Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 141 (2018) 263–270

Procedia Computer Science

www.elsevier.com/locate/procedia

The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2018)

3iVClass: a new classification method for Virtual, Augmented and Mixed Realities

Marc Parveau[a]*, Mehdi Adda[a]

[a]*Département de mathématique,d'informatique et de génie, Université du Québec à Rimouski, 300 allée des Ursulines, G5L 3A1, Québec, Canada*

Abstract

For several years, augmented and virtual reality technologies have attracted increasing interest in all areas. In the midst of this universe, the concept, already well known, of mixed reality has established itself as a distinct paradigm. However, and contrary to augmented and virtual realities, there is not a clear and one definition of what it is exactly and why it is different from the other concepts. In this article, we attempt to provide a new classification method to standardize the definition of virtual, augmented and mixed realities. First, a quick overview of existing taxonomies is made, then we present our classification which is based on three criteria we called *3iVClass* (Immersion, Interaction, Information). Finally, in order to verify its reliability, we used this classification to propose a definition of mixed reality.

Keywords: : MR; Mixed Reality; Augmented Reality; AR; VR; Definition; SLAM; 3iVClass;

1. Introduction

Currently, augmented, virtual and mixed realities are well-known paradigm studied since a long time. Unfortunately, while augmented and virtual realities have convergent definitions, a wide number of definitions are used to define the mixed reality paradigm resulting in confusion and misinterpretation. In order to avoid this problem, we are trying to unify those definitions by proposing a new classification method we call 3iVClass. This method is

* Corresponding author. Tel.: +33-670-917-597.
E-mail address: marc.parveau@uqar.ca

10.1016/j.procs.2018.10.180

based on three criteria which are common to each of the augmented, virtual and mixed realities paradigms. To begin, in section 2, we are studying how the augmented reality paradigm is defined and how this definition can be used to create our classification. After that, in section 3, the classification method is tested and used to define the mixed reality paradigm which doesn't have a standardized definition. Then, in section 4, to verify the reliability of this classification, a comparison of the different paradigms is realized.

2. 3iVClass principles

The idea behind this classification is to find a common and objective way to classify new “realities” paradigms. Indeed, if we take the mixed reality paradigm as an example, we can see that there is no clear and unique definition. To fully understand the origins of the 3iVClass, it is important to study the underlying technologies and concepts related to it, namely augmented, virtual and mixed realities. Then, augmented reality taxonomies are presented to understand how this paradigm is defined. Finally, and based on this definition, a new classification is proposed.

2.1. Augmented reality: the essential block of mixed reality

Augmented reality was introduced during the 60s by the computer vision pioneer Ivan Sutherland [1]. It was with the introduction of the personal computer during the 90s that this concept became a technology and a field of research of its own [2]. Augmented reality (AR) is conventionally defined as a technology that allows the superimposition of virtual information upon the real world, for example, by adding text or an image to what the user sees. Nevertheless, there are different interpretations and taxonomies [3, 4] which can be used to provide a based definition for the 3iVClass. The first taxonomy that we can cite was proposed by Azuma [3]. According to him, augmented reality can be defined following three distinct criteria:

- The combination of one or more virtual objects superimposed on a real environment.
- Real-time interaction between the physical environment and virtual incrustations.
- Virtual incrustations must be registered in three-dimensional space.

Another definition, proposed by Jun Rekimoto [5], defines Augmented Reality as a new type of human-machine interface just like GUIs (i.e. Graphical User Interface). According to the dictionary definition [6], an interface is "*a surface forming a common boundary of two bodies, spaces, or phases*". If we follow this definition, the user should be in direct contact with reality, which is not exactly the case since a head-up display device is required to provide the augmentation, which is basically adding virtual information to reality.

