

Physically-based Modeling of Virtual Content for a Natural way of Interaction in AR

J. Adrián Leal-Meléndez and Leopoldo Altamirano-Robles

Instituto Nacional de Astrofísica, Óptica y Electrónica,
Computer Science Department
Luis Enrique Erro No. 1, Tonantzintla, Puebla, México
{jal,m,robles}@ccc.inaoep.mx

Abstract. Nowadays the Augmented Reality (AR) has been spread into several practical areas of applications. Nonetheless, in many cases the interaction between the real and virtual content is in somehow limited. In this regard, most of the work has devoted to the study of AR using physical controls and gesture recognition. In recent years, several works have moved their attention to research more natural ways of interaction. One of the main problems of interest is the physically-based modeling of real objects. This is, export the physical properties of real objects to the virtual objects. Providing a user interface with such physical properties makes possible to improve the user experience. In this proposal, we are interested in modeling three physical properties (weight, center of gravity and material) into the virtual objects and also modeling the environment. Preliminary results, show that using a 3D sensor to generate a point-cloud-based mesh, it is possible to model a suitable environment that improve the spatial-awareness by allowing occlusion handling. Such modeling seems to be an ideal scenario to perform the physically interaction.

Key words: Physically-based modeling, natural interaction, augmented reality, 3D sensors.

1 Introduction

Augmented Reality (AR) has the potential to bring new user interfaces, in which space is not restricted to a screen and controls, even they could be unnecessary. The classical AR aims to superimpose synthetic information on a view of the real world where both elements, real and virtual, look as a part of the same 3D scene. In this way, a user can receive useful information in real time and in a most adequate place (real environment) in order to be guided in a determined task. In the last years, with the progress of portable technologies such as smartphones, cameras, and sensors for tracking, many AR applications have been developed in several fields. However, the interaction methods provided for such applications are restricted to the use of common input devices: mouse and keyboard. This situation produce a poor user-computer interaction environment.

In this regard, an ideal application of AR should have two main features: 1) allowing the user to interact with the virtual content in a more natural and intuitive way, and 2) the system must provide a kind of feedback mechanism to ensure the spatial-awareness (mixed space) of the user while navigation is performed in the environment. The aforementioned two points are key procedures to improve the immersion and user experience. In this context, different approaches have been used to give more natural ways of interaction with virtual content. There are two main approaches focused on improving the user experience: i) the use of physical objects to interact with the virtual content (tangible user interfaces, TUI), and ii) the use of gesture recognition. The first approach is based on using physical objects as a tool for interacting with virtual objects, this approach was one of the first that removed the dependence of the use of controls along the manipulation phase [11, 12, 15]. The main idea behind this approach, is to manipulate the real object and then reproduce this behavior in the virtual object, thereby when the physical object is manipulated the associated virtual one does too. In this way the user has the feeling of interacting with a real object. Because of the rise of the use of mobile devices, this approach adopted a new form; in this case, the manipulation of virtual objects is based on the use of tactile screens and sensors for tracking and tracing (GPS, accelerometers and gyroscopes). Researches in this approach provide a way of interaction with virtual content through the device, using it as an interface [21, 10, 7, 8, 6]. In spite of providing more intuitive ways for interacting, these works, still remain far from provide an enhanced user experience. The second approach is the one based in gesture recognition, in [13, 3, 2, 14] the authors propose more natural methods for interacting by matching gestures with actions applied to the virtual model. In these works the gestures do not imply contact with the virtual object. On the other hand, in [18, 9, 1] the authors consider the contact with the virtual object and they propose a modeling of that. Finally, the above two main approaches (the use of physical objects and gesture recognition) can be combined into an hybrid approach, in order to consider both of them. This approach is best known as multi-modal interfaces [4].

