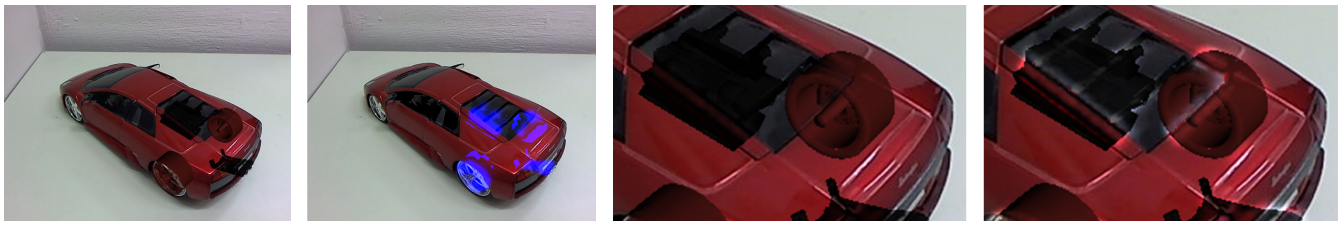
Adaptive Ghosted Views for Augmented Reality

Denis Kalkofen * Eduardo Veas† Stefanie Zollmann‡ Markus Steinberger§ Dieter Schmalstieg¶

Graz University of Technology



(a) (b) (c) (d)

Figure 1: Common Ghosted View versus Adaptive Ghosted View. (a) Clarity of an x-ray visualization may suffer from an unfortunate low contrast between occluder and occludee in areas that are important for scene comprehension. (b) Blue color coding of the important areas lacking contrast. (c) Closeup on areas that are lacking contrast. (d) Our adaptive x-ray visualization technique ensures sufficient contrast, while leaving the original color design untouched.

ABSTRACT

In Augmented Reality (AR), ghosted views allow a viewer to explore hidden structure within the real-world environment. A body of previous work has explored which features are suitable to support the structural interplay between occluding and occluded elements. However, the dynamics of AR environments pose serious challenges to the presentation of ghosted views. While a model of the real world may help determine distinctive structural features, changes in appearance or illumination detriment the composition of occluding and occluded structure. In this paper, we present an approach that considers the information value of the scene before and after generating the ghosted view. Hereby, a contrast adjustment of preserved occluding features is calculated, which adaptively varies their visual saliency within the ghosted view visualization. This allows us to not only preserve important features, but to also support their prominence after revealing occluded structure, thus achieving a positive effect on the perception of ghosted views.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities;

1 INTRODUCTION

Augmented Reality (AR) displays are able to uncover hidden objects, yielding the impression of seeing through things. Such x-ray visualization is a powerful tool in exploring hidden structures within their real world context. AR achieves this effect by overlaying otherwise hidden structure on top of the real world imagery. However, without considering the information which is about to be removed, the augmentation may lead to a number of perceptual problems [17].

Although the desired effect of seeing through things does not come natural to people, artistic and scientific illustrators have created many perceptually effective methods [12]. These methods often rely on artistic abstractions and have been studied in computer graphics under the term non-photorealistic rendering (NPR) [10]. One such technique, so-called *ghosted views*, selectively applies transparency to the occluding structure (occluder) to unveil the occluded object (occludee) [8, 7, 20, 18]. Thus, a ghosted view requires simultaneous uncovering of occludee and the preservation of important features of the occluder.

Ghosted views find their origin in illustrations, where the artist carefully chooses viewing angle, fidelity and the representation of distinctive features [12]. Most of these characteristics dinamically change with motion in the real world. Thus, a separation cannot be guaranteed without additional measures because the occluder in AR has arbitrary material and color. This makes it difficult to enforce the desired appearance, i.e., make the occludee and the important structures of the occluder simultaneously stand out to a sufficient degree. Conversely, both artistic illustrations and interactive VR applications have full control over the scene, whereby materials and colors for the three components – occluder, occludee, background – can be chosen so that sufficient contrast is achieved between any two of them. Previous work of Avery et al. [2] and Kalkofen et al. [17] apply a virtual line overlay to enforce visibility. However, this line overlay being predefined and added to the real world imagery all the time, makes the resulting ghosted view rather obtrusive. To retain as much of the real world as possible, we aim to make only minimal use of artificial fill-in colors. Hence, we require an adaptive choice of suitable colors, selected following an analysis of the current view .