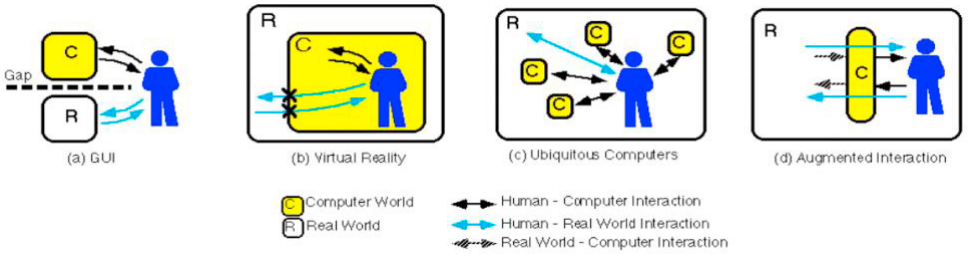

Fig. 1. Diagram of human-machine interactions, Rekimoto [5]

We can also note that, in Rekimoto’s diagram (Fig. 1), the proposed denomination for augmented reality (Fig. 1. (d)) is represented as "Augmented Interaction". In a certain way, it removes the gap between computers and reality that can be found in a conventional graphical user interface (GUI) [Fig. 1. (a)].

Paul Milgram and Hirohisa Kishino proposed a definition that diverges in some way with the previous one. According to their « taxonomy of mixed reality visual displays. » [4] written in 1994, augmented reality is in a continuum (i.e. space that is not interrupted) real-virtual that they call "mixed reality" and which represents the universe between the real world and the virtual world [Fig. 2].

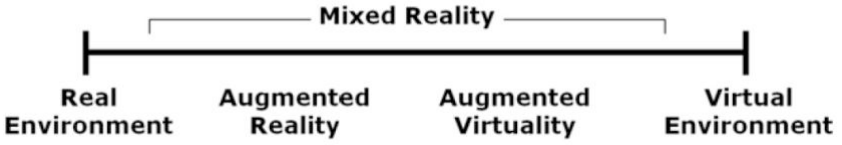

Fig. 2. Diagram of the mixed reality continuum of Milgram. [4]

In this representation of the continuum, the "real environment" represents the environment surrounding the user, that is to say, the outside world. The "virtual environment" represents an environment that would be generated entirely by computer and in which the user would be immersed without any outside interaction (Virtual reality).

Just like the augmented virtuality, which adds real content to the virtual environment, augmented reality is an element that is not covered by the Milgram's real-virtual continuum and that can tend to one side of it, or the other. The closest it is to the "real environment" the less the user experience will be immersive, the augmented content will be more limited. On the contrary, an application with an important quantity of virtual objects or a virtual world would be on the "virtual environment" side. However, there is no exact threshold on the “important quantity” of virtual objects and it is thus a subjective parameter.

2.2. Proposition of a new classification method: 3iVClass

As we saw in the presentation of the different taxonomies of augmented reality [3-5] , the way to define the paradigm is quite similar. Indeed, each definition describes the way the user is immersed into the experience, what type of interaction he has and finally what type of information is handled.

Because of theses similarities, our proposition is to create a new classification method based on those criteria. The following table [Table 1.] describes this method, we called 3iVClass.

Table 1. Classification criteria.

Criterion	Definition
Immersion	This criterion corresponds to the description of the immersive experience brought by the technology during the user experience.
Interaction	This is the description of the type of possible interaction provided by the technology.
Information	Corresponds to the types of information and data managed.

3. An attempt to normalize the definition of mixed reality based on the 3iVClass

Currently, the mixed reality paradigm is one of the most confusing one, because, as we will see it hereafter, there is no common definition of it. To test the reliability of our 3iVClass, we are attempting to give an objective definition for this paradigm.

3.1. Back to the first interpretations of mixed reality

After recalling the different concepts that make up the "mixed reality" continuum of Milgram, it is already possible to note that the term "mixed reality" is widely used. It is therefore interesting to ask if all the uses of this term refer to a single definition. To answer this question, it is necessary to study different usages in order to identify convergence and divergence points.

