

Context-Responsive Labeling in Augmented Reality

Thomas Köppel*
TU Wien, Austria

M. Eduard Gröller†
TU Wien, Austria
VRVis Research Center, Austria

Hsiang-Yun Wu‡
TU Wien, Austria

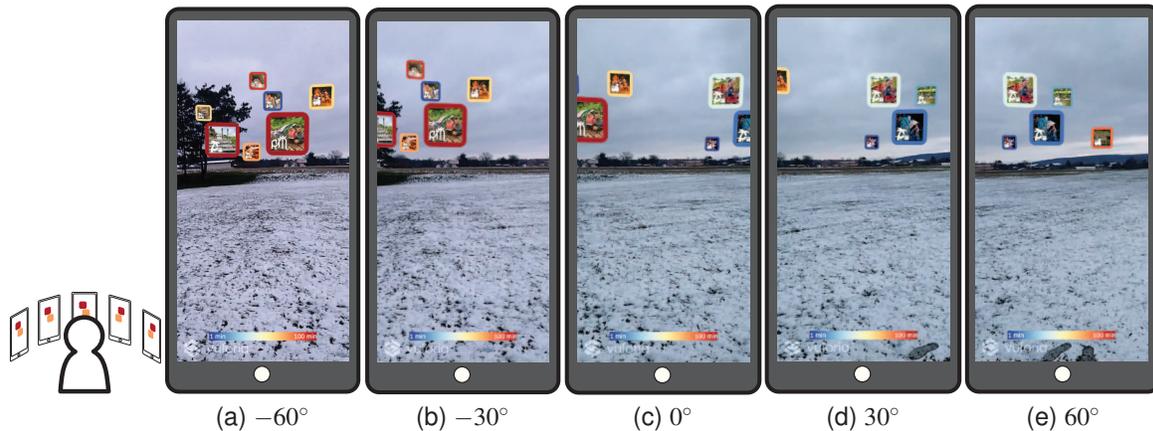


Figure 1: Responsive labeling of the *Tokyo Disneyland Dataset* in AR. The images present results of different viewing angles to show consistent label positions when the AR device is rotated toward different directions by the user.

ABSTRACT

Route planning and navigation are common tasks that often require additional information on points of interest. Augmented Reality (AR) enables mobile users to utilize text labels, in order to provide a composite view associated with additional information in a real-world environment. Nonetheless, displaying all labels for points of interest on a mobile device will lead to unwanted overlaps between information, and thus a context-responsive strategy to properly arrange labels is expected. The technique should remove overlaps, show the right level-of-detail, and maintain label coherence. This is necessary as the viewing angle in an AR system may change rapidly due to users' behaviors. Coherence plays an essential role in retaining user experience and knowledge, as well as avoiding motion sickness. In this paper, we develop an approach that systematically manages label visibility and levels-of-detail, as well as eliminates unexpected incoherent movement. We introduce three label management strategies, including (1) *occlusion management*, (2) *level-of-detail management*, and (3) *coherence management* by balancing the usage of the mobile phone screen. A greedy approach is developed for fast occlusion handling in AR. A level-of-detail scheme is adopted to arrange various types of labels. A 3D scene manipulation is then built to simultaneously suppress the incoherent behaviors induced by viewing angle changes. Finally, we present the feasibility and applicability of our approach through one synthetic and two real-world scenarios, followed by a qualitative user study.

Index Terms: Human-centered computing—Visualization—

*e-mail: tkoeppel@cg.tuwien.ac.at

†e-mail: groeller@cg.tuwien.ac.at

‡e-mail: hsiang.yun.wu@acm.org

1 INTRODUCTION

We schedule and plan routes irregularly in our everyday life. For example, we visit offices, go to restaurants, or see doctors, in order to accomplish necessary tasks. In some cases, such as visiting medical doctors or popular restaurants, one has to wait in a queue until being able to proceed. This is time-inefficient and most people try to avoid it. Normally, if a person needs to decide the next place to visit, he or she can extract knowledge about the targets of interest. Then a decision is made based on the corresponding experience or referring locations using a map. 2D maps are one of the most popular methods that describe the geospatial information of objects, to give an overview of the object positions in a certain area. With a 2D map for navigation, users need to remap or translate the objects on the map to the real environment, to understand the relationships and distances to these objects [14]. This inevitably strains our cognition. It is also the reason why some people cannot quickly locate themselves on a 2D map or find the correct direction or orientation immediately. *Augmented Reality (AR)* and *Mixed Reality (MR)* have been proposed to overlay information directly on the real-world environment with a lower complexity by instructing users in an effective way [10, 28]. In this paper, we use AR as our technique of choice for the explanation. Displaying texts or images in AR or MR allows us to acquire information encoded with geotagged data and stored in GISs. It is also known that using AR for guiding users in exploring the real environment can be more effective in comparison to a 2D representation [8].

In mixed environments, *points of interest (POIs)* are often associated with text labels [16, 20, 35] in order to present additional information (e.g., name, category, etc.). For example, an Augmented Reality Browser (ARB) facilitates us to embed and show relevant data in a real-world environment. Technically, POIs are registered at certain geographical positions via GPS coordinates. Based on the current position and the viewing angle of the device, the POIs are annotated and the corresponding labels are then projected to the screen of the user's device. Naive labeling strategies can lead to occlusion problems between objects, especially in an environment with a dense

arrangement of POIs. Additionally, properly selecting the right level of a label to present information can help to avoid overcrowded situations. Moreover, retaining the consistency between successive frames also enables us to maintain a good user experience and to avoid motion sickness. Based on the aforementioned findings, we summarize that a good AR labeling framework should address:

- (P1) **The occlusion of densely placed labels in AR space.** Occlusion removal has been considered as a primary design criterion in visualization approaches. It reflects user preferences and also allows the system to present information explicitly [45].
- (P2) **Limited Information provided by plain text.** As summarized by Langlotz et al. [20], labels in AR often contain plain text rather than other richer content, such as figures or hybrids of texts and figures.
- (P3) **Label incoherence due to the movement of mobile devices.** During the interaction with an AR system, the user may frequently change positions or viewing angles. This leads to unwanted flickering that impacts information consistency [16].

In this paper, we develop a context-responsive framework to optimize label placement in AR. By *context-responsive*, we refer to taking contextual attributes, such as GPS positions, mobile orientations, etc., into account. The system responds to the user with an appropriate positioning of labels. The approach contains three major components: (1) *occlusion management*, (2) *level-of-detail management*, and (3) *coherence management*, which are essential for the approach to be context-responsive. The *occlusion management* eliminates overlapping labels by adjusting the positions of occluded labels with a greedy approach to achieve a fast performance. Then, a levels-of-detail scheme is introduced to select the appropriate level in a hierarchy and present it based on how densely packed the labels are in the view volume of the user. We construct a 3D scene to manipulate and control the movement of labels enhancing the user experience.

We introduce a novel approach to manage label placement tailored to AR. It enables an interactive environment with continuous changes of device positions and orientations. A survey by Preim et al. [30] concluded that existing labeling techniques often resolve overlapping labels once the camera stops moving or the camera position is assumed to be fixed to begin with. Approaches often project labels to a 2D plane to determine the occlusions and then perform occlusion removal. However, object movement in 3D is not obvious in the 2D projections of a 3D scene, which leads to temporal inconsistencies that are harmful to label readability [35]. Čmolič et al. [6] summarized the difficulty of retaining label coherence due to many discontinuities of objects projected into 2D images. As in the sequence of snapshots in Figure 1, we treat labels as objects in a 3D scene and apply our management strategies for better quality control. In summary, the main technical contributions are:

- A fast label occlusion removal technique for mobile devices.
- A clutter-aware level-of-detail management.
- A 3D object arrangement that retains label coherence.
- A prototype to demonstrate the applicability of our approach [17].