To address the issue we propose a rendering technique that produces *adaptive ghosted views*. The technique analyzes the contrast of a ghosted view to identify fragments in the occluding structure that lack contrast to sufficiently stand out against the occludee. Hereby, an enhancement must be computed to achieve sufficient contrast. For example, the ghosted view in (Figure 1(c)) suffers from a lack of contrast between occluder and occludee. Our approach identifies adversely blended elements, before it increases

*e-mail: kalkofen@icg.tugraz.at
†e-mail: veas@icg.tugraz.at
‡e-mail: zollmann@icg.tugraz.at
§e-mail: steinberger@icg.tugraz.at
¶e-mail: schmalstieg@icg.tugraz.at
website: https://icg.tugraz.at/~denis/projects/agv

their contrast so that both occluder and occludee sufficiently stand out from each other (Figure 1(d)).

We validate our solution with a study comparing our technique to ordinary alpha blending and unmodified ghosted view. Our study revealed that low contrast between occluder and occludee makes them indistinguishable to such extent that ghosting is no more effective than simple alpha blending. We found that our technique enhances perception of important features that define the occluder and improves spatial understanding. We describe details of our real-time capable implementation on the GPU, and give selected examples of the application of our approach.

2 Related work

X-Ray visualizations provide a view on occluded objects by making the occluder partly or completely transparent. However, when using a simple uniform transparency modulation to reveal occluded objects, important visual cues of the occluder often get lost. Since such visual cues contribute to the interpretation of depth [21] and to the interpretation of shape, it is important to retain them in the visualization. Thus, in contrast to simple transparency modulation, ghosted view techniques vary transparency values non-uniformly over the occluding object based on a measurement of importance.

Out of the many cues for depth perception [21], our technique strives to preserve correct perception of occlusion and shape. We focus on monoscopic, pictorial cues, as they can be represented with any display device, and closely investigate their preservation to prevent the loss of legibility, frequently present due to the dynamics of AR.

The approach presented in this paper maintains the legibility of ghosted views in AR environments. In this section, we review the state of the art in ghosted view rendering and the state of the art of legibility maintenance in AR.

2.1 Ghosted views

Ghosted views have been successfully deployed in various application areas of computer graphics. For example, early work on illustrative rendering in 3D computer graphics showed ghosted views [8] that used a description of the visualization's intention as importance measurement [26]. Burns et al. [5] shows how view aligned wedge cuts can be overlaid with an edge-based ghosted view.

Illustrative ghosted views have also been applied to volume rendering. Bruckner et al. [4] and Krueger et al. [20] presented two approaches to generate ghosted views from volumetric data. While Krueger et al. analyzed curvature to modulate transparency values, Bruckner et al. use information about the shading of the object.

In AR, ghosted views have been generated based on an image analysis of the system's video feed [17, 29, 25] and based on a feature analysis of a given 3D CAD model, which is registered to its real world counterpart. For example, Bichlmeier et al. [3] applied the approach from Krueger et al. to 3D volumetric data registered within the AR environment. Kalkofen et al. [17] present edge based ghosted views derived from an edge filtering of either the video image or an offscreen rendering of a registered 3D model.

Image based ghosted views have largely relied on an analysis of visual saliency. The saliency map [16] determines how different a location is from its surround, hence how much visual attention it attracts [1]. Zollmann et al. [29] and Sandor et al. [25] both interpret visually salient features as important for scene understanding and retained them in the ghosted view.

A point common to all these approaches is that they try to address the question of which features of the occluder should be retained in the final visualization. However, our approach focuses on how these features will be retained in the final composition of occluder and occludee. The input to our approach is a state of the art ghosted view which can be generated from both a model and an image analysis.

