Data Markup Representation for Mixed Reality Contents

Hyejin Kim[1], Hyoseok Yoon[1], Ahyoung Choi[1], Woonhyuk Baek[1],
Ilgu Lee[1], Dongchul Kim[2] and Woontack Woo[1,2]

[1]GIST U-VR Lab., Gwangju 500-712, South Korea
{hjinkim, hyoon, achoi, wbaek, ilee, wwoo}@gist.ac.kr

[2]GIST CTI, Gwangju 500-712, South Korea
dckim@gist.ac.kr

Introduction

Mixed Reality (MR) is defined as the merging of real and virtual worlds on virtuality continuum [1], also shown in Figure 1. In other words, it includes Augmented Reality (AR) and Augmented Virtuality where real and virtual objects co-exist and interact in real time. MR contents are necessary elements to build Mixed Reality applications and real and virtual world information should be synchronized and co-exist or connected in the both world.

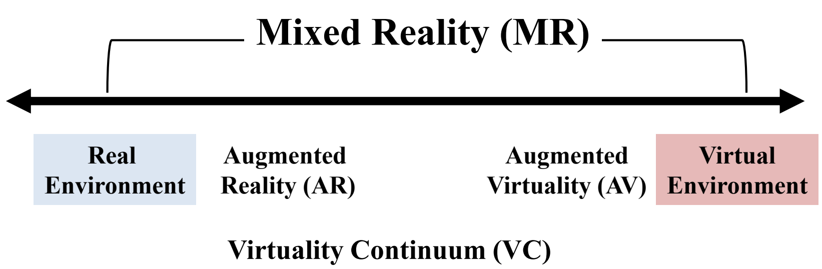

Figure 1. Mixed Reality defined on virtuality continuum [1]

Recently there are many research projects on Mixed Reality such as Ubiquitous Virtual Reality (U-VR or ubiquitous VR) [2] and Cross-Reality (X-reality) [3], which facilitate seamless information exchange and synchronization between the real and virtual world. Furthermore, AR contents have been increasing as AR browser applications such as Junaio [4], Wikitude [5] and Layar [6] appeared and became popular. Therefore, at this stage, we need a standard method for representing the increasing AR contents consistently for interoperation. For this reason, ARML [7] and KARML [8] provide markup language for representing AR contents. Although both approaches are based on KML [9], the former uses subset of KML with the object of data representation and the latter uses full KML for the purpose of visualization. These different characteristics raise incompatibility issues so that each representation operates on only its own approach. In addition, aforementioned approaches aim to particularly represent loosely geo-tagged contents positioned using GPS and compass sensor. As a result, it is not sufficient to represent Mixed Reality contents accurately registered in the real world, so we need a new standard on data markup representation for Mixed Reality contents. The new standard is required to encompass existing representation of AR contents and express MR contents with platform independency and easy flexibility to various MR applications.

Data Markup Representation for Mixed Reality Contents

First of all, we consider applications of Mixed Reality contents as follows.

Application 1) Augmenting location or fixed object with virtual information

The Namsan Tower, which is famous tower in Korea, is real world information and 3D Namsan Tower model is virtual world information. If a user looks at the Namsan Tower through a location sensor and camera equipped mobile device, the user can see the real Namsan Tower augmented with virtual 3D Namsan Tower model.

Application 2) Augmenting movable object with virtual information

A brochure of the Namsan Tower is real world information and 3D Namsan Tower model is virtual world information. When a user looks at the brochure with a mobile device equipped with camera, he or she can see the real brochure of the Namsan Tower augmented with virtual 3D Namsan Tower model registered continuously wherever the brochure is moved.

Application 3) Augmenting fixed object and movable object with virtual information

User A is real world information and 3D arrow model pointing from user A to his or her house is virtual world information. When user B looks at user A through a mobile device with a location sensor and camera, user B can see user A augmented with updated virtual arrow model indicating location and direction of user A's house from wherever user A is.

Application 4) Creating augmentation of other people's object with your own virtual information

A picture frame is real world information and belongs to user A. An image is virtual world information and belongs to user B. When user B want to see the picture frame of user A augmented with the image, first user A needs to permit the access, then user B can make the linkage between the picture frame and the image for augmentation.

Application 5) Playing virtual information augmented onto a real object with other people

A picture frame is real world information and an image augmented onto the picture frame is virtual world information. If this augmentation information is set as public access mode and user A moves the image's location onto the picture frame, user B can see the result of moving the image and vice versa.