In the article "Collaborative Mixed Reality" [7], proposed in 1999 by Mark Billinghurst and Hirokazu Kato, the

mixed reality is presented as the inclusion of virtual objects in the real environment. Billinghurst introduces an interesting notion by making a distinction between "Virtuality" and "Augmentation". According to him, the "Virtuailty" represents the integration of virtual objects, which can be observed and examined in the real environment whereas the "Augmentation" principle represents the simple virtual annotation of real objects.

In 2006, in the article "Virtual reality and mixed reality for virtual learning environments" [8] by Pan, Z., & al, a different definition is proposed: "*The incorporation of virtual computer graphics objects into a real three-dimensional scene, or alternatively the inclusion of real-world elements in a virtual environment* ". Through this definition, we find the previous notion of inclusion of virtual objects in the real environment, enhanced by the inclusion of real objects in the virtual environment. Here we see an evolution in the interpretation of the term, without having any precision about the proportion of each of the objects in relation to their environment.

Finally, in 2017 in their article "Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices" [9] C. Kress and J. Cummings presented the different types of Head-Mounted Displays (HMD) by category. A distinction between "merged reality" and "mixed reality" is also introduced. This difference resides in the mediate of representation of the real world. In the case of "merged reality" the world representation is done by "video see-through" contrary to "mixed reality" which uses an "optical see-through" mediate. In addition to this distinction, Kress and Cummings also point out that the main difference between an augmented and mixed reality device lies in its ability to map depth and understand its environment.

These various examples confirm, over the years, that the definition of the mixed reality is not fixed and remains in constant evolution. However, it is possible to observe common points between the different definitions:

- The first point of convergence is on the notion of virtual object integration within the real environment. Moreover, this is in line with the definition of the word "mixed" ("which includes two or more elements of different nature").

- A common distinction is also made on how to display these virtual objects in the environment. It is not a simple "Augmentation" but virtual objects in their own right that can be related, or not, to objects of the real and with whom they can interact.

- Finally, we can observe the importance of the notion of environment. Each article presents mixed reality as improving the user experience. The goal is to enrich his environment while avoiding disrupting the link he has with reality [10].

As we saw earlier, the concept of mixed reality is not recent, one of the first to have mentioned it is Milgram in 1994 by its real-virtual continuum [4]. This idea of mixed reality is also reused by Mann in his conception of what he calls "mediated reality" [11] which is defined as the action of modifying the real world via an intermediary. To fully understand what mediated reality represents, it is necessary to define the notions of "mediate" and "immediate".

According to the "*Merriam-Webster.com"* [12-13] dictionary, mediate is said of a thing, " *Connected indirectly through another person or thing; involving an intermediate agency*". Conversely, immediate is said of a relation or action "*without an intervening medium or agency; direct."*

As it is the case with the Milgram's continuum, we can see on the Mann's mediated reality graph [Fig. 3] that mixed reality is composed of augmented and virtual reality.

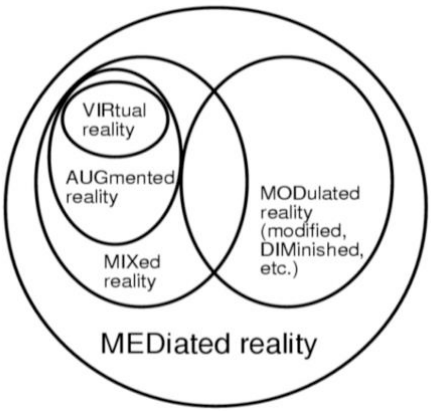

Fig. 3. Steve Mann's mediated reality model [11]

These three different realities, considered as mediated, implies the use of an intermediary in the user's perception of his environment. It is in this environment perception that we can identify the main criteria of differentiation between the virtual, augmented and mixed reality paradigms.

3.2. Mixed, virtual, augmented realities, SLAM as a main differentiator

We have seen previously that there is a difference between the « Augmentation » and the « Virtuality » [7], therefore, the mediator plays a key role in shaping information. Indeed, unlike a conventional augmented reality experience, in the mixed reality experiences the user can consider his environment in his various actions. The diagram below [Fig. 4] [14] is a representation of a user viewing a mixed reality environment through an HMD. One can view a mix of physical and virtual objects which are all interactive.