In order to provide a more natural and intuitive mechanism for interaction in AR, it is necessary going beyond the gesture recognition. In real life, while interacting with real objects there are a number of physical constraints which determine the behavior of such object, these restrictions also determine the user's behavior necessary to manipulate the object. Thus, according to the literature, a way to improve the user experience in AR is modeling the virtual content in such a way it can react to the real world. This is achieved by modeling physical properties into the virtual object, but also modeling the way in which the environment interact with the virtual content. In this way, the user experience is improved by giving to the user the feeling that virtual objects are part of the real world. Furthermore, this condition might allow a natural interaction directly with the user's hands. Having said that, this research is closer to the second approach. The research lies in designing methods to give to the user a

physically-based way of interaction in AR with the purpose of improving the user experience. The main problem addressed in this research is deepen next.

Problem: gestures are the most expressive form of communication between humans and computers. Moreover, despite of the approach based on gesture recognition has been used extensively to interact with virtual content in AR, most of these works usually do not consider physical properties to model its behavior. The latter result in a poor manipulation experience due to the lack of feedback; this is there are no feeling of working with real objects. On the other hand, it has been shown that considering physical properties while modeling the virtual content improved the immersion of an AR system, mainly along the interaction phase [16, 5, 18, 17].

Main objective: Proposing an strategy for natural interaction with virtual objects in an augmented reality environment that considers physical properties (weight, center of gravity, material) of real objects to model virtual objects.

Main contribution: the main contribution in this research is the design of algorithms for modeling real and virtual objects considering physical properties to provide a mechanism for a natural user-experience in AR.

Following the above three points, this research lies on a physically-based interaction in augmented reality; mainly in modeling physical properties to emulate the behavior of a real object when this is pushed by another one and, then apply this model to a virtual object. We are focused in modeling properties of rigid objects directly involved in the interaction task. As a first approach we have considered to model the rigid virtual objects through the use of the finite element method (FEM), The latter is because its capabilities could reproduce real-world behavior [22, 20, 23]. For the rest of the scene we have considered to build a simple mesh from a point-cloud¹ obtained with a 3D sensor. We are interested in modeling only certain specific-target objects, the rest of the scene can be treated as a simple mesh. The rest of this paper is organized as follows: in Section 2 describe the research methodology. In Section 3 we show the work done so far. Finally, in Section 4 an analysis of the obtained results is presented and future avenues of inquiry.

2 Research Methodology

The proposed methodology is divided into five main parts, where the main contributions are listed in 1(b), 2 y 3:

1. **Designing an spatial representation of the work area to allow collision detection and occlusion handling:** collision detection and occlusion handling among objects are important features to improve the user experience, while the interaction is performed through the virtual content. To make

¹ Point-cloud: in a three-dimensional coordinate system, is a set of points in which the points are usually defined by $[X, Y, Z]$ coordinates, and often are intended to represent the external surface of an object. A point-cloud may contain additional components as color and direction.

virtual objects react to a collision as a result of an applied force on them, it is needed to model both the virtual object and also the environment. Moreover, the modeling allows detect occlusion relationship, i.e., after reacting to a force, a virtual object could be behind of a real one. The modeling of the environment allow this condition.

- (a) **Creating a general model (mesh) in which the virtual object could be attached:** the main idea is to use a point-cloud-based representation using a 3D sensor to create a mesh geometry, which will cover the entire work area. The main problems working with a cloud point obtained from a 3D sensor are: i) there are areas that are not covered by the sensor that produces holes on the mesh, in which the virtual content can not be attached, and ii) there is an overlapping between the mesh and the rgb image.
- (b) **Modeling an specific horizontal flat surface, in which the virtual object should slide when a force is applied to it:** once the general model of the work area was created, the next step is to segment an specific flat surface, on which the virtual object will be slipped. We are focused in modeling this surface with two different materials: wood and glass.
- (c) **Modeling the obstacles in which the virtual object should collide:** in this point we are interested in select vertical surfaces, which will represent obstacles on the general geometry and model them like a rigid bodies.

This stage include an study of 3D reconstruction and physical modeling approaches as well as point-cloud correction techniques.

