INTERNATIONAL STANDARD

ISO/IEC 21145

First edition
2023-09

Information technology — Computer graphics, image processing and environmental data representation — Style representation for mixed and augmented reality

Reference number
ISO/IEC 21145:2023(E)

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

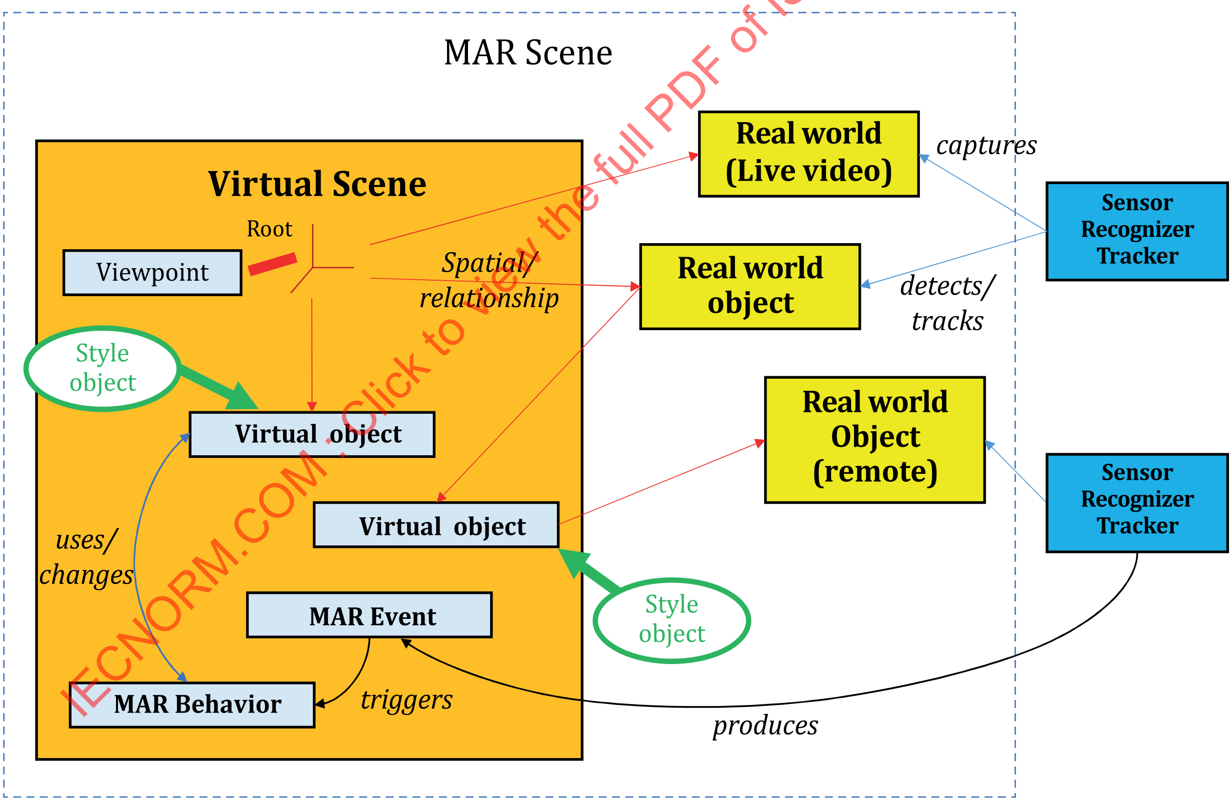
This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 24, *Computer graphics, image processing and environmental data representation*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Mixed and augmented reality (MAR) refers to a spatially coordinated combination of media/information components that represent, on the one hand the real world and its objects and on the other hand those that are virtual, synthetic and computer generated. MAR as an information medium strives to provide rich experience based on realism, presence and augmentation [ISO/IEC 18039].

