21	105 sec

| (a) | (b) | (c) | (d) | (e) |

Figure 6. 5 angles for HoloLens scanning

The evaluation scenario uses CloudCompare [6] which is an open source tool. We used a 3D modeling program Cinema4D to implement a 3D model with the most ideal surface information to be used in the experiment and regard it as a ground truth model. We have scanned from multiple angles to obtain minimum error information, and the three models from this scan are considered as compared models. Then, the average distance and standard deviation of these two data are calculated and compared with ground truth model. Comparisons are based on the color scale using C2M signed distances which represent distance differences in 256 levels of color levels by comparing the distance differences between the ground truth model and the compared model. This can be expressed by using Blue> Green> Yellow> Red for the distance values between the two models. Histogram and Gaussian distribution values are provided in CloudCompare with different color scale values as basic data.

3.2 Result and Discussion

At the same time as HoloLens is turned on, HoloLens starts to scan the space. At the end of this process, with the spatial information obtained, HoloLens prepares to arrange various hologram forms in the real world. Actual number of triangular mesh is set per 500 cubic meters, which is basically set on HoloLens. For this reason, we used SpaceCatcher [7] application to change the size to 100, 750, 2000 to scan the space and get the 3D model as a result. Fig. 7 is the result of applying a black frame to the 3D model of the result value. As the number of triangular meshes per cubic meter increases from 100 to 2000, it can be seen that the vertices and triangular mesh over object surface are increased. It is proved that many corresponding points on the surface can be secured and a large amount of information can be obtained to clearly identify the outline of the object. Fig. 8 shows the result of Heatmaps obtained by comparing with the existing ground truth model. Histogram and gauss distribution using the color-level distribution as the value of C2M signed distance are shown in Fig. 9 as comparative analysis data.
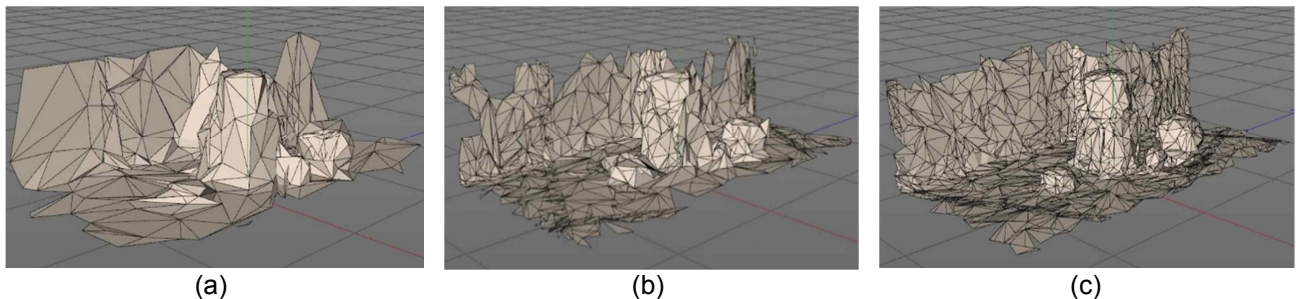

(a)   (b)   (c)

Figure 7. 3D model obtained by changing triangles mesh with HoloLens. Fig. (a) - (c) represents Model1 – Model 3 in Table 2

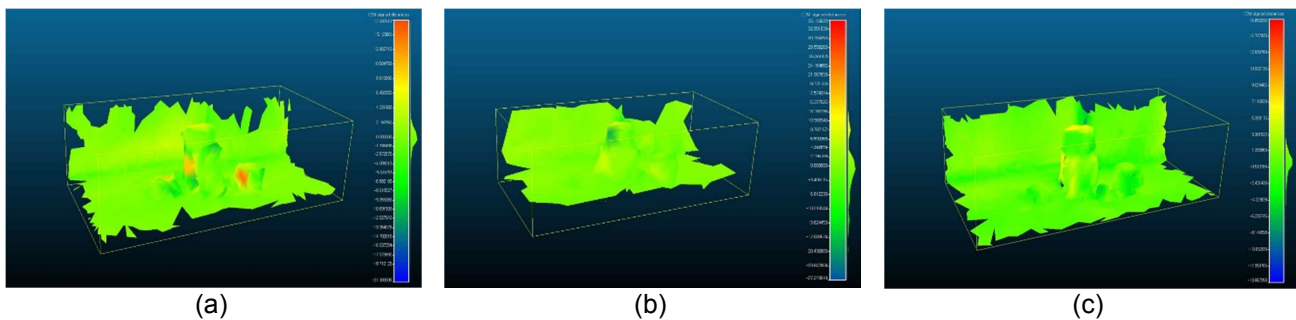(a)   (b)   (c)

Figure 8. Heat maps of distance distributions using C2M signed distance level for HoloLens compared data and ground truth model. Fig (a) – (c) represents Model 1 – Model 3 in Table 2

Before comparing the distance distributions of the two models when using CloudCompare, use the iterative closet point algorithm (IPC) to perfectly align the surfaces of the two models. The target error is set

to 1×10^{-8} for more accurate alignment [8]. Ground truth model and the compared model optimize the meshes of the two models with an octree level structure that is automatically selected as all vertices are computed in software [9]. The CloudCompare shows how each model computes its optimal alignment and then the average of the distances of the contrasted vertices of these two models and how they are distributed from the average. Mean value and standard deviation of Model1 - Model3 are compared with those of Table 4.