Besides these five cases, there are other various applications of Mixed Reality contents such as augmenting real world with virtual information changed with time, augmenting real world with synchronized virtual information and vice versa, and augmenting real world with virtual contents from mashup services.

In order to represent Mixed Reality contents applicable to above MR applications, we propose the general concept based on an idea depicted in Figure 2. As mentioned above applications, real world information can be anything in the real world such as space, objects, people, etc. We call representation of real world information with metadata *real contents*. Virtual world information, which is shown onto augmented real world information, is any digital information such as image, 3D model, text, video, audio, etc. Similar to *real contents*, representation of virtual world information with metadata is called *virtual contents*. And we provide a link between real contents and virtual contents to augment real or virtual world with corresponding virtual or real world information. We call this representation of link with metadata *real-virtual contents link*. Now, Mixed Reality content is general term for these three components – *real contents*, *virtual contents*, and *real-virtual contents link*.

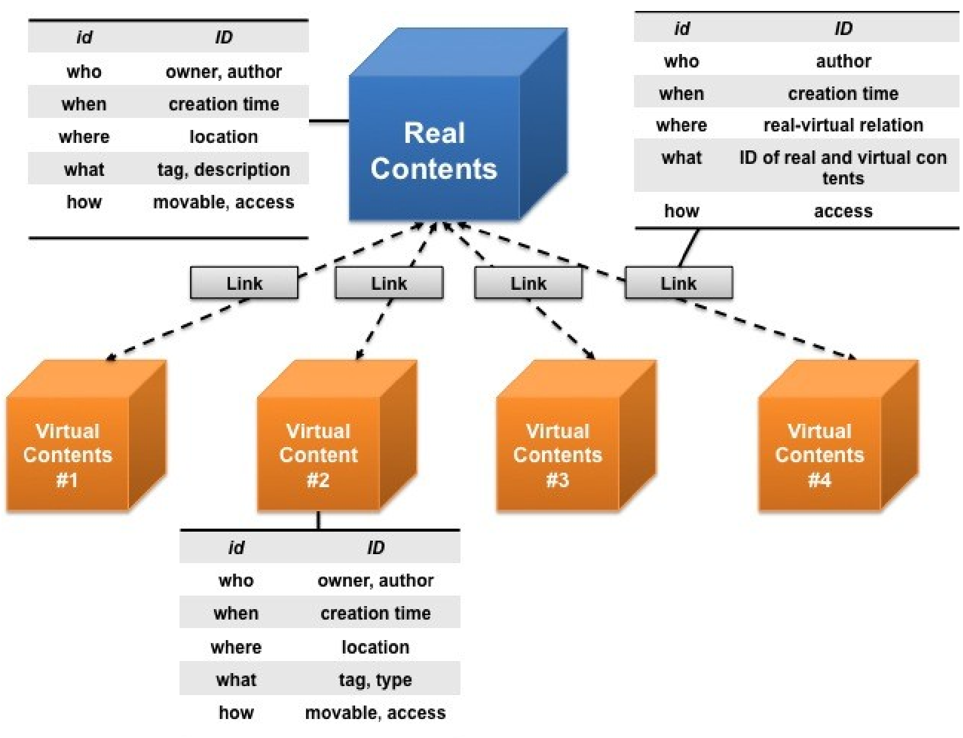

Figure 2. The diagram of mixed reality contents representation

Below are some issues to consider when we build Mixed Reality contents applications, and the tags or fields in a form of 4W1H(Who, When, Where, What, How) which is used to represent MR contents consistently. To identify each component in MR contents applications, each component needs to be automatically assigned with a unique ID.

1) Real contents representation

• Who: *Author* and *owner* who make the real world information as real contents.

• When: Real contents *creation time* for the application.

• Where: *Location* of real world information for location based augmentation or global coordinate based applications.

• What: *Description*, which is data for recognizing real world information in an object-based augmentation. *Tag*, which is an explanation of real world information for search based filtering application.

• How: *Fixed* or *dynamic*, is a property for movable objects based augmentation. *Private* or *public*, is a property for sharing application. *Responsive* or *non-responsive*, is a property for the application using intelligent response of real world contents according to environmental input sources.

2) Virtual contents representation

• Who: *Author* and *owner* who make virtual world information as virtual contents.