This document describes a set of information constructs for stylizing the MAR content and its objects. ISO/IEC 3721 specifies the MAR scene with the core objects and their attributes [ISO/IEC 3721] This document further refines ISO/IEC 3721 similarly to how CSS (Cascading Style Sheets)[2] augments HTML-5 (Hyper Text Markup Language)[1] for web document stylization. Among the various core MAR content objects defined in ISO/IEC 3721, virtual objects that augment real world objects may be subject to further stylization. Stylization refers to the act of making certain information follow a particular form for various purposes. As MAR contents become more widespread and sophisticated, there is an increasing need for its stylization and its separation from the core content for its efficient specification. The augmentation style can affect the whole MAR experience in terms of its naturalness (how harmonious objects appear in relation to the real world into which they are inserted), and conspicuousness and saliency (how augmentations stand out for visibility and readability). This in turn can be important for efficient information transfer through the MAR content. Style specification components are thus associated with and applied to the virtual objects used for augmenting the real world objects in the MAR content (see Figure 1).

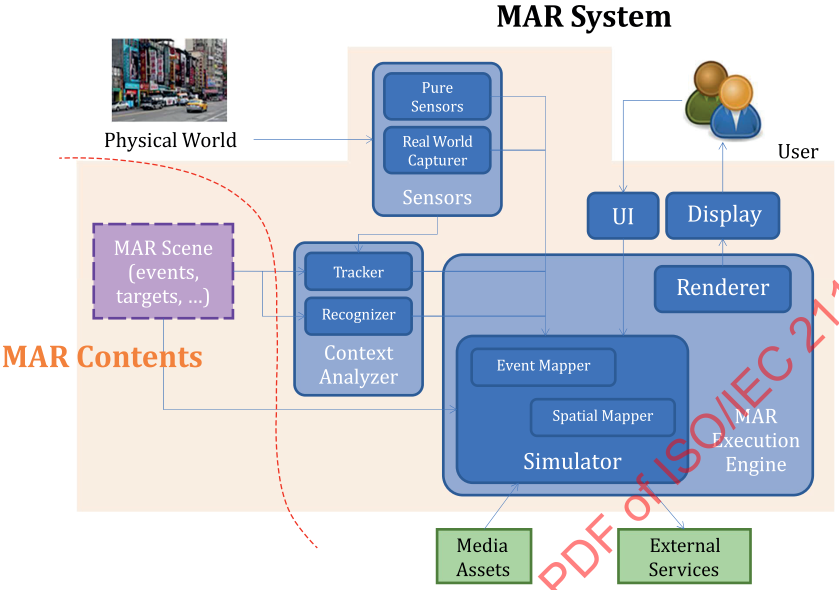


The real world objects are associated or augmented with the virtual scene/objects shown within the left box. The style objects represent the constructs for further specifying their "styles" for presentation to the users.

Figure 1 — The general structure of mixed and augmented reality (MAR) content [ISO/IEC 18039]

This work only establishes the information model, and neither promotes nor mandates the use of a specific language, file format, algorithm, device, or implementation method. It is to be considered as a minimum basic model and a sound basis that can be extended for other purposes in actual implementations (e.g. application standards, specific file formats). It is designed for ease, generality, and extendibility and is illustrated in this document with various examples and implementation results.

This document is based on ISO/IEC 18039 which specifies a reference architecture for the MAR system as a contents-browser/player. In ISO/IEC 18039, the MAR content is specified as the input to the MAR system that describes the scene and objects' behavior. The MAR system parses, simulates and renders the content to the display (See Figure 2).



The MAR Scene or equivalent content description is an input to the larger MAR system which will interpret it and render it to a display for user consumption [ISO/IEC 18039].

Figure 2 — The generic MAR system architecture

Information technology — Computer graphics, image processing and environmental data representation — Style representation for mixed and augmented reality

1 Scope

This document specifies:

1) Constructs for representing and specifying various augmentation and presentation styles. While augmentations can be in modalities other than the visual (e.g. aural, haptic), this work addresses the visual augmentation style only.

2) A model for how to associate the stylization constructs to the augmentation objects. Specifically, the MAR behavior object in ISO/IEC 3721 is extended for this purpose.

