Virtual Reality Callouts - Demonstrating Knowledge With Spatial-Related Textual Information

Robin Horst and Anika Degreif† and Marvin Mathy† and Ralf Dörner

RheinMain University of Applied Sciences, Faculty of Design Computer Science Media, Wiesbaden, Germany

Abstract
Virtual (VR) and augmented reality (AR) can bring an added value during the demonstration of knowledge, as for example within an interactive research demo. Callouts are short text strings which are connected by a line to a specific feature of an object. These visual annotations can be used during such demos and can be placed in different kinds of media, such as illustrations, technical drawings, images and videos. Callouts are also used in virtual 3D environments to anchor textual information to a specific point in space. Therefore they can be a valuable tool for virtually demonstrating knowledge. The alignment of callouts in such information rich environments is an elemental factor within the view management of the VR scene.

In this paper we propose a concept for interactive microlearning application for knowledge demonstration that uses callouts as a fundamental element. We distinguish three types of interactive callout-representations by their alignment relative to the user, for being static or dynamic in their position and orientation. Within an implementation of the different callout versions we show the feasibility and in a user study we indicate a user-preference towards static positioned callouts.

CCS Concepts
• Human-centered computing → Virtual reality; Information visualization; Visualization systems and tools; • Applied computing → Annotation;

1. Introduction
Short text strings which are connected by a line to a specific object are called callouts. The text is mostly outlined or enclosed by a certain shape - the callout body. Callouts are often utilized in

† These authors contributed equally to this paper.

© 2019 The Author(s)

DOI: 10.2312/cgvc.20191252
technical drawings, illustrations, videos or other visual media. They are used to give information about a specific feature of the object. They often transport very concrete information and are often used to demonstrate knowledge, for example in learning applications. The concept of callouts is widely used in learning methodology and it is self-contained, so that a common application lies in the context microlearning. Microlearning [Hug05a; Hug05b] is a current trend [Giu17] in learning and education methodologies that relies on dividing learning content in relatively small and independent learning units. A typical example for the setting of our work therefore is research demonstration (demo) at a science fair, where a professor demonstrates new research outcomes in a face-to-face situation. Different media can be used during such a demo. One of them is Virtual Reality (VR) [HD18]. Mobile VR head-mounted displays (HMDs) for example offer a possibility to be used during a demo. Callouts can also be used in VR (Fig. 1). Since the visual communication is constrained in such a setup, VR callouts, however, offer a possibility to explain and describe information in this closed virtual environment.

Especially in information rich environments can interaction with callouts be difficult [IF19; CGN15; MKY09]. Information rich environments are virtual environments that are augmented by additional information, such as text, numbers or graphs [BNC*03]. Preventing objects from overlapping or occluding each other is a major challenge in these environments. Occlusions or visual clutter can occur based on the user’s perspective and position in the scene. These visual annotations must be regarded in the view management the virtual scene. They can be aligned differently and the connecting line can be anchored at the callout body at different locations. View management generally describes the process of maintaining visual constraints, for example by preventing occlusion of callouts.

Furthermore is the design of a virtual scene with callouts itself a complex task in the authoring process of VR applications. When the users are free to move within the scene, callouts may not be readable from every point in the scene of from every angle. But users may want to explore the virtual object from several perspectives. So the placement and orientation of callouts within the scene are elemental within the view management.

Different types of callouts can help to simplify this VR authoring sub-task. Within this paper we strive to explore different types of VR callouts within a self-contained microlearning demo application, based on their alignment. With this paper we make the following contributions:

- Three different types of callouts are identified. We differentiate them by a taxonomy of static or dynamic positioning and orientation. We propose a static positioned and static oriented callout, a static positioned callout with dynamic orientation towards the user, as well as a callout with dynamic position and orientation in the 3D space. These three concepts allow to apply callouts in information rich environments considering different view management aspects.
- In a user study we point out how these types of callouts are appealing for users and which one they prefer. We also indicate specific challenge regarding the different versions.
- We propose a microlearning application used for the field of knowledge demonstration which incorporates callouts as the main aspect. With this application we show the feasibility of the callout concepts. All three callout types are embedded within this prototypical implementation.