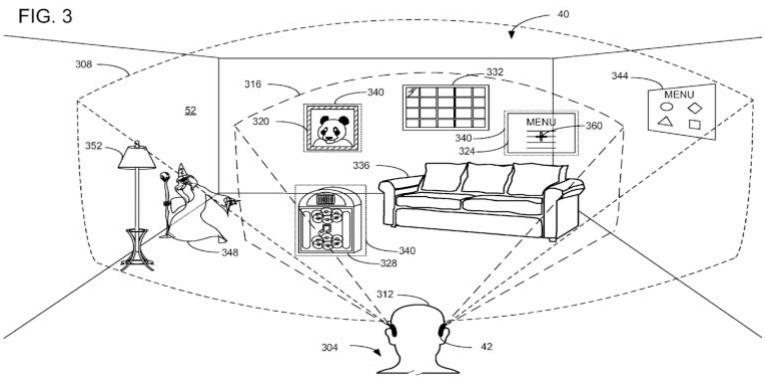

Fig. 4. Illustration of the of a mixed reality environment [14]

For example, the user may place a virtual object on a table or use the walls to install virtual windows [Fig. 4. Menu # 344]. These virtual objects can therefore be saved in three-dimensional space and time according to:

- A position in relation to the world: An object can have a fixed position and orientation, in the real environment, allowing the user to move around.
- A position relative to the user: The object can be fixed relative to the user. A displacement of the user then drives a proportional displacement of the object.
- A position relative to an object. The object will be positioned relative to another object, be it virtual or physical.

These methods of space recording involve the calculation and the "understanding" of the environment. However, it is possible to compare this process of spatial knowledge with the SLAM principle, "*Simultaneous Location and Mapping*". Originally, this is an extremely well-known problem since the 1980s, in the field of robotics, where one wonders whether a robot placed in an unknown environment is able to map it incrementally while knowing its position inside [15]. Hugh Durrant-Whyte pioneered in this area, defines the problem as « *SLAM is a process by which a mobile robot can build a map of an environment and at the same time use this map to deduce its location.* » [16]

It is possible to use this concept to define what the mixed reality is. In this context, it is no longer the robot equipped with sensors that will move, but the user equipped with a mediate of mixed reality. Therefore, this mediator must necessarily be equipped with sensors allowing it to perform such mapping. One can refer to the use of depth mapping technique by TOF (i.e. Time Of Flight) which is extremely common thank to its ease of implementation [9, 17-19]. Thanks to this understanding of space, the mediator is able to render a virtual representation of its environment to the user. For example, by recognizing an object or knowing its position in the environment, the user will be able to interact with it, whether it is a virtual or physical object, as long as the object is connected. Spatial mapping is a determining concept in the user experience since it puts the human at the center of the experience. It is therefore possible to create adapted interfaces that are suitable to the needs of the user in order to promote the acceptance of the experience while reducing the adaptation time. It is therefore by answering the SLAM problem that we identify the key differentiation point between virtual, augmented and mixed realities.

3.3. Definition of mixed reality according the3iVClass classification method

As we seen it before, we provided a classification method called « 3iVClass » based on three different criteria (immersion, interaction, information). In order to highlight the preceding points as well as the significant differences between the concepts of augmented, virtual and mixed realities, a definition of mixed reality is proposed according the 3iVClass:

- **Immersion:** User environment must be processed and interpreted (calculated) in real time. This process can be named as spatial mapping or spatial understanding.

This is the most important point, it allows the user to be at the center of his experience. It also limits augmented reality classic "mistakes" such as positioning errors or displaying through surfaces (and not on it). This spatial mapping brings an immersive feeling far more important improving in a significant way the user experience.

- **Interactions:** Natural user interactions must be processed in real time and in an immediate way

Once the space is mapped, the user must be able to interact with it. In order to stay on a user-centered experience, he must be able to interact in a way which is as natural as possible (gesture, voice, gaze) without any mediate (no controller).