The remainder of the paper is structured as follows: Section 2 presents previous work and relates our approach to existing research. An overview of our design principles and system is described in Section 3. In Section 4, we detail the methodology and technical aspects. The implementation is explained and use cases are demonstrated in Section 5, followed by an evaluation in Section 6. The limitations are explained in Section 7, and we conclude this work and provide future research directions in Section 8.

2 RELATED WORK

We present a novel responsive approach considering label occlusion, visual clutter, and coherence simultaneously. We discuss related work to identify our contributions by first covering general navigation techniques, and then specific labeling topics in different applications and spaces.

2.1 Spatial Identification and Navigation

Spatial cognition studies show how people acquire experience and knowledge to identify where they are, how to continue the journey, and visit places effectively [40]. Maps are classical tools used to detect positions and extract spatial information throughout human history [44], while modern maps often use markers to identify and highlight the locations of POIs. 2D maps may not be always optimal since the 2D information needs to be translated to the real environment [14].

An alternative, or maybe a more intuitive way, is to map the information directly to the physical environment. McMahan et al. [28] compared paper maps and Google Maps to AR or more specifically hand-held AR [33], which better supports people in terms of activating their navigation skills. Willett et al. [43] introduced embedded data representations, a taxonomy describing the challenges of showing data in physical space, and mentioned that occlusion problems have not yet been fully resolved. Bell et al. [3] proposed a pioneering view-management approach to project objects onto the screen while resolving occlusions or to arrange similar objects close to each other. Guarese and Maciel [14] investigated MR, to assist navigation tasks by overlaying the real environment with virtual holograms. Schneider et al. [32] investigated an AR navigation concept, where the system projects the content onto a vehicle's windshield to assist driving behaviors.

2.2 Labeling in Various Spaces (2D, 3D, VR, and AR)

Labeling is an automatic approach to position text or image labels in order to efficiently communicate additional information about POIs. It improves clarity and understandability of the underlying information [2]. Internal labels are overlaid onto their reference objects. External labels are placed outside the objects and are connected to them by leader lines. Recently, Čmolič et al. [6] have introduced *Mixed Labeling* that facilitates the integration of internal and external labeling in 2D. Labeling techniques have been extensively investigated in geovisualization, where resolving occlusions and leader crossings [21] are primary aesthetic criteria to ensure good readability. Besides 2D labeling, in digital map services, such as Google Maps and other GISs, scales have been considered to improve user interaction. Active range optimization, for example, uses rectangular pyramids to eliminate label-placement conflicts across different zoom levels [1, 46]. Labeling of 3D scenes has been mainly investigated in medical applications [29], usually focusing on complex mesh and volume scenes, as well as intuitiveness for navigation. Maass and Döllner [23] developed a labeling technique to dynamically attach labels to the hulls of objects in a 3D scene. Later they extended this billboard concept by taking occlusion with labels and scene elements into account [24]. The approach by Kouřil et al. [18] annotates a complex 3D scene, involving multi-instances across multiple scales in a dense 3D biological environment. Occlusion in these approaches is detected after projecting objects into 2D. It is hard to maintain coherence.

Handheld Augmented Reality has become useful as the computing power of mobile devices has increased. One advantage of using AR is to overlay information directly on the real world that the user is familiar with. For example, White and Feiner [41] proposed *SiteLens*, a situated visualization that embeds relevant data of the POIs in AR. Veas et al. [38] investigated outdoor AR applications, where they focused on multiple-view coordination and occlusion with objects in the background. Labels are not fully researched

here. As referred to in most of the following papers, occlusions between labels have been considered as a primary issue in AR applications [13, 16, 20]. Grasset et al. [13] proposed a view management technique to annotate landmarks in an image. Edge detection and image saliency are integrated to identify unimportant regions for text label placement. Jia et al. [16] investigated a similar strategy, with incorporating human placement preferences as a set of constraints to improve the work by Grasset et al. [13]. Two prototypes are implemented for desktop computers due to the poor temporal performance on mobile devices. Tatzgern et al. [35] developed a pioneering approach that considers labels as 3D objects in the scene to avoid unstable labels due to view changes. The approach estimates the center position of an object and moves labels along a 3D pole, which attaches to the object. Another proposed scenario constrains label movement to a predefined 2D view plane. This technique is limited to annotating objects in front of the camera.

Existing work tends to directly solve label occlusions in 2D or to project labels from 3D to 2D and apply 2D solutions. These techniques cannot avoid label inconsistencies [6]. In contrast to existing approaches, we handle labels as objects in the 3D scene. This allows us to compensate for incoherent label movement caused by viewing angle changes of the device. We integrate the labeling technique into 3D to retain stability and introduce additional visual variables, including text, images, icons, and colors, to enrich the corresponding visual representation. Our label encoding also varies in order to balance information provided by POIs. More design choices will be explained in Section 3.

3 CONTEXT-RESPONSIVE FRAMEWORK

Based on the taxonomy by Wiener et al. [42], our approach supports aided and unaided wayfinding tasks. We can directly highlight the destination label and assist users to combine decision-making processes, memory processes, learning processes, and planning processes for finding the overall best destinations. The effort to identify objects in AR is low [14] because real-world objects can be directly annotated [3, 13] and AR navigation is less user-focus demanding compared to other map techniques [28]. The responsive framework is inspired by Hoffswell et al. [15], who proposed a taxonomy for responsive visualization design, which is essential to present information based on the device context. In principle, our design has three major components, including (1) occlusion management, (2) level-of-detail management, and (3) coherence management, each of which aims to solve the problems (P1-P3), respectively. We first introduce the encoding of labels beyond plain text, followed by an overview of the presented approach.

3.1 Label Encoding

The label encoding reduces the limitations in existing work and solves (P2). We introduce additional types of labels than merely text labels as concluded by Langlotz et al. [20]. We use color to encode scalar variables of each POI [25, 27]. In general, the users can choose a color scheme and a scale according to their preferences. A label consists of several of the following components:

- a text tag containing the name of the POI,
- an iconic image (photo) of the POI,
- an icon encoding the type of the POI, and
- a color-coded rectangle representing a scalar value of the POI.

In Figure 2, labels concerning the *Tokyo Disneyland Dataset* are shown. POIs are attractions in this case. Attractions can be categorized into three types, i.e., *thrilling*, *adventure*, and *children*, each of which is depicted through a type icon. Figure 2(a) provides an explanatory label annotating an attraction of the dataset. The text tag depicts the name of the attraction and the waiting time (e.g.,

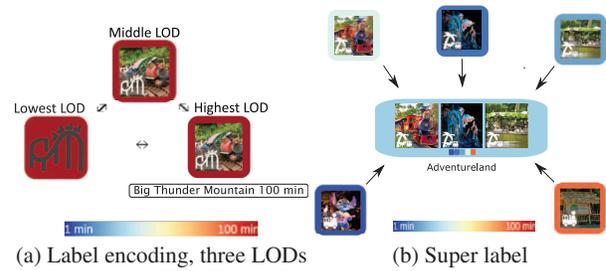


Figure 2: An example label encoding (*Tokyo Disneyland Dataset*).