2. **Physically-based modeling of virtual objects:** in this stage we are focused in reproduce the behavior of an object, mainly when a perpendicular force is applied on one side. The physical properties considered to model the virtual object are the following: i) weight, ii) center of gravity (contact area which is divided into six zones, see Fig 1), and iii) material (resistance and friction). Initially, we propose to work with three objects with well-defined shapes, which are: cube, pyramid, and sphere. We propose to combine two approaches: i) to use a Finite Element Method to describe the physical constraints and the geometry itself, and ii) to use semantic rules to describe the behavior. Both approaches are studied along this stage and an analysis of the advantages and disadvantages of each approach will be made.
3. **Modeling the tool to interact with the virtual content (user's hand):** once the work area and the virtual objects have been modeled, the last element to complete the interaction phase, is the tool used to interact with the virtual content. We are focusing in only one interaction way: push an object; so we propose to model the hand as a rigid body, where only the tip of the index finger will have a slight deformation when the contact occurs. Figure 2 shows the shape of the model. In order to match the model with the real hand, we propose a color-based approach to segment the hand in a real video sequence. Then a template-matching approach is used to select

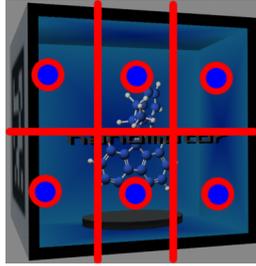


Fig. 1. Contact area

a more appropriate shape according to the size. During the interaction task we assume that the shape is constant.



Fig. 2. Push model hand

4. **Designing a feedback mechanism to guide the user during the interaction task:** when working with virtual objects in an interaction task is necessary to provide some sort of feedback to advise the user that is handling with a virtual object or maybe to guide him/her to the right position to be touched. We are interested in forms of haptic feedback; to this end, we will study different approaches on visual, auditory and tactile feedback. The integration of such mechanism to the interaction process will also be made in this stage.
5. **Evaluating the proposed strategy for physical-interaction in AR:** two applications will be designed in order to evaluate the proposed strategy. The first one will be designed so as to measure the difference of the final positions between the real and virtual objects after a force is applied on each one, as well as the difference of the paths followed to reach the final positions. This application will take place in the early stage of the research, in which the user will be asked to apply a force on a virtual object to show its behavior. In the later stage, we will focus on the second application, which aims to evaluate the user experience and usability. In order to achieve that a user study will be performed through this application.

We expect to achieve the goals of this research by following this methodology. In the next section we describe the preliminary results obtained so far.

3 Preliminary Results

This section describes the preliminary results of our research. It also describes the work performed following the proposed research methodology.

3.1 Modeling a spatial representation of the work area

As a first approach, we propose to use a point-cloud from a 3D sensor to create a representation of the environment. Figure 3 shows a common point-cloud obtained from the sensor. Figure 3(a) shows the original point-cloud, which has over 307,000 points, composed by $[x, y, z]$ and color data. We can see the more black pixels being holes in which the virtual object cannot be attached. Figure 3(b) shows a correction of the point-cloud to fill holes. To do this, we used the inpainting-telea algorithm [19].

The original cloud have many points and it is impractical to be used in a real-time application, Figure 3(c) depicts an example of reducing the point-cloud density. By using thresholding, we can delete points out of the work area; however, the cloud remaining still has unnecessary points (there is not a significant difference in the z component of many grouped points). In this point, we are interested in methods that allows reducing the redundant information without sacrificing significant geometries. After the cloud is reduced, the normal² of the faces (triangular faces) must be calculated and added to it; so, now the cloud has five components. A first candidate to be used in this point is the moving least squares method (MLS), which is used to reconstruct a surface from a set of points and to calculate the normal of the faces. This produces a smoothed cloud.