- **Information:** Virtual object must be registered in space and time and can be placed according to the user position, the environment, or any other objects. Every virtual or physical object must be intractable in real time.

Unlike augmented reality, the goal is no longer simply "Augmenting" the real environment. It is therefore necessary to have the ability to display virtual objects in three dimensions and in real-time, in and with respect to the environment of the user. These 3D objects are often wrongly called "Holograms", because they are supposed to be perceived without the help of a particular device. In this case, the objects are computer generated and displayed by a mediator. The most common way to achieve this is by means of an HMD.

The diagram below [Fig. 5] shows the three essential components of this technology. We can thus see that the association between the man/machine interactions, the computer perception of the environment and the so-called "conventional" reality form the mixed reality.

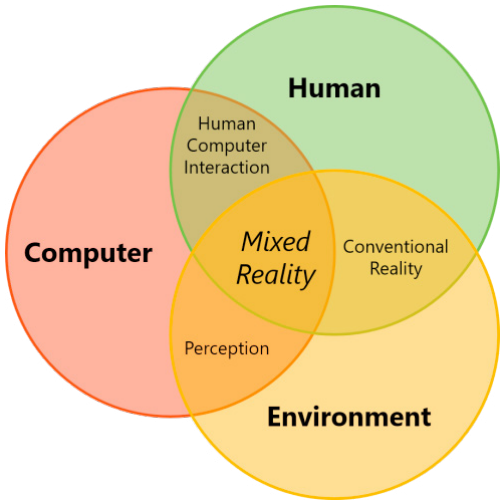

Fig. 5. Mixed reality principal components. [20]

At this point, a definition of mixed reality that can be made based on the "3iVClass" is as follows:

- Mixed reality is a paradigm that combines technologies that, by mapping the user's space, display 3D-embedded virtual content registered in space and time. Virtual objects can be positioned relative to the real environment, the user, or any other virtual or physical object. In addition, the mixed reality experience must be user-centered and offer natural and immediate interactions.

4. Application use cases for 3iVClass

In order to verify the relevance of our classification, we propose a comparison of the three different augmented, virtual and mixed realities paradigms.

- **The immersive character:** While the goal of these technologies is to improve the user experience, the way to achieve it is different. Indeed, for augmented reality, the immersion is limited, the user has the opportunity to view contextual information, such as text or virtual objects, related to its environment, we can for example imagine the display of a virtual object when a QR code is scanned. Conversely, Mixed reality gives an experience far more immersive thanks to the spatial mapping capability.

- **User interactions**: In augmented reality, the user is not able to interact with real-world objects directly. Tags, QR codes or recognition algorithms can be used but the actions will be passive. Thus, the information will be interpreted by the augmented reality device without the interaction of the user. Mixed reality will make a way to interact naturally and freely using gestures, gaze or voice.

- **The type of information and data managed**: In both technological cases, the data managed is virtual, it is possible to take for example contextual information in augmented reality (text, point of interest, etc.). For mixed reality, virtual object will exist in space and time and their positions will depend on the user's environment.

The table below [Table 2.] shows a summarized comparison of augmented, virtual and mixed realities paradigm based on the 3iVClass.

Table 2. Virtual, Augmented and Mixed realities differentiation according the « 3iVClass » classification method.

"3iVClass" criteria	Augmented Reality	Virtual Reality	Mixed Reality
Immersion	Augmentation of the real world using virtual annotations.	Fully virtual environment.	Spatial mapping in real time allowing data contextualization
Interaction	Mediate interaction with physical object.	Mediate interaction with virtual object (using controller [Fig. 4.])	Immediate interaction with virtual and physical objects
Information	Virtual annotation within the real environment.	Virtual object registered in 3D space.	Virtual object registered in 3D space.
	Not time-persistent.	Non time-persistent.	Time persistent.
	Decorrelated from user space.	Decorrelated from user space.	Correlated to the user space.