This paper is organized as follows. The next section shows related work concerning callouts. In the third chapter we describe the demo setting, the callout concepts and the highlights of the prototype implementation. Within a following user study, data about the callout concepts are collected and results are discussed. In the last section we draw conclusions about our work and indicate beneficial future research.

2. Related Work

Fekete and Plaisant [FP99] formally introduced excentric labeling with visual annotations as callouts within the field of data visualization. Callouts in virtual environments, such as VR or augmented reality (AR) are applied in a variety of use-cases, such as annotating cultural heritages with information [CGN15], 3D street maps [VFW13]. Existing work that we relate to deals with the technical creation of visual annotations in relation to the visibility and arrangement of them in terms of the view management for VR and AR scenes.

A current study by Ichihashi and Fujinami [IF19] describes to use machine learning techniques in a view management process to assess textual overlays in AR environments. They point out several features to incorporate to assess the visibility of text overlays and linkage lines and restrict their work to spatial AR. Work by Caggianese et al. [CGN16] evaluate two common approaches (Depth Ray and SQUAD) for implementation and interaction with annotations in an egocentric wearable AR setup. They could indicate that both disambiguation techniques for annotations lead to a certain fatigue within information rich environments. Makita et al. [MKY09] by contrast explores the impact of the actual size and shape of AR-annotations together with their actual arrangement. They show different implementations and propose a specific arrangement technique of callouts based on computer vision methodology. They consider recognition and segmentation relating to the occlusion of objects and assign a penalty score for areas of the image. A low penalty score indicates a good position for these callouts in spatial AR.

Shibata et al. [SNS*08] propose a view management method for labels that incorporates algorithms for seven challenges/functions, which are: (1) occluding real objects, (2) mutual overlap, (3) getting out of frame, (4) adjusting the size, (5) coloring annotations, (6) adjusting transparency and (7) highlighting a real object. Tatzgern et al. [TKGS14] explore callout annotations for AR and furthermore incorporate the aspect of consistency over time. These and further studies that investigate in different aspects of the view management of callouts (e.g. [MKY09; SZR13; FMS92; LWMS12; NBLB15; UMKT05]) indicate the importance of arranging callouts relating to visual clutter, visibility, occlusions or interactions. However they mostly address the field of spatial AR.

We also found few literature about creating and handling visual annotations similar to callouts in VR. Zhang and Sun [ZS05] propose a dynamic labeling method that can be applied for both AR and VR. In their system they avoid visual clutter and overlapping.
or occlusions by an adaptive placement, view-driven label filtering and structured label searching method. Bell et al. [BFH01] introduce an algorithm for application within a similar scenario. They do also refer to images and other additional information within the general view management. Objects within the scene are approximated here by rectangular shapes. These shapes are incrementally combined to point out possible empty space that can be used for the automatic arrangement of additional information. After all, these studies do either not incorporate all aspects of a callout (e.g. the connection line or a callout body) and do not explore different orientations towards the user in there alignment approach. They mostly place the text directly at a certain spot or object which is always oriented towards the user.

Besides visual annotations in AR and VR, there also exist studies that explore the influence of callout arrangement within non-immersive virtual environments. Chen et al. [CPB04] compares labels placed within world space and screen space. The latter is a similar integration of these labels as in head-up-displays (HUDs). They show that HUD-like labels better support native search tasks in information rich environments. Goldstein [Gol09] and Polys [PKB07] generally come to the same conclusion for such non-VR environments, but also point out some positive aspects relating to a world space integration, for example tight coupling of information to the objects. Maass and Döllner [MD06] describe an iterative approach for placing callouts in virtual landscapes. They take one specific point of view in consideration and show that their approach can be used to place callouts without overlapping for this specific view. Schwartges et al. [SMHW15] explore labeling methods for linear objects, such as streets in interactive 3D maps. They use billboards and present a force directed algorithm that has only little impact on the performance of their prototype system while it reduces the total number of overlapped labels, compared to a an algorithm with a fixed leader length.