• When: Virtual contents *creation time* for the application.

• Where: *Location* of virtual world information on the web for accessing the data.

• What: Data decoding *type*, which is used for rendering virtual world information. *Tag*, which is an explanation of virtual world information for search based filtering application.

• How: *Fixed* or *dynamic*, *private* or *public*, *responsive* or *non-responsive* for the same reasons as mentioned in real contents representation.

3) Real-virtual contents link representation

• Who: *Author* who make real-virtual contents link.

• When: Real-virtual contents link *creation time.*

• Where: *Relation* between real and virtual world information for augmenting real world with virtual world and vice versa.

• What: *ID of real contents* and *ID of virtual contents* connected by composing this link.

• How: *Private* or *public* for the same reason as mentioned in real contents representation.

By using 4W1H in our MR content representation, it provides a consistent and logical structure and association between real content, virtual content and link. As shown in Table 1, with 4W1H, it is intuitive to identify different attributes of MR content. Similarly, POI core principles [10] which highlights Where can be distributed in a form of 4W1H.

Table 1. 4W1H structure

4W1H	MR Content	POI [10]
WHO	-Owner, author	-"meta data" primitive (authorship)
WHEN	-Creation time	-"meta data" primitive (creation time)
WHERE	-Location	-"id" primitive -Location primitive, (is at) -"contained within", "contains", "adjacent-to"
WHAT	-Tag, description (how to detect)	-"name" primitive (human description) -"categorization" primitive
HOW	-Movable, access	-"meta data" primitive (visibility rights)
Extensibility	-Real Content, Virtual Content, Link	-(method of payment, opening closing hours, 3D content, images, multimedia…)

Examples/Prototypes

In this section, we show an example using the proposed representation for the third Mixed Reality application scenario mentioned above. First and second scenarios can be similarly represented.

Example 3) Augmenting fixed object and movable object with virtual information


```
<realcontent id="{4CDCF772-DAB2-4C7D-A1EB-7B781DF4DFB7}">
  <who>
    <owner id="{7DD22961-78EE-4B66-8766-F6B388F6B9DD}"/>
    <author id="{7DD22961-78EE-4B66-8766-F6B388F6B9DD}"/>
  </who>
  <when>2010-09-24T09:34:00 </when>
  <where>
    <latitude>37.551425</latitude>
```

```
    &lt;longitude&gt;126.988&lt;/longitude&gt;
  &lt;/where&gt;
  &lt;what&gt;
    &lt;description&gt;
      &lt;tag&gt;&lt;![CDATA[User A]]&gt;&lt;/tag&gt;
      &lt;img&gt;http://uvr.gist.ac.kr/imgs/usera1.jpg&lt;/img&gt;
      &lt;img&gt;http://uvr.gist.ac.kr/imgs/usera2.jpg&lt;/img&gt;
      &lt;img&gt;http://uvr.gist.ac.kr/imgs/usera3.jpg&lt;/img&gt;
    &lt;/description&gt;
  &lt;/what&gt;
  &lt;how&gt;
    &lt;movable&gt;dynamic&lt;/movable&gt;
    &lt;access&gt;public&lt;/access&gt;
  &lt;/how&gt;
&lt;/realcontent&gt;

&lt;virtualcontent id="{3EA3D510-4673-486E-83BD-2C578302DCCB}"&gt;
  &lt;who&gt;
    &lt;owner id="{BC920170-B300-4449-A047-6FD496B79D88}"/&gt;
    &lt;author id="{7DD22961-78EE-4B66-8766-F6B388F6B9DD}"/&gt;
  &lt;/who&gt;
  &lt;when&gt;2010-09-24T09:40:00&lt;/when&gt;
  &lt;where url="http://uvr.gist.ac.kr/userahousedirection.collada"/&gt;
  &lt;what&gt;
    &lt;description&gt;
      &lt;tag&gt;&lt;![CDATA[arrow from user A to user A's home]]&gt;&lt;/tag&gt;
      &lt;type&gt;&lt;![CDATA[collada]]&gt;&lt;/type&gt;
    &lt;/description&gt;
  &lt;/what&gt;
  &lt;how&gt;
    &lt;movable&gt;dynamic&lt;/movable&gt;
    &lt;access&gt;public&lt;/access&gt;
  &lt;/how&gt;
&lt;/virtualcontent&gt;

&lt;link id="{CE5BA114-04F6-47DB-B955-E34B228B0EA9}"&gt;
  &lt;who&gt; &lt;author id="{7DD22961-78EE-4B66-8766-F6B388F6B9DD}"/&gt; &lt;/who&gt;
  &lt;when&gt;2010-09-24T09:41:00&lt;/when&gt;
  &lt;where relation="absolute"&gt;
    &lt;latitude&gt;37.551425&lt;/latitude&gt;
    &lt;longitude&gt;126.988&lt;/longitude&gt;
  &lt;/where&gt;
  &lt;what&gt;
     &lt;rid&gt;4CDCF772-DAB2-4C7D-A1EB-7B781DF4DFB7&lt;/rid&gt;
     &lt;vid&gt;3EA3D510-4673-486E-83BD-2C578302DCCB &lt;/vid&gt;
  &lt;/what&gt;
  &lt;how&gt; &lt;access&gt;public&lt;/access&gt; &lt;/how&gt;
&lt;/link&gt;
```