Big Thunder Mountain 100 min). The iconic image shows a photo of the train of the attraction and the type icon indicates that it is a thrilling attraction. The colored (rectangular) backgrounds of the labels encode the corresponding waiting times.

3.2 Pipeline of the Approach

Figure 3 gives an overview of our approach. We first position labels of POIs in AR (Figure 3(a) as a top view and (b) as a front view) and perform the proposed three management strategies. We process the objects in the 3D scene using a Cartesian world coordinate system, where the xz -plane is parallel to the ground plane. Figure 3(a) depicts a top view of our coordinate system, the x -axis and z -axis define the ground plane and the y -axis is vertically upwards from the ground plane. The input to our system is a set of POIs $P = \{p_1, p_2, \dots, p_n\}$ and a set of labels $L = \{l_1, l_2, \dots, l_n\}$, for example, manually selected by the users or downloaded from an online database. In the *positioning labels in AR* preprocessing (Section 4.1), for each POI p_i , the corresponding label l_i is initially placed perpendicularly to the ground plane in the world coordinate system (Figure 3(b)). Currently, each POI p_i has one associated label l_i describing the attributes of the POI. We also assume that the (x, z) -coordinates of each annotated POI are more important than the y -coordinate, since the (x, z) -coordinates are essential to indicate the relative positions of the POIs as suggested by prior work [3, 14].

The *occlusion management* strategy (Section 4.2) addresses (P1) and resolves occlusions of labels considering the current configuration of the device. The labels are first sorted by distance from the device into a list S , from the nearest to the farthest positions. With this information, we resolve occlusions starting with the closest label and using a greedy approach (see Figure 3(c)). The greedy approach arranges the lowest y -positions of the labels to be visible iteratively. This allows effective execution of the occlusion-handling on mobile devices, where the computation powers are limited compared to desktop computers. The occlusion strategy provides a solution to otherwise inconsistently moving labels when the viewing angle of the AR device changes [35].

In the *level-of-detail management* (Section 4.3), we introduce four distinct types of label encodings for (P2) to represent three levels-of-detail (LODs, Figure 2(a)) of an individual label and one *super label* to indicate an aggregated group of labels for visual clutter reduction (Figure 2(b)). The *level-of-detail management* depicts a different amount of information for each label (see Figure 3(d)). The LOD of a label l_i is selected according to the distance of the annotated POI to the device and the label density in the view volume. For convenience, we assume that close labels get at least as much screen space as distant labels, since it is natural to show objects larger when they are close by. However, different configurations can be also incorporated by adding rays in the occlusion detection. Super labels (Figure 2(b)) are representative labels that depict a set of aggregated labels in order to reduce visual clutter. Figure 2(b) gives an example of a super label for the *Tokyo Disneyland Dataset*. The themed area *Adventureland* is aggregated and the blue background color of the

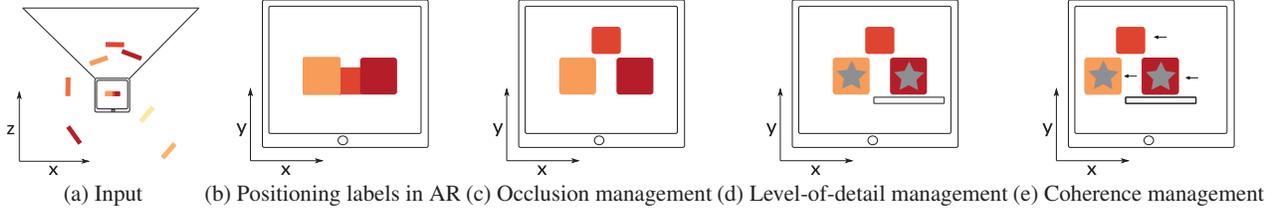


Figure 3: The input scenario (a), positioning of labels in AR (b), and the three management strategies of our approach (c)-(e).

super label encodes the average waiting time. A color legend at the bottom of the label presents the individual waiting times of the aggregated attractions in this themed area.

Positioning labels in AR, occlusion management, and level-of-detail management are smoothly updated in the *coherence management* module (Figure 3(e)). To avoid flickering that inevitably reduces coherency [16], the labels are not moved or changed immediately, but follow a common animation policy, by strategically updating changes over time (Section 4.4) to solve problem (P3).

4 CONTEXT-RESPONSIVE LABELING MANAGEMENT

Our approach positions labels in AR space, followed by a context-responsive computation. Here we introduce occlusion removal, perform level-of-detail strategies, and enforce coherent label placement. In this section, we will detail the proposed technique.

4.1 Positioning Labels in AR

In a preprocessing step, we map the geographical locations from the real world to our Cartesian AR world space. This considers the GPS position of the user's device, the GPS location of the POIs, and the compass orientation of the device [39].

The labels are oriented towards the user's position by aligning the normal vectors of the labels with the AR device in the AR world space. Once this initial label positioning is done, a perspective projection from the AR world space into the screen space of the device is performed. In doing so, we can position the labels in AR spatially relative to the position of the user to support exploration and navigation as shown by Guarese et al. [14]. In principle, existing frameworks, like the *AR + GPS Location SDK* package [11] or the *Wikitude AR SDK* package [12] can be used to map real-world objects to the AR world space. Unfortunately in our experiment, the techniques are not stable due to the inaccurate GPS sensor [37] or compass [19] data of mobile devices. To test and assess the quality of the coherence strategies for the *occlusion management* and *level-of-detail management*, we predefine the positions of labels at the (x, z) -coordinates in the AR world space. The existing libraries do not provide stable label positions, which would lead to a less coherent behavior that is not relying on the proposed *coherence management*. Once the labels are placed, we order the labels based on their distance to the user for future computations.

4.2 Occlusion Management

Showing many labels simultaneously on a mobile device will, unfortunately, lead to occlusions of labels, especially if the annotated POIs are close to each other or even hidden by other labels (Figure 4(a)). Point-feature labeling has been extensively investigated due to its NP-hardness, even when looking for an optimal solution just in 2D [5]. In our setting, occlusions change over time, since the users move. Fast responsive management strategies are required to update the scene regularly. Viewing angle and position changes of the user need to be accounted for to guarantee smooth state transitions and to eliminate unwanted flickering. We perform the entire occlusion handling in the 3D scene, overcoming the label positioning inconsis-

tencies caused by viewing angle changes. The occlusion handling consists of two steps, occlusion detection and shift computation.

4.2.1 Occlusion Detection

We employ ray tracing to detect occlusions, which is different from existing approaches [2]. As the labels have been sorted by the distance to the user, the occlusions are detected and solved iteratively from label l_1 to label l_n of the sorted list S . For each label l_i , the origins of four rays are set to the location of the user's device in AR. The rays run through the corner points of label l_i as shown in Figure 4(b). If another label is hit during the ray traversals, an occlusion occurs. To ensure that all possible occlusions will be detected, we assume that labels closer to the viewer are either larger or as large as labels farther away. This allows us to use just four rays to detect 3D occlusions effectively. The approach works for rectangular shapes or rectangular bounding boxes of polygonal shapes and could be extended to polygons or 3D objects (e.g., buildings in MR). Other configurations can be accommodated by increasing the number of rays. Figure 4(b) gives an example, where label l_1 (orange) is in front of label l_2 (red). In this case, the corner ray 1 of label l_2 collides with label l_1 , indicating that label l_1 occludes label l_2 . Since we assume that closer labels are always larger or as large as farther away labels, no occluding labels will be missed during the occlusion detection.