The currency of related work generally indicates that the view management relating to callouts is still an unsolved challenge. Much effort is put in methods for the field of AR so that the alignment of callouts in VR provides space for potential investigations.

3. Virtual Reality Callouts

We integrated the VR callouts within a virtual knowledge demo setting used for microlearning. In this demo we aim to display several features of one object of interest (OOI) This virtual setting is described briefly in the next subsection to clarify where our callouts are applied. The three concepts of callouts that we differentiate between are then introduced thereafter. In the last subsection we will also discuss how the callouts are implemented in a prototype.

3.1. VR Knowledge Demo Setting

Our callout-based VR demo is graphically represented in the scheme in Fig. 2. The OOI that that is annotated is placed in the middle of the demo room, on a pedestal. Depending on the type of callout they are placed around and above the OOI (Fig. 1). Around the pedestal there are four teleport spots for facilitating to watch the OOI from every angle. Despite having these spots the users can also move freely. These spots are also used as points for interaction with the other media in the room. On the top there is an interactive canvas that allows to get textual additional information based on which callout is selected. On the left there is an additional image displayed relating to the OOI and the selected callout. In the back there is a menu for an interactive quiz which can used after a demo to quickly self-check on the demonstrated knowledge. In this quiz we implemented gamification elements, such as a scoreboard. On the right there is the callout menu which allows the user to change the type of callout. A users point of view can be seen in Fig. 3, which also shows the prototype implementation of the demo room.

3.2. Callout Concepts

In this work we identified and propose three different callout concepts:

- Static callout
- Rotation callout
- HUD callout

A static callout is a simple callout whose position and rotation remain static and are independent of the user’s position. This callout concept is illustrated in Fig. 4 b). An implemented example is shown in Fig. 1 left and Fig. 3. While standing in the anticipated position and looking the right direction this type of callout can be seen as expected (Fig. 4 b) left). When the user moves to another place and looking in another angle at the callouts, the text on the callouts is not oriented towards the user and may be cut off by the view frustrum (Fig. 4 b) right). He has to move back to the predefined position to view the callouts correctly.

A rotation callout basically behaves similar to the static callout when he user is at the initially designed position (Fig. 4 c) left). When the user looks from a shifted angle at the OOI (Fig. 4 c) right and Fig. 1 middle) then the callouts rotate towards the user and stay orthogonal. The text therefore is always readable when the callout falls within the view frustrum. But the callout itself can still be cut off by the view frustrum or it can be located even outside this view depending on the orientation of the user.

A HUD callout, however, assures that the body of the callout and
Figure 3: Different types of information that are displayed during the demo of an object of interest. The content changes depending on the currently selected callout. From left to right: An additional image, callouts, textual information and the 3D object representation that the callouts are attached to.

it’s text can always be fully seen, even if the user is looking away from the OOI (Fig. 1 right). From a conceptional point of view, the body of the callouts are not placed within three-dimensional space, but they are reduced on a 2D HUD-like layer. Regarding 3D space they are dynamic in their position and orientation. Therefore they are always visible and readable for the users – independently of the users position and orientation, as illustrated in Fig. 4 d). The connecting line is drawn from this HUD callout into 3D space so that the spatial anchoring of callout information is still preserved in this type of callout.