3) Other miscellaneous functionalities and abstractions that support the stylization of augmentation objects.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 18039, *Information technology — Computer graphics, image processing and environmental data representation — Mixed and augmented reality (MAR) reference model*

ISO/IEC 3721, *Information technology — Computer graphics, image processing and environmental data representation —Information model for mixed and augmented reality content — Core Objects and their Attributes*

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 18039, ISO/IEC 3721, and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1.1
stylization
act of enforcing certain information presentation to follow a particular form or behave in a given way for various purposes such as to highlight it, emphasize it, or make it more readable and perceptible

3.2 Abbreviated terms

AR	augmented reality
AV	augmented virtuality
CSS	cascading style sheet
HTML	hypertext markup language
MAR	mixed and augmented reality
MAR-RM	mixed and augmented reality reference model
MR	mixed reality
UI	user interface
VR	virtual reality

4 Principles and requirements

The MAR content model should provide reasonable generality, being able to express various genres of mixed reality contents, especially AR and AV under a consistent and unified representational framework. Other requirements are as follows:

— independent of particular sensor/device model, algorithm or implementation platform,

— be able to handle virtually any digital information and media, both static and dynamic, as augmentation such as text, images, videos, animation, HTML document elements, etc.,

— provide useful abstract and declarative constructs for often used content functionalities and minimize manual scripting or programming,

— make use of existing standard constructs, where possible,

— flexible/extendible to accommodate new future requirements.

To fulfil these requirements, this document will continue to adopt the component based approach, as defined in ISO/IEC 3721. Components are collections of objects that perform or represent similar operations, displays, or functions. The component framework is ideal for introducing and modelling new capabilities. New objects/styles can be added or expanded as needed. The components form a structured environment that ensures consistent behavior, usability and generality.

5 MAR content model: Target for stylization

To properly set the context for this work and aid understandability of this document, we first briefly outline the information model for MAR content [ISO/IEC 3721] to which the stylization is applied.

MAR contents as defined in ISO/IEC 3721 are technically realized as virtual reality contents with the scene containing "special type" objects that represent the real physical world and objects. In other words, one can think of a MAR scene as technically a virtual scene with "placeholders" for real world physical objects. The "placeholders" may be defined logically or spatially with respect to the virtual scene. These so called "special type" placeholder objects may need additional information of how they can be spatially registered into the virtual scene implementation. In AR, one form of MAR, virtual objects are augmented onto real world physical objects; in AV, vice versa. Thus, depending on the type of MAR content, both virtual and real world physical objects may be subject to stylization.

Just like the virtual reality environment or contents, the scene graph or scene tree structure would be the most natural and suited representation in which the MAR content is represented as a hierarchical and spatial organization of various types of objects. While the objects are in principle organized by their spatial relationship ultimately with respect to the assumed root coordinate system, logical association and aggregation are possible too.

Figure 3 illustrates the basic concept in which a virtual object (or the root of the virtual scene) provide a place for a certain real world physical object (remote or local) to map to in the MAR scene. For example, a virtual object can be spatially defined with respect to a real world physical object (or its placeholder). The dotted arrows in the figure indicate spatial relationships among the captured/tracked real world physical objects with respect to the virtual scene/objects. The dash-dotted objects are associated with the virtual objects to indicate their style in visualization. Association relations may exist among various objects that can refer and relate to each other for different purposes. The association may be one directional or two directional (the dashed or solid arrows). For example, styles can be assigned to content objects though mutual associations. The figure shows the overall class diagram as the gross information model for MAR contents.