4.2.2 Shift Computation

Once the occlusions are detected, we can iteratively shift the labels greedily in the order of increasing distance. Since the labels are shifted from the closest to the farthest one, the label l_i will be located either at its initial (x, z) -coordinates or above the previous label l_{i-1} along the y -axis. Figure 4(c) illustrates the basic shift of label l_2 . The blue lines represent the corner rays for occlusion detection and the gray line shows the traversed ray for calculating the occlusion free position of label l_2 . Figure 4(d) depicts an occlusion-free result after shifting label l_2 , where the shift distance d is $|y'_2 - y_2|$.

Szirmay-Kalos et al. [34] proved that the ray-tracing approach at least requires a logarithmic computation time in the worst case based on the number of scene objects. On the other hand, modern platforms already provide real-time ray-tracing [36]. In our approach, the *occlusion management* takes $O(n^2)$ if labels are aligned in a sequence along the current viewing direction. The current label l_i possibly needs to be shifted above each label in front of it. We show a comparison with different label alignments in Section 5. The greedy label placement terminates as soon as no other label in front occludes label l_i .

4.3 Level-Of-Detail Management

Labels occupy space that is a scarce resource on a mobile device, especially if many labels should be shown simultaneously. To reduce unwanted visual clutter, we introduce an LOD concept for labels [26] and incorporate a *level-of-detail management* in the pipeline (Figure 3(d)). The LOD is also computed based on the sorted distances of labels and the label density. The LOD selection consists of two steps: LOD calculation and super label aggregation.

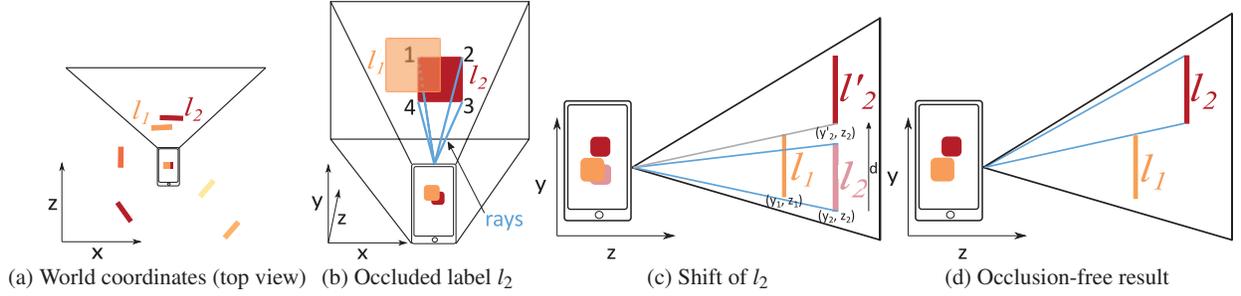


Figure 4: Illustration of the *occlusion management*. (a) The labeled 3D scenario in top view. Label l_1 and label l_2 are in the current view volume. (b) Occlusion by l_1 (transparent and orange), which is in front of l_2 (red). The ray at corner 1 of l_2 intersects the occluding, label l_1 . (c) Label l_2 is shifted above the gray ray by the distance d to resolve the occlusion. (d) The blue corner rays do not collide with a label in front anymore.

In our implementation, the lowest LOD occupies the least space and includes a colored rectangle and an icon (Figure 2(a)). The middle LOD presents a colored rectangle, the icon, and an iconic image (photo) of the POI (Figure 2(a)). The highest LOD contains a text tag and occupies the most space (Figure 2(a)). The level-of-detail for each label changes when the user navigates through the scene.

4.3.1 LOD Calculation

The LOD for each label depends on the distance to the user and the label density. For each label, a virtual view volume aligned to the (x, z) ground plane is constructed to mimic that the user would look into the direction of each label. The horizontal distance along the (x, z) ground plane and the vector from the user to the position of each label are used. If the angle between these two vectors is above a threshold t (45° by default in our system), the label is located outside the aligned view volume. We split the view volume and each label below the threshold m_1 (20° by default) receives the highest LOD until one label exceeds the angle m_1 . The remaining labels are displayed in the middle LOD until reaching the threshold m_2 (30° by default). If a label exceeds m_2 , it will be displayed in the lowest LOD. The threshold angles can be changed according to user preferences. The LODs of all labels are consistent when the viewing angle of the device changes for the current user position. The *level-of-detail management* provides coherent movement when rotating the AR device. The LODs for the labels are updated if the user moves.

4.3.2 Super Label Aggregation

To further reduce visual clutter, we introduce super labels that group individual labels (see Section 3.1). The position of a super label is calculated as the average (x, z) -positions of the individual labels that are part of the aggregation in the 3D scene. A predefined grouping (i.e., themed areas of amusement parks) of labels is necessary to compute the super labels, while unsupervised clustering algorithms can also be directly applied. We do not aggregate labels of the closest predefined group considering the position of the user. Individual labels in the close surroundings of the user are always displayed and not aggregated supporting the exploration process. We only aggregate individual labels to super labels if the user is located outside of the respective label group.

4.4 Coherence Management

To avoid unwanted flickering, we incorporate smooth transitions for each movement and change. Smooth transitions are implemented if positions of labels change to be occlusion-free during the interaction with the system, if LODs of labels change, or if labels are aggregated to super labels. We investigated ten different easing functions, including linear, and various quadratic and cubic equations, for the

transitions to further increase the coherency. For comparison, we refer readers to the supplementary videos. We believe that the ease-in ease-out sine function (Eq. 1) represents the best easing function as it provides harmonic transitions. The easing function can be changed based on user preferences. Let $t_{transition}$ be the duration for a transition to be completed. The variables t_{start} and $t_{current}$ indicate the start time and the current time during the transition. The function $e(t_{current})$ represents the easing function for a smooth transition:

$$e(t_{current}) = -0.5 * (\cos(\pi * \frac{t_{current} - t_{start}}{t_{transition}}) - 1). \quad (1)$$

4.4.1 Smooth Occlusion Transitions

Due to the interaction of the user, occlusion-free label positions may vary from one frame to the next. If the labels would simply be displayed at the newly calculated positions, the labels might abruptly change their positions, which destroys the users' experience since the labels do not move in a coherent way. To allow the user to better keep track of the labels, we implemented smooth transitions from the previous locations of the labels to the newly calculated ones. We interpolate original positions and the newly calculated positions of the labels. The position for label l_i is updated every frame until it reaches its destination. Let $p_{goal}(l_i)$ be the new occlusion-free label position and $p_{start}(l_i)$ the label position at the start of the transition. We calculate the current label position for label l_i :

$$\vec{p}(l_i) = \vec{p}_{start}(l_i) + (\vec{p}_{goal}(l_i) - \vec{p}_{start}(l_i)) * e(t_{current}). \quad (2)$$

4.4.2 Smooth LOD Transitions

If the LOD for a label changes, the transition needs to be smoothed to avoid flickering and allow a coherent user experience. The LODs of labels change over time, and we adapt the alpha channel to achieve a smooth transition. In this way, the iconic images, the icons, and the text tags fade in or out using

$$\alpha(l_i) = \begin{cases} e(t_{current}), & b = 1 \\ 1 - e(t_{current}), & b = 0, \end{cases} \quad (3)$$

where $\alpha(l_i)$ is the alpha value of the iconic image, the icon, or the text tag of label l_i . Since our easing function e (in Eq.(1)) returns a value between 0 and 1, the result can be used to set the alpha channel in Eq.(3). The variable b indicates, if the object should become invisible ($b = 0$) or if the object should become visible ($b = 1$).