2.2 Legibility of Augmented Reality visualizations

The identification of important image structure is one part of the problem for the ghosted view visualization. The next important point is to ensure that such information (since it is selected, sparse and important information) is legible in all means. Legibility refers to a characteristic of any visual representation, either graphical or textual, that it must have discernible features to clearly convey its meaning [13]. Legibility is limited by issues that affect perception, such as foreground-background interference and visual clutter. Foreground-background interference occurs when overlaying an image (foreground) onto another (background) with similar color properties. In AR, the color properties of the real world (e.g., captured video) constantly change, complicating legibility. Furthermore, with increasing number of objects, information density increases and objects overlap to such an extent that identifying them might become impossible [27, 19, 23].

Labeling in AR is a typical application where research has focused on dynamically adapting the visualization so that the user is not disturbed by the final rendering and the actual content is well visible. For instance, Gabbard et al. [9] applied a contrast enhancement between video and text labels to increase their legibility. Rosten et al. [24] extracted corner features to avoid placing labels on important parts of the video image. Recently Grasset et al. applied saliency analysis for adaptive label placement [11].

In this paper, we propose a parametric approach that is based on prominent visual features that represent structure (e. g., visual saliency) and that adapts to maintain the visual distance between occluder and occludee.

3 Method

The goal of our method is to ensure that the image information required to understand the ghosted view is clearly visible. To achieve this, using state of the art techniques we obtain a map identifying important structure – Importance Map (*IM*) – and a ghosted view $G_0(x,y)$. We then analyze the resulting ghosted view, comparing it with *IM*. If this analysis determines that fragments important for scene understanding (i.e., those in *IM*) have become indistinguishalbe after blending, we increase the contrast of these fragments in order to improve their visibility, yielding an *adaptive ghosted view* $G_A(x,y)$.

Our method to generate G_A is shown in Figure 2. It involves three stages:

1. In the first stage, we generate a ghosted view using state of the art rendering techniques. Therefore, we first compute an importance map $IM(x,y)$. The $IM(x,y)$ guides the computation of *transparency* values used to blend the occluder $O(x,y)$ and the occludee $D(x,y)$, yielding the initial ghosted view G_0 (Section 3.1).

2. In the second stage, we *analyze* the initial ghosted view to find fragments which were identified as important before blending but became indistinguishable thereafter, i.e., are not salient enough in the initial ghosted view (Section 3.2).

3. In the third stage, we visually *emphasize* the fragments identified in the second stage (Section 3.3).

3.1 Transparency computation

To turn occluding structures transparent, the algorithm must identify pixels in the video feed which belong to occluding structures. Using a uniform transparency does not lead to a suitable depth perception and introduces clutter that makes it difficult to see the occludee [6]. Therefore, ghosted views generally compute individual transparency values for every pixel of the occluder, revealing the occludee only at fragments that are not important for perceiving the shape of the occluder. To capture the importance of individual