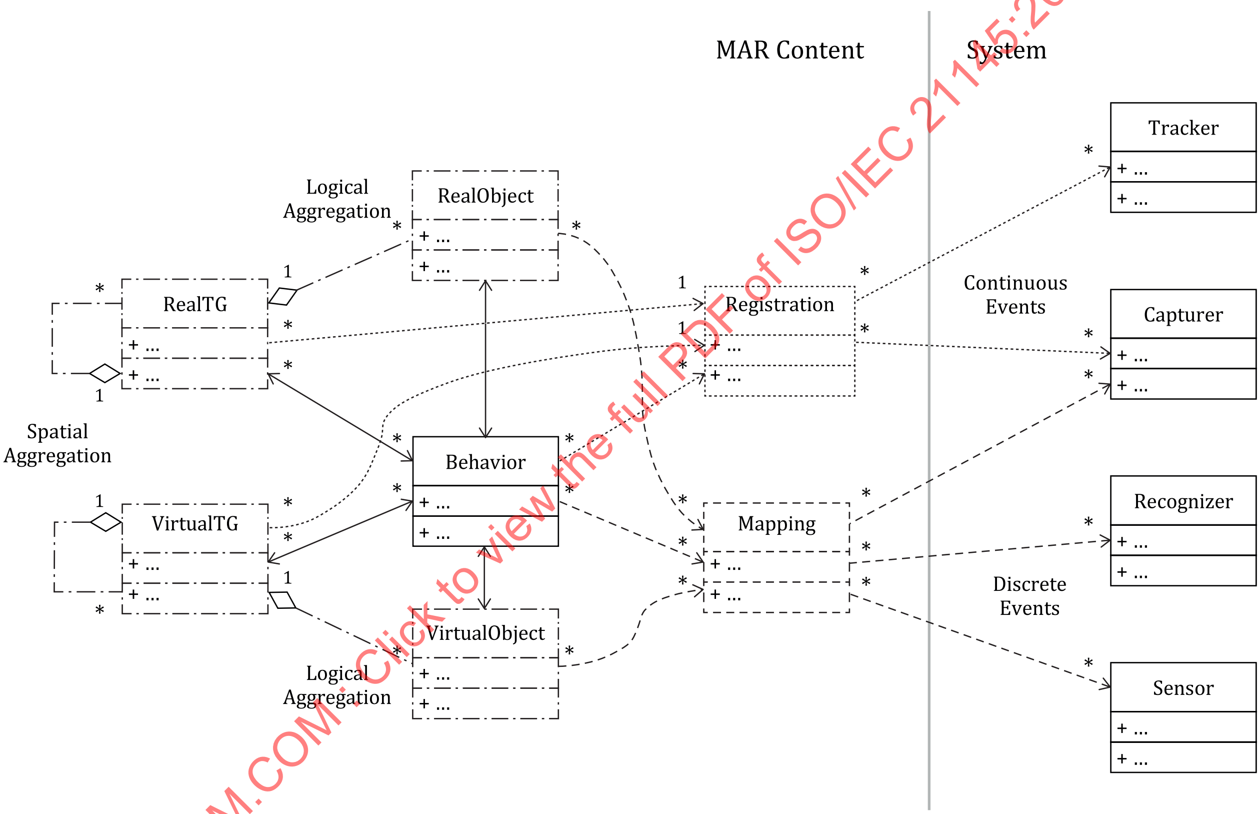


Figure 3 — Relationships among classes in the MAR scene structure [ISO/IEC 3721].

6 MAR stylization as MAR behavior components

6.1 MARSNode

Stylization refers to the act of enforcing that certain information (content elements) follow a particular form (or behave in a given way) for various purposes. Thus, the stylization components derive out of the “MAR Behavior” object of ISO/IEC 3721. MAR behavior objects are refined to express stylization. In ISO/IEC 3721, all MAR content components are subclasses of the abstract MARSNode, including the MAR behavior node as the stylization components. MARSNode represents the abstract superclass that can take a place in the hierarchical (tree) representation of the scene. For more details, refer to ISO/IEC 3721.

6.2 Behavior::MARSNode

Behavior class specifies the dynamics of the objects in the scene (e.g. time based or event based behavior). Therefore, such behaviors can be expressed simply by relying on the usual scripts. Behavior nodes, in addition to housing such flexible scripts, serve to abstract some typically used MAR "augmentation" behaviors (i.e. how virtual augmentation objects associated and spatially registered to a real physical object behave when reacting to input and external events). The main purpose of such an abstraction is ease of use, simplicity and therefore quick authoring.