4.4.3 Smooth Aggregation Transitions

If labels are aggregated to super labels, individual labels will be moved to the respective super label positions in the scene. Simultaneously, we fade in the super labels and fade out the labels by

interpolating the alpha channels. If individual labels are aggregated, the labels move towards their super label and disappear. If an aggregation is split up again, coherency is achieved analogously. If the alpha channel of a super label is decreased, the individual labels reappear over time and move back to their respective positions (Eqs.(4), (5), and (6)). Let l_i be a label that will be aggregated into a super label l_s . The alpha values of l_i and l_s and the position of l_i are computed as follows:

$$\alpha(l_i) = 1 - e(t_{current}) \quad (4)$$

$$\alpha(l_s) = e(t_{current}) \quad (5)$$

$$\vec{p}(l_i) = \vec{p}_{start}(l_i) + (\vec{p}(l_s) - \vec{p}_{start}(l_i)) * e(t_{current}) \quad (6)$$

5 EXPERIMENTAL RESULTS

To assess the applicability of our technique, we investigate three different use cases, including a (1) *Synthetic Dataset*, a (2) *Local Shops Dataset*, and the (3) *Tokyo Disneyland Dataset*. The *Synthetic Dataset* shows different variations of label layouts. The *Local Shops Dataset* provides a real-world example, where the labels are close and next to each other. The *Tokyo Disneyland Dataset* presents another real-world scenario, where the labels are spread out in the 3D scene. We use Unity as the visualization platform [36] and incorporate the Vuforia engine [31] to arrange objects in AR. The images shown in this section were taken using a Xiaomi Mi A2 device (Qualcomm Snapdragon 660 processor and 4 GB RAM) with Android 10 in portrait mode.

5.1 Synthetic Dataset

We study three different label layouts of the *Synthetic Dataset* (Figure 5) and compute the execution time measured on the mobile device Xiaomi Mi A2. The three layouts are a circle layout, a grid layout, and a line layout, which are computationally increasingly expensive. This assumption is based on the fact that if more labels are hidden in the current viewing direction, more occlusion removal steps are necessary. Figure 6 gives the execution times of all layouts in milliseconds based on a variation of label numbers. The labels in this dataset have a height and width of 120 world space units by default in Unity.

The circle layout (Figure 5(a)) requires the least computation times to resolve occlusions since many labels are initially arranged without occlusion issues. The radius of the circle layout is set to 1,000 world space units in this experiment. The grid layout (Figure 5(b)) distributes the labels equally leading to densely placed labels in the scene. In our setting, the number of labels per row is equal to \sqrt{n} , where n is the total number of labels in Figure 6. If \sqrt{n} is not an integer, the layout contains one partial label row in the grid. The size of the grid is $4,000 \times 4,000$ world space units and includes both near and far labels in the world space. The line layout (Figure 5(c)) represents the worst case example. The labels are located one after another, which leads to the maximum number of $i - 1$ shifts for each label l_i . The labels are placed 90 world space units behind each other. As shown in Figure 6, resolving occlusions for the grid layout leads to higher computation times than the circle layout, but lower computation times compared to the line layout.

5.2 Local Shops Dataset

The *Local Shops Dataset* contains shop locations, types of shops, and number of people inside a shop (per m^2) of a strip mall (Figure 7). The icons indicate the respective shop types (e.g., clothing, shoes, and groceries). Considering the current COVID-19 regulations, we encode the number of people per m^2 , to identify the customer density

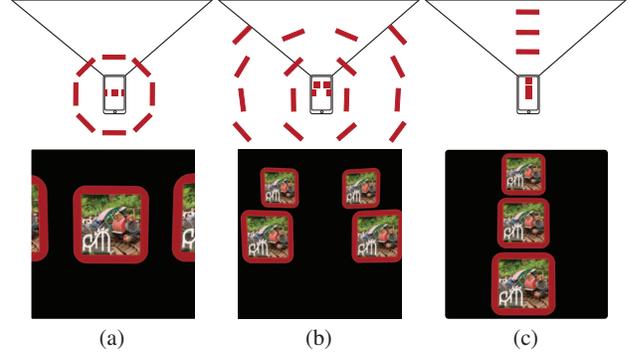


Figure 5: An example of the *Synthetic Dataset* in top view with the displayed results beneath. Labels are arranged on a (a) circle, (b) grid, and (c) line.

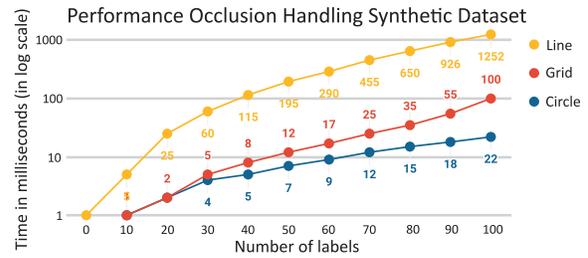


Figure 6: Computation times for removing occlusions.

or COVID-19 safety measure in the shop in real-time. In Figure 7, we use a color scale from white to red. The text displays the name and measure accordingly. Figure 7 gives an explanatory result, in which the device is tilted. As shown here, the placement of the labels is thereby not influenced. The rectangular labels remain parallel to the ground.

5.3 Tokyo Disneyland Dataset

The *Tokyo Disneyland* is one of the most popular amusement parks in the world. Many visitors often need to line up for hours to enjoy a specific attraction, and many magazines and blogs guide visitors to optimize their one-day visit [4]. The amusement park consists of 35 big attractions, which we mark all as POIs in our system to give an overview of the park. In the park, themed areas, such as the *Westernland*, are subregions grouping several attractions for convenience. We use the themed areas of the amusement park to aggregate labels and present the area using the corresponding super label.

Once the *positioning labels in AR* has been preprocessed, labels might initially be occluded. Figure 8 compares the results of the same position and viewing angle. Initially, the labels are occluded as shown in Figure 8(a) and the respective occlusion-free result is given in Figure 8(b). Since the occlusion-free results are independent of the viewing angle of the device, no incoherent label movement occurs when the user rotates the device. The occlusions are resolved for all the labels around the users as explained in Section 3 and Section 4.2. Labels closer to the user are more likely to stay close to their initial positions than labels that are farther away. The two closest labels in Figure 8 are *Big Thunder Mountain* and *Mark Twain's Riverboat* showing an iconic image of a train and a boat. The positions of these two labels are not changed. Labels that are occluded by these two labels will be shifted upwards during the *occlusion management*. Figure 9 depicts the transition of a super label to its individual labels. The super label represents the *Westernland* themed area of the *Toko*

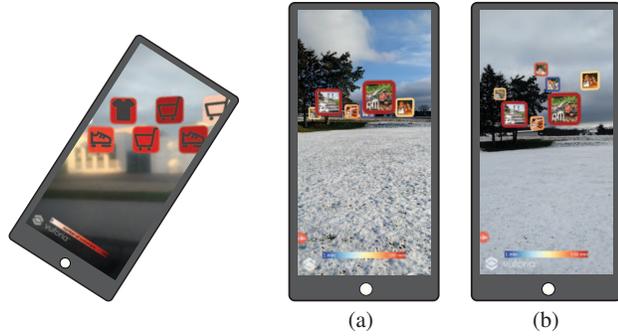


Figure 7: An example with a 45° tilted mobile device. Figure 8: Occlusions that occur in (a) are resolved in (b).