3.3. Prototype Implementation

For our prototype implementation we used the Unity3D game engine as a foundation. Generally Unity provides two options for implementing callouts: 1) Either working with 3D objects or primitives, as 3D text or textured cubes for example, or 2) using 2D UI elements, as text, that are drawn on a virtual 2D canvas-plane which is placed in the environment. The benefit for using a UI canvas is the fast prototyping of the interaction with the elements. Predefined button elements can be utilized for example. The advantage of working with separate 3D objects is that they can also be transformed separately without work-around the 2D canvas object. So if a canvas is rotated in unity then all elements on it are rotated as well. For our prototype we still decided to implement the callouts as UI elements, but chose to create a distinct canvas for each callout.

Usage of separate canvases with a total size of the callout-body also facilitates the visualization of the corresponding line. For placing these lines seamless at the borders of the callout body the most-right or most-left position of the canvas border is utilized. A vector calculation is used to check whether the annotated OOI is to the left or to the right of the callout. Derived from this the appropriate side for attaching the line to the body is chosen. For rendering the line we used an object of the Unity line renderer class to create 3D lines that connect the 2D canvas with a point in 3D relating to the OOI. This line was updated in the game loop update cycle, since the positions and orientations of the proposed callouts can change in two of the three versions.

The static callouts were implemented as stated above, each as a 2D canvas with a text component. This canvas was positioned
during the VR authoring process before the usage of the resulting application. A rotation callout was created in the same way. But during run time it was checked if the callout body faces the actual users point of view (the virtual camera simulating the eyes). To check this, a vector was calculated from the camera to each individual callout separately and it was compared to the normal vector of the canvas. If it did not face the right direction the canvas gets rotated towards this vector, until the normal of the callout is equal to the vector. The HUD callouts were also created and aligned before the usage of the application. The canvases of the HUD callouts were placed as a child of the camera, to keep them in the view of the user at any time. The lines were updated in each frame.

As for the use-case of a microlearning application for knowledge demonstration we decided to use a mobile and standalone HMD to be independent on stationary hardware. We used an Oculus Go with the compatible Oculus Go controller. For prototyping the interaction regarding the VR hardware we used the OVR camera rig and the Oculus SDK in Unity, as proposed by Oculus. Alternatively the Virtual Reality Toolkit (VRTK)† could have been used to provide multi-platform compatibility. To facilitate future authoring of the application, additional data like texts and images are loaded in real time from JSON-files. Separate threads were used to load the data during run time, so that the performance on the mobile VR hardware was not affected.

4. Evaluation

We conducted a user study for evaluating the demo concept and draw conclusions about the proposed callout versions. This study involved 16 participants (11 male; nine in the age of 18-30 years, three in the age of 30-50 years and four over 50 years). All considered themselves technically affine. There were ten that stated to have used VR technology before and six without experience. An Oculus Go standalone VR headset was used for the study together along a hand-held Oculus Go controller. The content of the demo was the construction of a hard disk drive, which was completely new to all participants.

We utilized thinking aloud methodology [ES84] to protocol statements about the usability of the prototype during the free-exploration user test. A custom online questionnaire was used to capture the participants’ opinions about the callouts and the VR demo. Besides items concerning demographic information and experience, the study was based on five main-questions (translated into German as native language of the participants): (1) Which callout version did you like best? (2) How useful did you find the callouts in the VR demo? (3) How satisfied were you with the learning context in our demo? (4) How useful did you find the gamification elements? (5) How did you find the length of the demo? Participants furthermore were asked to distribute 100 points among the three types of callouts.

4.1. Analysis of the User Study Results

The questionnaire outcome about the VR demo is shown in Fig. 5. A 5-point likert scale was used for these items. The overall usefulness of callouts within our proposed VR demo setting was rated with Ø 4.12 and SD ≈0.95. There was one outlier with two points, also shown in the descriptive statistic. The satisfaction of the learning context in our demo was rated with Ø 3.56 and SD ≈0.96 with one outlier with 5 points. The subjectively perceived usefulness of the gamification elements within our concept was rated by the participants with Ø 4.25 and SD ≈0.57 with one outlier at three points as shown in Fig. 5. Additionally, participants rated the length of the VR demo. One participant stated that it was to short, while nine stated that it was adequate for the scope of the microlearning content.