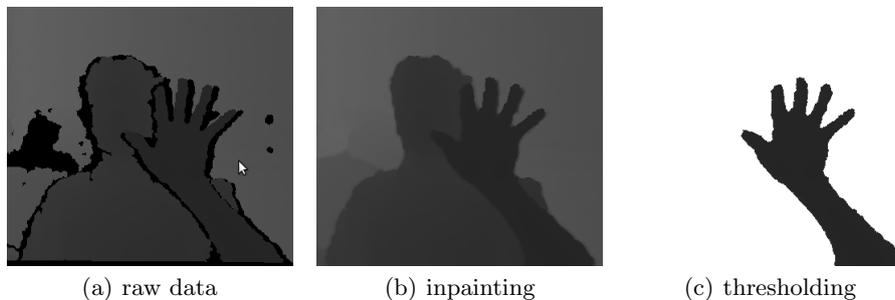


Fig. 3. Point-cloud from kinect

Once we have the five components, the surface can be reconstructed. Figure (4) shows only the hand area that is reconstructed using the greedy triangulation

² In geometry, the normal is an object such as a line or vector that is perpendicular to a given object. In this case it can be seen as the direction of the faces.

algorithm, provided by PCL library³. The reconstructed surface enables the occlusion handling and lays the foundation for collision detection. Only the hand area was reconstructed because we are only using the threshold filter for reducing the cloud. Thus, this is still computationally expensive.

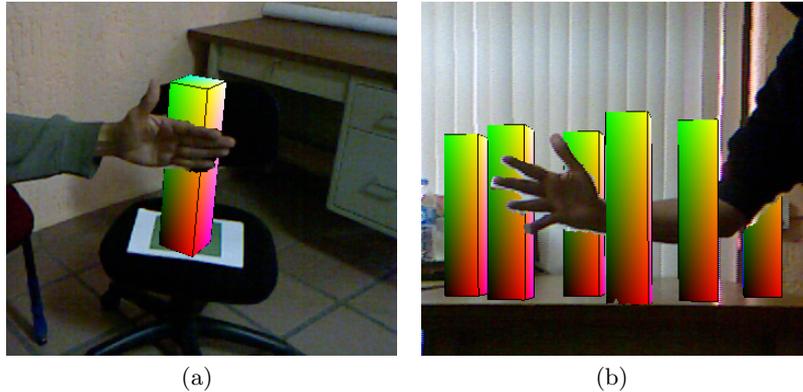


Fig. 4. Enabling occlusion

3.2 Evaluating the proposed strategy

The evaluation will be conducted within the context of an application. In this point of the research, we have already defined a user study, which will be driven through this application. Furthermore, this study will help us to define the features that should be present in the application itself.

Objective: with this study we want to evaluate the proposed strategy for interacting in AR, in terms of realism and intuitiveness. The two main factors which we are interested in measuring are:

1. **User experience:** involves the person's behaviors, attitudes, and emotions about using our application.
 - Intuitiveness: involves the feeling of naturalness and human-likeness in interacting with the virtual content.
 - Captivation: involves the feeling of being immersed and captivated in the interaction with the AR.
 - Responsive: involves the feeling about the response time of the system to a particular action.
 - Tangibility: involves the feelings spatial-awareness and coherence of the content according with the environment.
2. **Usability:** involves the distribution of content and clarity with which the interaction is designed.

³ <http://www.pointclouds.org>

- Learnability: involves the easy to accomplish a task the first time the user has contact with the application.
- Efficiency: involves the speed with which a task is performed.
- Errors: involves the number of errors made by the user in a task.

External factors: there are some external factors that can influence in the evaluation, we are interested in control the next two.

- **User experience:** the user experience with AR systems is a key factor in performing a task.
- **Incentives:** the willingness of the user to perform a task influenced in the result, a change in user motivation usually results in a change in what he/she does. In this point, we are studying the best way to encourage the participants without causing bias.

Users: the users play an important role in this evaluation, in order to cover as many aspects as possible and considering that the application is a game, the participants will be divided into three groups: 1) inexperienced users in augmented reality systems and video games, 2) inexperienced users in augmented reality but experienced in video games and 3) experienced users in augmented reality systems and video game.

Setting: each participant will be given a tutorial before performing the task. Then he/she will be asked by the system to perform certain tasks. While the user is performing the tasks, an observer will note for feedback purpose. When all tasks are completed, the participant will be asked to answer a poll covering all aforementioned aspects.