Disneyland.

Figure 10 presents different LODs of the themed area *Westernland*. Figure 10(a) shows all labels in the lowest LOD consisting of a colored rectangle encoding the waiting time and an icon indicating the attraction type. This LOD provides the simplest overview of the attractions, and it presents the least amount of information as only the attraction types and the color encodings are included. Figure 10(b) illustrates the middle LOD adding an iconic image to the encoding. In this case, the type icon is less dominant than in the lowest LOD. Figure 10(c) depicts the highest LOD by adding a text tag stating the name and the exact waiting time of the attraction in minutes. This LOD provides the most detailed information. However, higher vertical stacking of labels is necessary to resolve occlusions compared to the lowest and middle LOD during the occlusion handling. Figure 10(d) presents the label placement of the themed area *Westernland* once the dynamic LOD selection is enabled. This solution constitutes a compromise concerning the presented amount of information and label displacement. It includes detailed information about close attractions and keeps the vertical stacking of labels low compared to the highest LOD. The preferred LOD might vary depending on the use case and the user’s preference (see Section 6). Each LOD has its benefits and drawbacks with the dynamic LODs being the most versatile one as they present detailed information about close labels and avoid excessive vertical stacking (see Section 6). Figure 11 exemplifies lateral translations of the user and the resulting label arrangements. Figure 11(a) and Figure 11(c) correspond to the initial positions. In Figure 11(b) and Figure 11(d), the user moved laterally to the right. The label positions are updated smoothly depending on the movement of the user.

6 QUALITATIVE EVALUATION

We conducted an online survey to evaluate the effectiveness and the applicability of our approach. Primarily, we aim to confirm the appropriateness of the selected design principles. It is based on users’ preferences by examining task performance in terms of required time and result accuracy. Our hypotheses of the study are summarized as follows:

- (H1) The design principle, removing label occlusions, has higher priority in comparison to showing precise positions of labels.
- (H2) Rich label design in AR leads to a better POI exploration and decision-making experience in contrast to plain text labels.
- (H3) Users can perform faster route planning tasks using our system compared to conventional maps.

We further decompose our hypotheses into four main tasks as summarized in Table 1 for an online questionnaire. In the future, we

plan to do an in-person user study as one of our primary attempts. For each measurable task, time and accuracy were collected. After each task, we also asked participants to provide reasons regarding their experience when performing the task. At the end of the entire questionnaire, we requested general feedback and collected some personal information for further analysis (e.g., age, educational background, experience with AR devices, and so forth). Privacy agreements have been received prior to the user study and the collected data is carefully stored without identifications of the participants. In total, we recruited 30 participants who are experienced with visualization techniques and graduate students of visual computing participated in the survey. The age of the participants ranges from 24 to 64 years with the majority of participants being in the late twenties or the early thirties. One limitation of the user study comes from the limited access to the general audience, while experience in visual computing will help the participants to answer the questions smoothly. We performed a within-subjects study design, where we tested all variable conditions for a participant in order to analyze individual behaviors in more depth. Questions in each task are also randomized to avoid a learning effect. For more details, we refer to the accompanying supplementary materials.

Tasks	Goal of the investigation and question samples
Task 1	Impact of occlusion on attribute tasks and comparative tasks Q1: What is the waiting time of an attraction? Q2: Which attraction has the minimal waiting time?
Task 2	Effectiveness of levels-of-detail Q3: Which themed area has the minimal waiting time? (with LOD variations) Q4: Which LOD do you prefer?
Task 3	Effectiveness of 2D maps and our AR encoding Q5: Choose the attraction with the minimal waiting time in the specified themed area
Task 4	Combinatorial features in our system Q6: Provide your feedback to different configuration settings

Table 1: Overview of the tasks in the user study.

(H1) demonstrates the importance of resolving label occlusions in AR. As described in Section 2, existing work concludes the importance of resolving occlusions in AR to support the decision making process by the users [13]. In Task 1, we show participants a few snapshots (see supplementary materials) of our system, and ask the participants to determine the waiting time of the specified attraction (Q1) and select the attractions with minimal waiting times (Q2). Three participants managed to select the correct waiting times if occlusions occurred, and the participants stated that the waiting times were not recognizable in such a situation. Figure 12 summarizes task completion time and accuracy. The time needed to answer the questions could be decreased (Q1 from 33.26 s to 12.6 s, Q2 from 21.39 s to 13.7 s) and the number of correct answers could be increased (Q1 from 10 % to 86.67 %, Q2 from 3.33 % to 100 %) when showing results with our *occlusion management* (Figure 12). 24 participants explicitly stated that it was difficult or impossible to select the correct answers if information is occluded, and 24 participants agree that the occlusion-free positioning eases decision-making processes when investigating the labels.

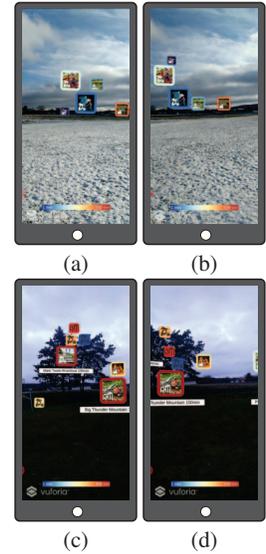
For hypothesis (H2), we design questions in Task 2, where participants need to take several attributes into account to answer the questions. In Q3, the participants were asked to select a themed area of the amusement park with the lowest average waiting time. We showed participants images with labels of different LOD settings, including text labels, labels with the lowest LOD, and super labels. The time needed to answer questions for this task is summarized in Figure 12(a). The participants, in general, spent more time if only text labels are present (52.05 s on average) since they probably like to calculate the correct number to answer the questions



Figure 9: Transition from a super label to the individual labels for each POI over time.



(a) Lowest LOD (b) Middle LOD (c) Highest LOD (d) Dynamic LODs



(c) (d)

properly. If we present information using the lowest LOD, a shorter time (23.52 s on average) was required in comparison to pure text labels. Using super labels achieved a similar performance, participants spent 21.61 s to answer the questions. If the waiting time is depicted using text labels or labels in the lowest LOD, the themed area with the minimal average waiting times was correctly selected by 73.33 % of the participants. 90 % of the participants selected the correct answers if the super labels were shown (Figure 12(b)).

In Q4, we ask participants which LOD they prefer. We presented text labels, labels in one of the three LODs, and labels in dynamic LODs as computed by our *level-of-detail management*. The dynamic LODs were chosen as the favorite approach by 40 % of the participants. 53.33 % of the participants preferred the highest LOD. Participants, who selected the dynamic LODs as their favorite design, emphasized that the vertical stacking of labels is reduced while detailed information about close attractions is preserved. The participants who chose the highest LOD as their favorite design appreciated the detailed information that can be used in decision making. It is surprising that they were not disturbed or annoyed by the excessive vertical stacking of the labels. The dynamic LODs avoid this excessive vertical stacking while presenting more information about close labels and less information about far labels. To check vertical stacking, Figure 13(a) compares the highest LOD and dynamic LODs. The more information is included for a label, the higher is the chance the label needs to be shifted upwards and stacked. We recorded the y-coordinate from the highest label of the two methods as a representative value for each themed area. The height of the stacked labels can be effectively reduced when using the dynamic LODs.