5. Conclusion

Based on the existing multiple interpretations of the scientific community about virtual, augmented and mixed realities, this paper proposes a clear and objective definition of each of those three paradigms. According to the existing taxonomies, we proposed a classification method based on three criteria called 3iVClass. These three criteria make it possible to group together the various components of the virtual, augmented and mixed realities according to: 1) the type of Immersion to which the user is confronted; 2) the means of Interaction which he has at his disposal and finally; 3) the type of Information and data handled. We tested this classification on the most confuse paradigm, namely mixed reality, in order to give a standardized definition. It allowed us to highlight the major divergences between augmented, virtual and mixed realities, which are mainly are found in the response to the well-known SLAM problem. In other words, it is the ability to interact with both physical and virtual objects which are registered in time and according to the user's environment. We can clearly identify the benefits of such a classification, which allows us to easily establish a reliable definition method for each of these realities or future similar paradigms.

References

[1] Sutherland IE, editor. (1968) "A head-mounted three dimensional display". *Fall joint computer conference, part I*.
[2] Starner T, Mann S, Rhodes B, Levine J, Healey J, Kirsch D, et al. (1997) "Augmented reality through wearable computing". P*resence: Teleoperators and Virtual Environments* **6(4)**:386-98.
[3] Azuma RT. (1997) "A survey of augmented reality". *Presence: Teleoperators and virtual environments* **6(4)**:355-85.
[4] Milgram P, Kishino F. (1994) "A taxonomy of mixed reality visual displays". *IEICE TRANSACTIONS on Information and Systems* **77(12)**:1321-9.
[5] Rekimoto J, Nagao K, editors. (1995) "The world through the computer: Computer augmented interaction with real world environments". *Proceedings of the 8th annual ACM symposium on User interface and software technology* : ACM.
[6] Merriam-Webster. https://www.merriam-webster.com/dictionary/interface: Merriam-Webster; 2018. "Interface".
[7] Billinghurst M, Kato H, editors.(1999) "Collaborative mixed reality". *Proceedings of the First International Symposium on Mixed Reality*.
[8] Pan Z, Cheok AD, Yang H, Zhu J, Shi J. "Virtual reality and mixed reality for virtual learning environments. Computers & graphics. 2006;30(1):20-8.
[9] Kress BC, Cummings WJ, editors. (2017) "Invited Paper: Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices". *SID Symposium Digest of Technical Papers; 2017: Wiley Online Library*.
[10] Kress BC, Cummings WJ, editors. (2017) "Optical architecture of HoloLens mixed reality headset". *Digital Optical Technologies.*
[11] Mann, S. (1999). Mediated reality. *Linux Journal*, *1999*(59es), 5.
[12] Merriam-Webstercom. (2018) "mediate". *Merriam-Webstercom.com*
[13] Merriam-Webstercom. (2018) "immediate". *Merriam-Webstercom.com*
[14] Kinnebrew PT, Kipman A. "Mixed reality presentation". *Google Patents*; 2012.
[15] Fox D, Burgard W, Dellaert F, Thrun S. (1999) "Monte carlo localization: Efficient position estimation for mobile robots". *AAAI/IAAI*.**(343-349)**:2.
[16] Durrant-Whyte H, Bailey T. (2013) "Simultaneous localization and mapping: part I". *IEEE robotics & automation magazine*. **13(2)**:99-110.
[17] Nowozin S, Adam A, Mazor S, Yair O. (2018) "Depth from time-of-flight using machine learning". *Google Patents*.
[18] Dunn M, Pavan D, Ramirez P, Rava S, Atiqah S. (2018) "An automated method to extract three-dimensional position data using an infrared time-of-flight camera". *MDPI Proceedings*. **2(6)**:502-207.
[19] Ganapathi V, Plagemann C, Koller D, Thrun S, editors. (2010) "Real time motion capture using a single time-of-flight camera". *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on: IEEE*.
[20] Bray B, Zeller M. (2016) "Mixed reality principal components". *Microsoft Windows Dev Center: Microsoft Doc*.