A more accurate evaluation also will be addressed in order to give foundations that allows the comparison with other works. This include measuring the differences between the final positions (translation and rotation) of the real and virtual objects after an interaction task, as well as measuring the differences between the paths followed to get the final position.

4 Conclusions

In this paper, we present our research proposal to address the problem of natural interaction in augmented reality. The main motivation to do this work is to advance in the state-of-the-art with respect to interaction approaches in AR using a physical-based modeling to reproduce the behavior of real objects into virtual objects. To achieve a realistic interaction we consider modeling both the virtual and the real objects. Although our research is in the early stage, the preliminary results showed in Figure 4 give evidence that it is possible to improve the spatial-awareness of the user. The user can determine at all time the right position of the objects in a mixed space, that suggest evidence that the modeling of both real and virtual elements is a right direction to achieve and realistic interaction in an AR environment. The future paths of our research are focused on exploring video-games techniques of physical modeling and propose methods to bring them to a real interaction. We are also interested in propose novel feedback mechanisms to guide the user along the interaction.

References

1. Akman, O., Poelman, R., Caarls, W., Jonker, P.: Multi-cue hand detection and tracking for a head-mounted augmented reality system. *Machine Vision and Applications* 24(5), 931–946 (2013), <http://dx.doi.org/10.1007/s00138-013-0500-6>
2. Bai, H., Gao, L., El-Sana, J., Billinghamurst, M.: Markerless 3D gesture-based interaction for handheld Augmented Reality interfaces. In: *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*. pp. 1–6 (Oct 2013)
3. Bellarbi, A., Benbelkacem, S., Zenati-Henda, N., Belhocine, M.: Hand gesture interaction using color-based method for tabletop interfaces. In: *Intelligent Signal Processing (WISP), 2011 IEEE 7th International Symposium on*. pp. 1–6 (Sept 2011)
4. Billinghamurst, M., Kato, H., Myojin, S.: Advanced Interaction Techniques for Augmented Reality Applications. In: Shumaker, R. (ed.) *Virtual and Mixed Reality, Lecture Notes in Computer Science*, vol. 5622, pp. 13–22. Springer Berlin Heidelberg (2009), http://dx.doi.org/10.1007/978-3-642-02771-0_2
5. Corbett-Davies, S., Green, R., Clark, A.: Physically Interactive Tabletop Augmented Reality Using the Kinect. In: *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*. pp. 210–215. IVCNZ '12, ACM, New York, NY, USA (2012), <http://doi.acm.org/10.1145/2425836.2425880>
6. Ha, T., Woo, W.: ARWand: Phone-Based 3D Object Manipulation in Augmented Reality Environment. In: *Proceedings of the 2011 International Symposium on Ubiquitous Virtual Reality*. pp. 44–47. ISUVR '11, IEEE Computer Society, Washington, DC, USA (2011)
7. Ha, T., Woo, W.: ARWand for an augmented world builder. In: Lécuyer, A., Steinicke, F., Billinghamurst, M. (eds.) *3DUI*. pp. 207–208. IEEE (2013), <http://dblp.uni-trier.de/db/conf/3dui/3dui2013.html#HaW13a>; <http://dx.doi.org/10.1109/3DUI.2013.6550251>; <http://www.bibsonomy.org/bibtex/2874b307c0dda8c24479f4bab168ba97f/dblp>
8. Ha, T., Woo, W.: Poster: A pilot study on stepwise 6-DoF manipulation of virtual 3D objects using smartphone in wearable augmented reality environment. In: *3D User Interfaces (3DUI), 2013 IEEE Symposium on*. pp. 137–138 (March 2013)
9. Hürst, W., Wezel, C.: Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications* 62(1), 233–258 (2013), <http://dx.doi.org/10.1007/s11042-011-0983-y>
10. Kasahara, S., Niyama, R., Heun, V., Ishii, H.: exTouch: Spatially-aware Embodied Manipulation of Actuated Objects Mediated by Augmented Reality. In: *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. pp. 223–228. TEI '13, ACM, New York, NY, USA (2013), <http://doi.acm.org/10.1145/2460625.2460661>
11. Kato, H., Billinghamurst, M., Poupayev, I., Imamoto, K., Tachibana, K.: Virtual Object Manipulation on a Table-Top AR Environment, ISAR. *Presence* pp. 111–119 (2000)
12. Kato, H., Tachibana, K., Tanabe, M., Nakajima, T., Fukuda, Y.: MagicCup: a tangible interface for virtual objects manipulation in table-top augmented reality. In: *Augmented Reality Toolkit Workshop, 2003. IEEE International*. pp. 75–76 (Oct 2003)
13. Lee, B., Chun, J.: Interactive Manipulation of Augmented Objects in Markerless AR Using Vision-Based Hand Interaction. In: *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*. pp. 398–403 (April 2010)