For hypothesis (H3), we aim to compare the decision making effectiveness when using 2D paper maps or our AR encoding in Task 3. We again measure the task completion time and accuracy between using a Tokyo Disneyland map and our visualization. As a preprocessing, we first removed other POIs (e.g., shops or restaurants) and left the 35 big attractions from the official 2D map of the amusement park, to increase the fairness of the comparison. More details about the task are included in the supplementary material. 60 % of the participants selected attractions with minimal waiting times of a themed area when using the 2D map and 83.33 % when the AR encoding was employed (Figure 13(b)). The average time that the participants needed to select an attraction using the 2D map was 58.79 s while they spent 32.18 s on average when using our approach, which clearly shows a reduced effort for POI selection (Figure 13(c)).

In the feedback session, participants are allowed to freely comment on the presented approach. Videos are shown highlighting the dynamic behavior of our tool when the user interacts with the system. Two participants mentioned that they prefer 2D maps compared to AR since 2D maps give a global top view. However, they performed the tasks in the user study better with the AR setting. We believe that both 2D maps and AR systems have strengths and weaknesses depending on the tasks and use cases. In our study, we have proven that for navigation purposes, AR systems could be more practical. Two participants also suggested to combine 2D maps together with AR systems as done by Veas et al. [38]. This could allow us to exploit the advantages of both approaches and achieve a similar result as in *Google Maps* and *Google Street View*. Other participants would prefer super labels combined with the highest LOD. This could reduce

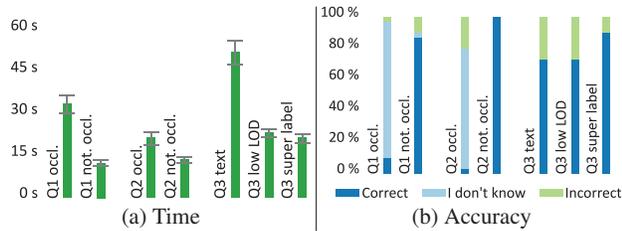


Figure 12: (a) Task completion times (in seconds) and (b) accuracy of Q1 to Q3. The error bars represent the standard errors.

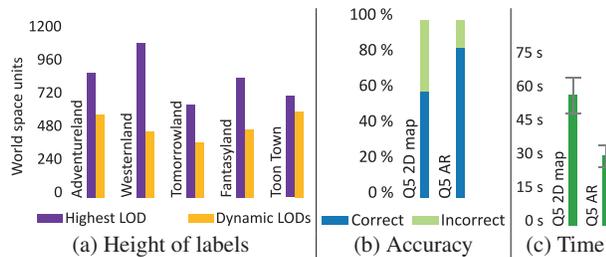


Figure 13: (a) Combined height of the stacked labels. (b) Accuracy and (c) task completion times (in seconds) of Q5. The error bars show the standard errors.

visual clutter, but might lead to a higher vertical stacking of labels compared to dynamic LODs. We, therefore, allow users to adjust the thresholds for switching LODs, to accommodate this preference. The occlusion handling and the smooth transitions were positively mentioned by participants in the general feedback. Examples include: "I really like the occlusion management, to my eyes, it's almost seamless." or, "Active occlusion handling is much superior to no occlusion handling.". Super label aggregation has been another popular and specifically mentioned feature. Participants appreciate the overview on the themed areas by giving feedback such as, "I like the super label transitions if there are many attractions because it gives a good overview of an area.", and "I like the super label transitions the most.". Overall, all participants expressed interest to use our system for navigation purposes.

7 LIMITATIONS

The limitations of our system are inherited from the hardware, especially the accuracy of mobile GPS. The position and particularly the rotation data from the available Xiaomi Mi A2 smartphone and the Google Nexus C tablet are not consistent based on our experience. A less coherent behavior of our system follows as the sensor data from each of the two devices is not stable. This, unfortunately, limits the capability to fully utilize the application, while we also envision that this will sooner or later be solved by newer technologies. To remove the errors, we thus present the results using predefined label positions in AR 3D world space. This allows us to avoid those errors induced by the hardware (e.g., changes in the device position and viewing angle) that could influence the coherence of labels. It will be interesting to collaborate with researchers focusing on high-precision GPS positioning systems.

Another limitation is that the data could contain many POIs with long text descriptions. If each label should be large enough to show the text, not much background information could be depicted eventually. The current aggregation of labels to super labels is straightforward and can be easily extended based on the use cases. One important decision criterion for the *occlusion management* and the *level-of-detail management* is the position of the user. The ordering of the labels based on the position of the user influences

the resulting labeling. Furthermore, one limitation is the loss of the global overview using AR compared to 2D maps as mentioned by related work [3, 13, 14] and two user study participants. Users need to interact with the system and look into different directions to see all the labels. The AR view only depicts the labels that are currently in front of the user in the respective view volume. We could in the future introduce additional labels on the sides of the screen to provide hints to invisible objects.

8 CONCLUSION AND FUTURE WORK

We present a context-responsive labeling framework in Augmented Reality, which allows us to introduce rich-content labels associated with POIs. The label management strategy suppresses label occlusions and incoherent label movements caused by transitions and rotations of the device during user interaction. The framework presents an alternative approach for spatial data navigation. The *level-of-detail management* takes the position of the user and label density in the view volume into account. The computed levels-of-detail for each label avoid excessive vertical stacking of labels, while still retaining basic information, which depends on the object distance. To further reduce visual clutter, we introduce the concept of super labels, which group a set of labels. Smooth transitions have been implemented in our *coherence management* to avoid flickering and enable stable label movement. The evaluation shows the applicability of the proposed approach.

As future direction, techniques will be investigated to overcome the drawbacks of seeing only the labels that are in the current view volume. The user should still anticipate POIs outside the view volume and retain a global overview of the annotated scene as with 2D maps. One possibility would be including the technique presented by Lin et al. [22] to depict labels that are currently outside the view volume and place hints at the display border of the device. Considering the positioning accuracy, it would be interesting to include so-called *Dual-Frequency GPS* [9] or *Continuous Operating Reference Stations* (CORS) [7] as investigated by related work to improve the sensor accuracy of mobile devices [19]. A selection scheme with the integration of service providers (e.g., OpenStreetMap or Google Maps with large POI data) could improve the system usability.

ACKNOWLEDGMENTS

Part of the research was enabled by VRVis funded in COMET (879730) a program managed by FFG.

REFERENCES

- [1] K. Been, M. Nöllenburg, S.-H. Poon, and A. Wolff. Optimizing active ranges for consistent dynamic map labeling. *Computational Geometry*, 43(3):312–328, 2010.
- [2] M. A. Bekos, B. Niedermann, and M. Nöllenburg. External labeling techniques: A taxonomy and survey. *Computer Graphics Forum*, 38(3):833–860, 2019.
- [3] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *Proceedings of the 14th annual ACM symposium on user interface software and technology*, pp. 101–110, 2001.
- [4] T. Bricker. Tokyo Disneyland planning guide. Disney Tourist Blog, 2020.
- [5] J. Christensen, J. Marks, and S. Shieber. Placing text labels on maps and diagrams. *Graphic Gems*, 4:497–504, 1994.
- [6] L. Cmolik, V. Pavlovec, H.-Y. Wu, and M. Nöllenburg. Mixed labeling: Integrating internal and external labels. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–14, 2020.
- [7] P. Dabove and V. Di Pietra. Towards high accuracy gnss real-time positioning with smartphones. *Advances in Space Research*, 63(1):94–102, 2019.
- [8] A. Devaux, C. Hoarau, M. Brédif, and S. Christophe. 3D urban geo-visualization: in situ augmented and mixed reality experiments. In *ISPRS Technical Commission IV Symposium*, vol. IV-4, pp. 41–48, 2018.