Participants stated their opinions about their callout preference in two separate questions. In the first questions eight participants chose the static callout, five chose the rotating callouts and three the HUD callout as their preferred one. In the second question there were 100 points to distribute between the different versions. They were distributed with a slider interaction in the online questionnaire. The static callout received Ø 44.81 points with a SD of ≈23.07. The rotating version received Ø 32.43 points with a SD of ≈19.23. The last callout type (HUD) received Ø 22.75 points with a SD of ≈20.99. The descriptive statistic in Fig. 6 shows a box-whisker-plots for each of these ratings.

We conducted a Kruskal-Wallis one-way analysis of variance [KW52] to make an assumption about the difference of these ratings. We chose this test to handle more than two groups with tied values and ordinal data as a non-parametric alternative for the F-test (analysis of variance).

The Kruskal-Wallis test did confirm a significant difference between the three callout ratings at a confidence level of 95%. The statistics value of ≈50.49 was higher than the critical value of 5.99 (after chi² with 2 degree of freedom and confidence of 95%). The rank-sums for the groups were 278.5, 391 and 506.5 respectively for the static, the rotation and the HUD version.

The thinking-aloud notes indicate especially, that some partic-

---

† https://vrtoolkit.readme.io/ (Accessed 04.06.2019)
4.2. Discussion of the Results

The results of the user study shown in Fig. 5 indicate a positive acceptance of our demo concept. High values for the callout and gamification usefulness furthermore indicate this strongly. The learning aspect of the demo was rated above the neutral value of 3 of the 5-point likert scale, but still indicates potential optimization. With an average value of 3.5 it is only slightly above and as well has a high SD.

As for the ranking and rating of the different types of callouts, the Kruskal-Wallis test did confirm a significant difference. The absolute values point out the static version to be the preferred one. The rotating version was rated as the second. This indicates a trend towards the two non-translating callouts and against the HUD callout. While the rating shows two high values for the static and the rotation type (44.81 and 32.43), the HUD callout has only 22.75 points in average.

Both the ranking and the first question about the preferred callout do indicate this trend. The box whiskers plot in Fig. 6 furthermore shows that the median of these two non-translating callouts lies higher than the competing one. The median of the HUD callout clearly is the lowest. Also the rank sums of the groups show a tendency towards the static and the rotating callout with 278.5 and 391 with average ranking respectively. The HUD callout has the highest value of 506.5.

The questionnaire outcome overall indicates that the HUD callout was perceived significantly worst, while the non-translating callouts were similarly ranked. This hypothesis is additionally supported by the comments during the think-aloud test, since the participants mentioned that HUD callout did hinder their intention to interact with other objects in the scene.

5. Conclusions and Future Work

In this paper we have proposed three types of callouts which differ from each other regarding their alignment towards the user - the static and dynamic positioning and rotation. We have shown a prototypical implementation of the microlearning application for demonstrating knowledge. All three callout versions were implemented. Within our user study we have indicated that there was a preference towards static positioned callouts. Walking with the controller-input at specific positions and changing view directions to read the callouts were widely accepted instead of keeping the callouts within the view.

In future work, we will investigate if the number of callouts is a potential factor for choosing certain callout types over another. It might be that few callouts are accepted in a HUD-like callout type, while ten could restrict the view of users to much. We will furthermore explore the authoring process of callouts. Since microlearning and knowledge demo experts have mostly no expertise in creating VR applications, the application of VR in these fields would be facilitated if they could create callouts themselves. We already outsourced additional content in JSON files. In future work we will strive to provide a simple application and authoring interface to create content for our demo application which are appropriate for domain experts.

Acknowledgements

The work is supported by the Federal Ministry of Education and Research of Germany in the project Innovative Hochschule (funding number: 03IHS071).

References