Table 4. Mean distance distributions and standard deviation

Object	Mean value	Std. dev
Model1	-1.4450	8.4393
Model2	0.5005	4.2025
Model3	0.0993	3.0190

In Fig. 9, it can be seen that the distance distribution of the corresponding points becomes closer to the center from Model 1 to Model 3. The standard deviation value of the Gaussian distribution graph shows the degree to which the values of all the classes of the three experiments are converged to the average value. Table 4 shows the mean value and standard deviation of each comparison value. After the two models are computed, mean distance value cannot be used as a measure to compare the accuracy of the ground truth model with the accuracy of the compared model. However, the standard deviation value decreases gradually from Mode 11 to Model 3. This result shows that as the number of triangular meshes scanned increases, the surface information of the ideal object can be obtained without any deviation in depth when recognizing the surface.
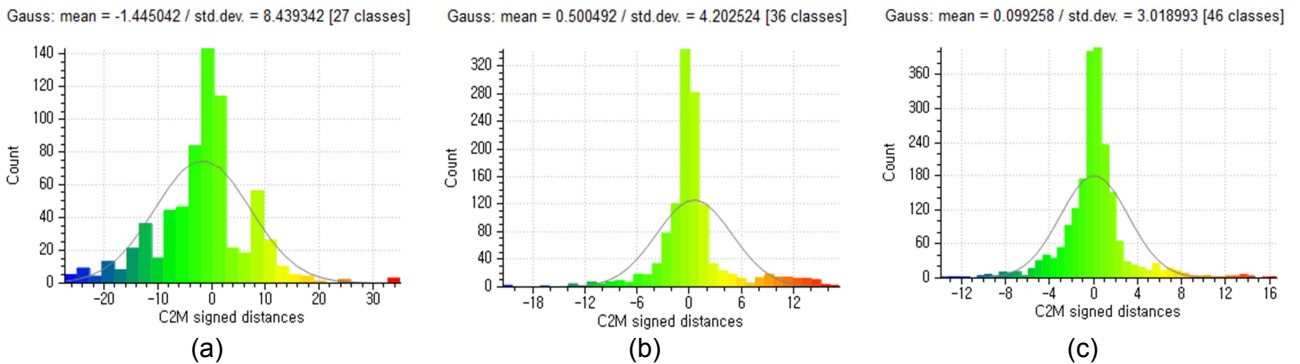

Figure 9. Gauss graphs of the distance distributions for HoloLens compared data and ground truth data. Fig (a) – (c) represents Model 1 – Model 3 in Table 2

4. Conclusion

In this paper, compared model that 3D model obtained by changing the number of triangular meshes per cubic meter used in spatial mapping in HoloLens is compared with ground truth models with ideal surfaces using CloudCompare which is an open source program for comparison and analysis. In HoloLens basic setting, HoloLens have triangular mesh per 500 cubic meters for spatial mapping. However, this value does not have much effect on the change of the number of triangular mesh in a flat and wide plane such as a wall or a floor, but it causes various errors in the case of a small or thin object. Therefore, in this study, we conducted experiments on the accuracy of scanning small and complex surfaces in real space with the scan data obtained by changing the total number of 3 steps in total of 100, 750 2000 values. As a result, As a result, it has been proved that when HoloLens starts spatial mapping, the density of the scan increases with high value of the triangular mesh per cubic meter in HoloLens, and thus the surface information of the object

can be obtained more accurately.

Acknowledgment

References

[1] Allen G. Taylor, *Develop Microsoft HoloLens Apps Now*, Apress, pp. 247, 2016.

[2] M. A. Hoffman, "The future of three-dimensional thinking," *Science*, Vol. 353, No. 6302, pp. 876, Aug. 2016.

[3] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-Flight Cameras in Computer Graphics," *Comput. Graph. Forum*, Vol. 29, No. 1, pp. 141–159, 2010.

[4] R. Furlan, "The future of augmented reality: HoloLens Microsoft's AR headset shines despite rough edges," *IEEE Spectrum*, Vol. 53, No. 6, pp. 21–21, Jun. 2016.

[5] Jason M. Odom, *HoloLens Beginner's Guide*, Packt, pp. 355, 2017.

[6] CloudCompare, (version 2.7.0), *http://www.cloudcompare.org/*.

[7] SpaceCatcher, *http://www.spacecatcher.madeinholo.com/*.

[8] J. Hafeez, S.-C. Kwon, S.-H. Lee, and A. Hamacher, "3D surface reconstruction of smooth and textureless objects," *Emerging Trends & Innovation in ICT (ICEI), 2017 International Conference on*, pp. 145–149, 2017.

[9] T. Bakirman et al., "Comparison of low cost 3D structured light scanners for face modeling," *Appl. Opt.*, Vol. 56, No. 4, pp. 985–992, Feb. 2017.