- [9] A. Elmezayen and A. El-Rabbany. Precise point positioning using world's first dual-frequency gps/galileo smartphone. *Sensors*, 19(11):2593, 2019.
- [10] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsomboon, and M. Billingham. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81–98, 2019.
- [11] D. Fortes. AR + GPS Location. <https://assetstore.unity.com/packages/tools/integration/ar-gps-location-134882>, 2020. [Online; accessed 09-June-2020].
- [12] W. GmbH. Wikitude Cross Platform Augmented Reality SDK. <https://www.wikitude.com/products/wikitude-sdk/>, 2020. [Online; accessed 25-October-2020].
- [13] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Image-driven view management for augmented reality browsers. In *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 177–186, 2012.
- [14] R. L. M. Guarese and A. Maciel. Development and usability analysis of a mixed reality gps navigation application for the Microsoft Hololens. In *Proceedings of the Computer Graphics International Conference*, pp. 431–437, 2019.
- [15] J. Hoffswell, W. Li, and Z. Liu. Techniques for flexible responsive visualization design. pp. 1–13, 04 2020.
- [16] J. Jia, Y. Zhang, X. Wu, and W. Guo. Image-based label placement for augmented reality browsers. In *Proceedings of the 4th International Conference on Computer and Communications (ICCC)*, pp. 1654–1659, 2018.
- [17] T. Köppel. git repository. <https://github.com/1327052/ARContextLabeling.git>. [Online; accessed 10-February-2021].
- [18] D. Kouřil, L. Čmolík, B. Kozlikova, H.-Y. Wu, G. Johnson, D. Goodsell, A. Olson, M. E. Gröller, and I. Viola. Labels on levels: Labeling of multi-scale multi-instance and crowded 3D biological environments. *IEEE Transactions on Visualization and Computer Graphics*, 25:977–986, 2019.
- [19] T. Kuhlmann, P. Garaizar, and U.-D. Reips. Smartphone sensor accuracy varies from device to device in mobile research: The case of spatial orientation. *Behavior Research Methods*, 2020.
- [20] T. Langlotz, T. Nguyen, D. Schmalstieg, and R. Grasset. Next-generation augmented reality browsers: rich, seamless, and adaptive. *Proceedings of the IEEE*, 102(2):155–169, 2014.
- [21] C. Lin. Crossing-free many-to-one boundary labeling with hyperleaders. In *2010 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 185–192, 2010.
- [22] Y.-T. Lin, Y.-C. Liao, S.-Y. Teng, Y.-J. Chung, L. Chan, and B.-Y. Chen. Outside-in: Visualizing out-of-sight regions-of-interest in a 360 video using spatial picture-in-picture previews. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 255–265, 2017.
- [23] S. Maass and J. Döllner. Dynamic annotation of interactive environments using object-integrated billboards. In *Proceedings of the 14th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, pp. 327–334, 2006.
- [24] S. Maass and J. Döllner. Seamless integration of labels into interactive virtual 3D environments using parameterized hulls. In *Computational Aesthetics in Graphics, Visualization, and Imaging*, 2008.
- [25] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, 5:110–141, 04 1986.
- [26] K. Matkovic, H. Hauser, R. Sainitzer, and M. E. Gröller. Process visualization with levels of detail. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 67–70, 2002.
- [27] R. Mazza. *Introduction to information visualization*. Springer Science & Business Media, 2009.
- [28] D. D. McMahon, C. C. Smith, D. F. Cihak, R. Wright, and M. M. Gibbons. Effects of digital navigation aids on adults with intellectual disabilities: Comparison of paper map, google maps, and augmented reality. *Journal of Special Education Technology*, 30(3):157–165, 2015.
- [29] S. Oeltze-Jafra and B. Preim. Survey of labeling techniques in medical visualizations. In *Proceedings of the 4th Eurographics Workshop on Visual Computing for Biology and Medicine, VCBM '14*, pp. 199–208, 2014.
- [30] B. Preim and P. Saalfeld. A survey of virtual human anatomy education systems. *Computers & Graphics*, 71:132–153, 2018.
- [31] PTC. Vuforia: Market-Leading Enterprise AR. <https://www.ptc.com/en/products/augmented-reality/vuforia>, 2020. [Online; accessed 02-June-2020].
- [32] M. Schneider, A. Bruder, M. Necker, T. Schluesener, N. Henze, and C. Wolff. A field study to collect expert knowledge for the development of AR HUD navigation concepts. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, pp. 358–362, 2019.
- [33] M. Sereno, X. Wang, L. Besançon, M. J. McGuffin, and T. Isenberg. Collaborative work in augmented reality: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 2021.
- [34] L. Szirmay-Kalos and G. Márton. Worst-case versus average case complexity of ray-shooting. *Computing*, 61(2):103–131, 1998.
- [35] M. Tatzgern, D. Kalkofen, R. Grasset, and D. Schmalstieg. Hedgehog labeling: View management techniques for external labels in 3D space. In *Proceedings of the 21st IEEE Virtual Reality (VR)*, pp. 27–32, 2014.
- [36] U. Technologies. Unity Website. <https://unity.com/de>, 2020. [Online; accessed 02-June-2020].
- [37] M. Uradziński and M. Bakuła. Assessment of static positioning accuracy using low-cost smartphone gps devices for geodetic survey points' determination and monitoring. *Applied Sciences*, 10(15):5308, 2020.
- [38] E. Veas, R. Grasset, E. Kruijff, and D. Schmalstieg. Extended overview techniques for outdoor augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):565–572, 2012.
- [39] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Müller, J. Wiegardt, H. Hortner, and C. Lindinger. Pervasive information acquisition for mobile ar-navigation systems. In *Proceedings 5th IEEE Workshop on Mobile Computing Systems and Applications*, pp. 13–20, 2003.
- [40] D. Waller and L. Nadel, eds. *Handbook of Spatial Cognition Hardcover*. Amer Psychological Assn, 2012.
- [41] S. White and S. Feiner. Sitelens: situated visualization techniques for urban site visits. pp. 1117–1120, 04 2009.
- [42] J. M. Wiener, S. J. Büchner, and C. Hölscher. Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation*, 9(2):152–165, 2009.
- [43] W. Willett, Y. Jansen, and P. Dragicevic. Embedded data representations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):461–470, 2017.
- [44] H.-Y. Wu, B. Niedermann, S. Takahashi, M. J. Roberts, and M. Nöllenburg. A survey on transit map layout – from design, machine, and human perspectives. *Computer Graphics Forum*, 39(3):619–646, May 2020.
- [45] H.-Y. Wu, S. Takahashi, D. Hirono, M. Arikawa, C.-C. Lin, and H.-C. Yen. Spatially efficient design of annotated metro maps. *Computer Graphics Forum (EuroVis 2013)*, 32(3):261–270, June 2013.
- [46] H.-Y. Wu, S. Takahashi, S.-H. Poon, and M. Arikawa. Scale-adaptive placement of hierarchical map labels. In *Proceedings of the 19th EG/NGTC Conference on Visualization (EuroVis2017)*, pp. 1–5, June 2017.