14. Lee, S., Lim, Y., Chun, J.: 3D interaction in Augmented Reality with stereo-vision technique. In: *Advanced Communication Technology (ICACT)*, 2013 15th International Conference on. pp. 401–405 (Jan 2013)
15. Maier, P., Klinker, G.: Evaluation of an Augmented-Reality-based 3D User Interface to Enhance the 3D-Understanding of Molecular Chemistry. In: Foley, O., Restivo, M.T., Uhomobhi, J.O., Helfert, M. (eds.) *CSEU*. pp. 294–302. SciTePress (2013), <http://dblp.uni-trier.de/db/conf/csedu/csedu2013.html#0002K13>; <http://dx.doi.org/10.5220/0004349502940302>; <http://www.bibsonomy.org/bibtex/2ce5a6bb4c6284a3f8f59db9c77f30655/dblp>
16. Piumsomboon, T., Clark, A., Billingham, M.: Physically-based Interaction for Tabletop Augmented Reality Using a Depth-sensing Camera for Environment Mapping. In: *Proc. Image and Vision Computing New Zealand (IVCNZ-2011)*. pp. 161–166. Auckland (Dec 2011), http://www.ivs.auckland.ac.nz/ivcnz2011_temp/uploads/1345/3-Physically-based_Interaction_for_Tabletop_Augmented_Reality_Using_a_De.pdf
17. Song, P., Yu, H., Winkler, S.: Vision-based 3D Finger Interactions for Mixed Reality Games with Physics Simulation. In: *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. pp. 7:1–7:6. VRCAI '08, ACM, New York, NY, USA (2008), <http://doi.acm.org/10.1145/1477862.1477871>
18. Talvas, A., Marchal, M., Lécuyer, A.: The god-finger method for improving 3D interaction with virtual objects through simulation of contact area. In: *3DUI*. pp. 111–114 (2013), <http://dx.doi.org/10.1109/3DUI.2013.6550206>
19. Telea, A.: An Image Inpainting Technique Based on the Fast Marching Method. *Journal of graphics, gpu, and game tools* 9(1), 23–34 (2004)
20. Umetani, N., Takayama, K., Mitani, J., Igarashi, T.: A Responsive Finite Element Method to Aid Interactive Geometric Modeling. *Computer Graphics and Applications*, IEEE 31(5), 43–53 (Sept 2011)
21. Wagner, D., Schmalstieg, D.: First steps towards handheld augmented reality. In: *Wearable Computers, 2003. Proceedings. Seventh IEEE International Symposium on*. pp. 127–135 (Oct 2003)
22. Wu, J., Dick, C., Westermann, R.: Efficient collision detection for composite finite element simulation of cuts in deformable bodies. *The Visual Computer* 29(6-8), 739–749 (2013)
23. Xu, G.: Finite Element Methods for Geometric Modeling and Processing Using General Fourth Order Geometric Flows. In: Chen, F., Jüttler, B. (eds.) *Advances in Geometric Modeling and Processing, Lecture Notes in Computer Science*, vol. 4975, pp. 164–177. Springer Berlin Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-79246-8_13