Behavior nodes are driven by external events/data produced by the MAR system components (such as the sensor, capturer, tracker and recognizer – see ISO/IEC 3721) and simulated/executed by the MAR simulation engine (see ISO/IEC 18039). The behavior node will have associations with the content objects (which exhibit the behavior) as well as other objects affected by the behavior simulation/execution. Typical examples of often used augmentation behavior are simple activation/deactivation of the augmentation to be visible, simple animation, highlighting effects, and changing of transparency or color. See Table 1.

In this document, we extend this concept to express "augmentation style".

Table 1 — A prototypical class specification for Behavior::MARSNode

Behavior			
Access type	**Data type**	**Attribute/Method name**	**Explanation**
private	int	type	specifies the specific type of the behavior
private	bool	isTriggered	enabling or disabling the triggering of this Behavior node (0 for disabling it, and 1 for enabling it)
private	Mapping[]	eventSource	list/array of source Mapping node(s) that delivers and filters the event/data to this node from the MAR System components
private	MARSGNode[]	sourceObjects	list/array of source objects referred and used for behavior definition
private	MARSGNode[]	targetObjects	list/array of MARSNodes that are affected by this Behavior node
private	VirtualEvent*[]	virtualTriggers	list/array of virtual events produced by the associated Mapping node for triggering this Behavior node
private	RealEvent*[]	realTriggers	list/array of real world physical events produced by the associated Mapping node for triggering this Behavior node
private	string*	scriptFilePath	file path to a script if any
public	Behavior*	Behavior()	constructor
protected	void	init()	initialization
private	Void	triggerBehavior()	trigger the behavior as specified in this node by updating and retrieving the most recent events/data from the associated Mapping nodes and update the MAR scene structure
public	String	getScriptFilePath()	return the script file path
public	Void	setScriptFilePath(string path)	set the script file path

Two simple usage examples are shown below. In the first example, a simple augmentation behavior, just making the virtual object appear (be visible) at some default location relative to the marker is specified. This "visibility" behavior is applied to the object, "v1" and is triggered by the event, "content_marker1_present". The second example illustrates the use of a script, which is triggered by a predefined event, "mouse_click" and applied to the object, "reg21" (the script uses this object as its argument).


```
% an event called "content_marker1_present" triggers virtual object "v1" to be visible
<behavior id="b1"
    type = 1 % this behavior is simple visibility behavior
eventSource = [ … ]
    realTriggers=["content_marker1_present"]
    targetObject = ["v1"] />
```

```
% a mouse click event triggers a script based behavior for virtual object "reg21"
<behavior id="b2"
    type = 0 % this behavior is script driven behavior
    eventSource = [ … ]
    realTriggers=["mouse_click"]
    targetObject = ["reg21"]
    scriptFilePath = "../script.py" />
```

6.3 BehaviorStyleVisual

Augmentation objects exhibiting a certain presentation style, especially those that are dynamic and time-varying or situation based, can be expressed by extending the Behavior node. There are three such subclasses: BehaviorStyleVisual, BehaviorStyleAural, and BehaviorStyleHaptic, each representing styling in the respective modalities. This document addresses only the stylization for visual augmentation; the detailed information model for the latter two are not included. See Figure 4.