For the fourth scenario, user A makes the picture frame as real contents whose owner/author is user A and the access property of how field is set to public. Then, user B makes the image as virtual contents whose owner/author is user B. Finally, user B also makes the linkage between the picture frame and the image as real-virtual contents link. In the fifth scenario, user A and user B can edit and save real-virtual contents link where the access property of how field is set public.

Benefits of Using Mixed Reality Contents Data Markup Representation

Existing AR contents representation methods such as ARML, KARML also represent necessary information to augment real world with virtual world information according to the real world. However, it is limited in several aspects. First of all, it is difficult to represent real and virtual world for object-based augmentation because it uses global coordinate's location like longitude and latitude when representing real world information rather than local coordinate's location for object-based recognition. Therefore it is impossible to augment real fixed and movable objects with virtual world information. Also, it is difficult to extend the range of applications beyond loosely geo-tagged information visualization. Last but not least, our MR content representation provides structured fields that can be used for various operations within MR applications, such as filtering by provider, time or location for different application scenarios.

There are a number of benefits of the proposed data markup representation for Mixed Reality contents. Firstly, it covers typical existing representation of AR contents and represents MR contents which other methods does not cover. Secondly, it is based on a clear concise concept of Mixed Reality so that it can be applied to various Mixed Reality applications easily. Finally, sharing the proposed xml based representation provides interoperability between MR applications over different platforms.

Conclusion

In this paper, we explained our position and the points to be specially considered for data markup representation of Mixed Reality contents. In summary, we tried to make a clear and concise representation that match a concept of Mixed Reality. Also, we considered how to encompass existing representation of AR contents and extend new representation of MR contents. Additionally, we wanted to achieve interoperability of the representation, which means platform independency and flexibility to various MR applications.

For this reason, we proposed data markup representation for Mixed Reality contents, which is useful to various MR applications. It is divided into three parts, "real world information representation called *real contents*", "virtual world information representation called *virtual contents*" and "real and virtual world link representation called *real-virtual link contents*".

Acknowledgement

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA), under the Culture Technology (CT) Research & Development Program 2011.

Reference

[1] P. Milgram and A. F. Kishino, “Taxonomy of Mixed Reality Visual Displays”, IEICE Transactions on Information and Systems, E77-D(12), pp. 1321-1329, 1994.

[2] Y. Lee, S. Oh, C. Shin and W. Woo, “Ubiquitous Virtual Reality and Its Key Dimension”, IWUVR09: International Workshop on Ubiquitous Virtual Reality, pp. 5-8, 2009.

[3] J. Lifton, M. Laibowitz, D. Harry, N. Gong, M. Mittal and J. A. Paradiso, “Metaphor and Manifestation - Cross-Reality with Ubiquitous Sensor/ Actuator Networks”, IEEE Pervasive Computing, vol. 8, no. 3, pp. 24-33, 2009.

[4] Junaio, http://www.junaio.com/

[5] Wikitude, http://www.wikitude.org/

[6] Layar, http://www.layar.com/

[7] ARML, http://www.openarml.org/

[8] KARML, https://research.cc.gatech.edu/polaris/content/karml-reference

[9] KML, http://code.google.com/intl/ko-KR/apis/kml/

[10] Core Principles of a POI format, http://www.w3.org/2010/POI/wiki/Core_Principles