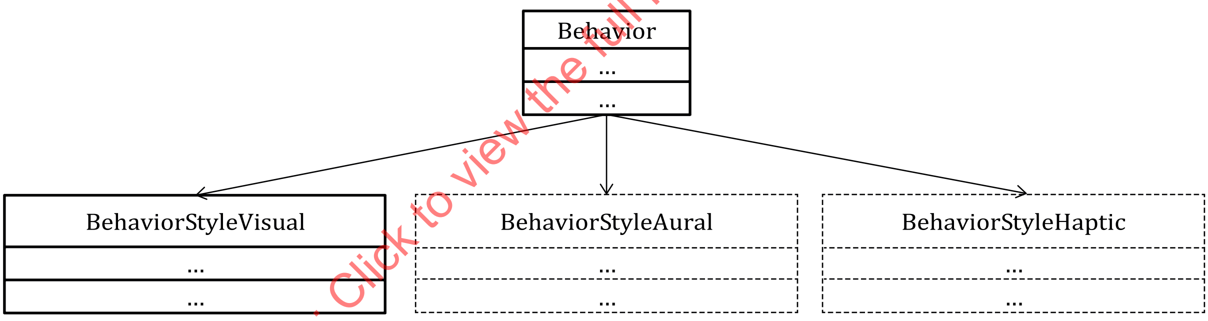


Figure 4 — Subclasses of Behavior node for visual augmentation style specification

BehaviorStyleVisual contains the attributes that are common to main visual styling information. However, the actual required data may differ depending on the particular type of the augmentation object being stylized. For instance, the transform attribute specifies the position and orientation of the augmentation object relative to the object being augmented. If the augmentation object is a 2D object, the transform attribute requires four real number entries for the rotation matrix and two for the translation vector. If the augmentation object is a 3D object, then the same attribute will have nine real numbers for rotation and three for translation. Therefore, a visual object_type is specified first, such as text, image, video, 3D model, billboard, etc.

Similar type-dependent attribute value specification and understanding will be needed for the other attributes such as size, opacity, view and emphasis. Size refers to the size of the augmentation relative to that of the object being augmented. Visibility controls whether the augmentation is visible or not, and also its opacity. View specifies for which view into the augmentation the given style is valid. For now there are six principal directions (front, back, top, bottom, right, left). For instance, a blinking 2D image augmentation may be made visible only when viewed from the front direction (or near front direction). Finally, the emphasis attribute is used to specify some typical emphasis style with default parameter values, such as blinking, reverse video, edge/border highlighting, and contrast (with background or environment). See Table 2 for more details.

Table 2 — A prototypical class specification for BehaviorStyleVisual::Behavior

Attribute/Method name	Explanation
object_type	specifies to which type of visual augmentation object the styling is applied to: text, image, video, 3D model, billboard, etc., and is identified with an unique integer
transform	specifies the initial and default position and orientation of the augmentation object relative to the object being augmented 2D objects: four real numbers for rotation, two for position/translation vector 3D objects: nine real numbers for rotation, three for position/translation vector
size	specifies the initial and default size of the augmentation object relative to the object being augmented Text: integer 2D object, image, video, billboard: two real number scale factors in local x and y with respect to those of the object being augmented 3D objects: three real number scale factors in local x, y and z with respect to those of the object being augmented
visibility	specifies the visibility and transparency of the augmentation object and is indicated with a real number between 0 and 1
view	list/array of camera-to-object views for which this stylization is valid, there are only six discrete views – front, back, top, bottom, left and right.
color	three integers indicating the red, green and blue component of a color applied globally to the augmentation object
font	font type (applicable to text stylization only)
font_size	size of the text (applicable to text stylization only) as a single integer
emphasis	specifies a variety of emphasis and highlighting visualization patterns: border, edge, silhouette, blinking, reverse video, contrast, motion and is identified with a unique integer code

While the BehaviorStyleVisual construct can be used for a variety of 2D and 3D augmentation objects, they can only cover simple and often used style behaviors – more complex ones should be encoded using the scripts.

7 Example usage

Here are three examples of how various style specifications can be made for the augmentation objects in a simple way using the BehaviorStyleVisual construct. In the first example, a style specification object “st1” is created and associated with the augmentation object, “text1” which is 3D text. The 3D text is stylized with a specific font type, size and color with its location translated 100 pixels in the y direction (by the local coordinate system of the object being augmented), and rotating around the y axis.

```
% Augment by event content_marker1_present with 3D text
% translated 100 in y direction, times roman font of size 10 and RGB of 100,15,44
rotating around y axis
<bahaviorStyleVisual id="st1"
    type = “3DText”
    targetObject=“text1“
    realTriggers=“content_marker1_present”
    transform=[1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 100, 0]
    font="Times_Roman"
    font_size=“10”
    color="100, 15, 44“
```