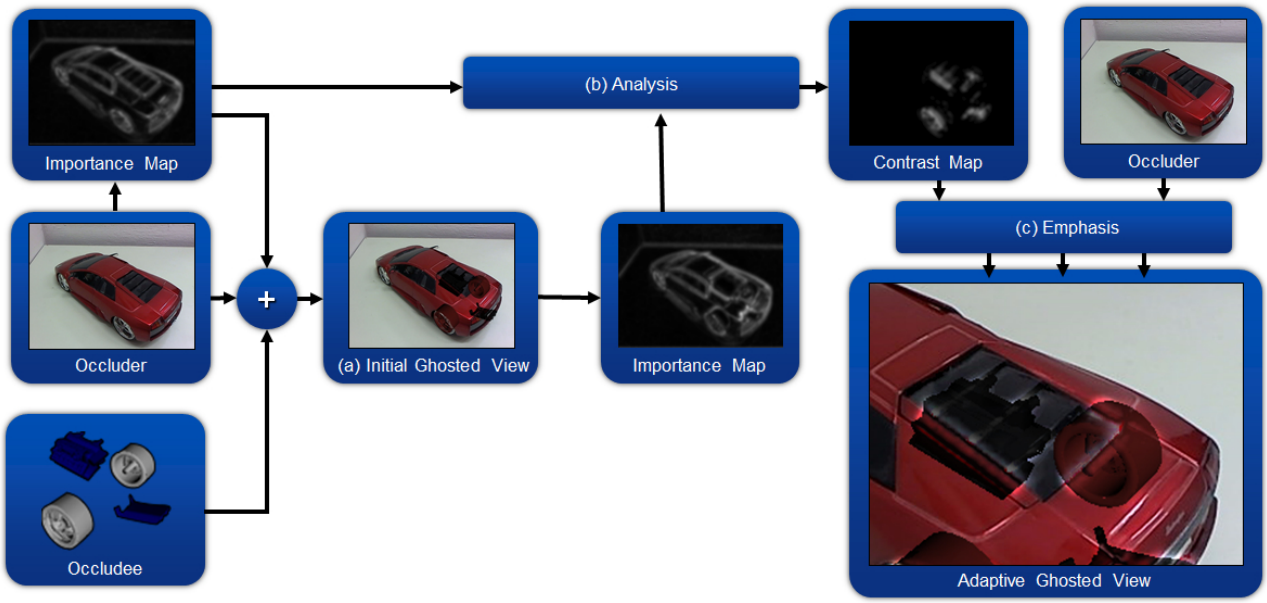


Figure 2: Method overview of adaptive ghosted views. (a) Occluder and occlude are combined to a ghosted view view by using state of the art transparency mapping of the importance map of the occluder. The resulting ghosted view is missing important features. Therefore, another importance map is calculated on the resulting ghosted view. (b) This importance map is compared to the importance map of the occluder. Elements that were identified to be important for the occluder, but are no longer detected as important in the ghosted view will be emphasized. (c) After applying the emphasis on the adaptive ghost, important structures are again clearly visible.

fragments, a so-called *importance map IM* is computed first, from which the the transparency values can be derived. Different strategies for computing *IM* exist.

Model-based importance map The *IM* is generated using a 3D model of the occluder to determine importance values according to the principal curvature. The rationale of this approach is that strongly curved portions of the occluder define its shape and should be more opaque, while flat areas can be made more transparent [14, 15].By choosing different settings for scale and bias when converting curvature into importance, the occluder can be made more opaque (dense occluder) or transparent (sparse occluder). A model-based technique can only be used if a registered 3D CAD model of the occluder exists.

Image-based importance map Since AR often lacks a complete 3D CAD model, image-based approaches have been proposed, which produce the *IM* from an analysis of the video image only. To generate image-based ghosted views, image information that is likely to be important for scene understanding is extracted from the video image and preserved in the final ghosted composition. Various methods use edges alone or combined with saliency maps to determine important image regions [18, 29, 25].

For our technique, we use a curvature analysis $C(x,y)$ if there is a registered 3D CAD model available. In case there is none, we run a saliency analysis $S(x,y)$ on the video frame to generate the *IM*:

$$IM = \begin{cases} C, & \text{if } model-based \\ S, & \text{if } image-based \end{cases} \quad (1)$$

We map the importance values to alpha values depending on a user defined transparency value t which defines how sparse or dense the occluder will rendered ($\alpha = t \cdot IM$). Finally, we use the alpha values to generate the initial ghosted view $G_0(x,y)$:

$$G_0 = O \cdot \alpha(x,y) + D(1-\alpha(x,y)). \quad (2)$$

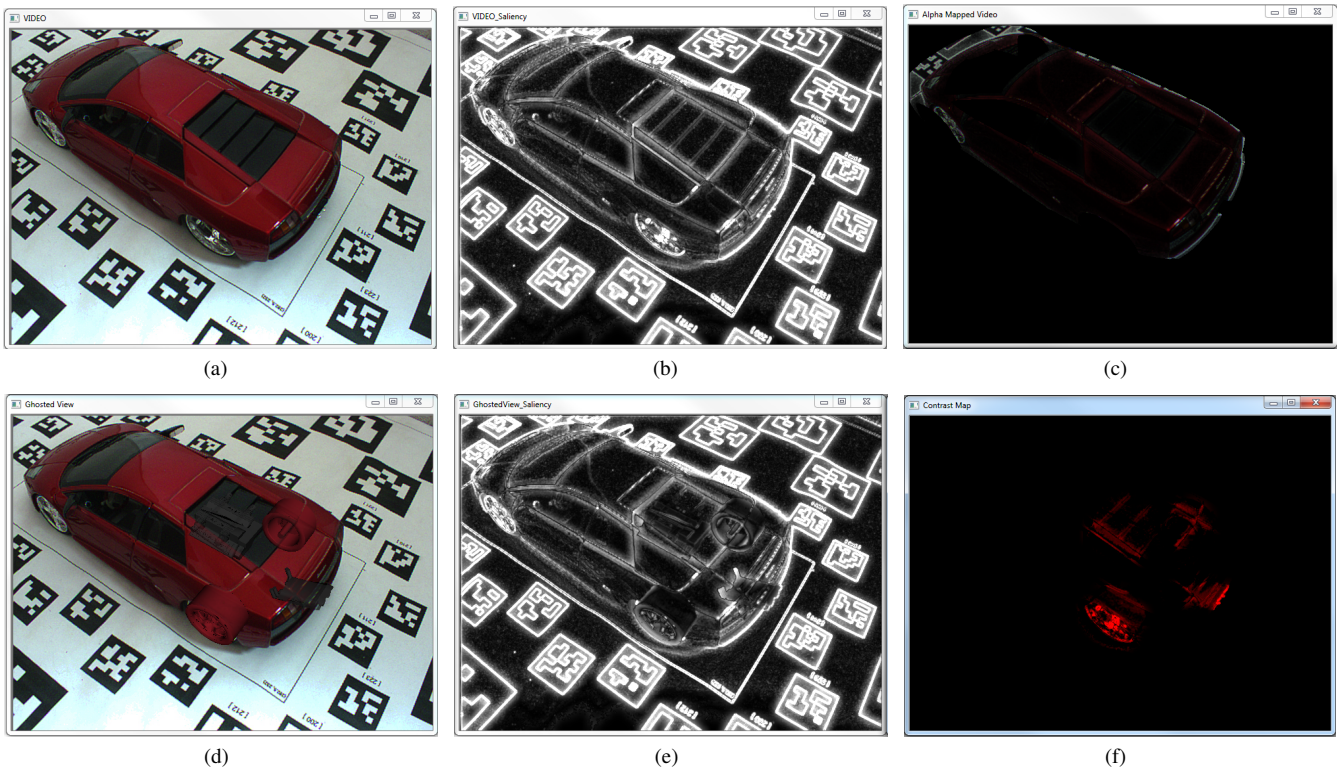
3.2 Analysis

The problem with both image and object-based ghosted views is that the method has no control over the properties of the final composition. In particular, the method does not guarantee that highly important structures remain visible (salient) after blending or that a viewer can discriminate occluder from occludee. For instance, Figure 1(a) and (c) show a ghosted view in which fragments which are important for scene understanding are less salient. This is due to a lack of contrast between the occluder and the occludee. To overcome such deficiencies we compare the saliency map of the initial ghosted view with the saliency map of the incoming video, with the goal to determine a contrast map C which represents at each pixel the amount of missing contrast. We define the amount of missing contrast as the difference between the contrast of the initial ghosted view $C_{initial}$ and the desired amount of contrast $C_{desired}$.

$$C = max(0, C_{desired} - C_{initial}), \quad (3)$$

By using the saliency map of the initial ghosted view $S(G_0)$ for $C_{initial}$ and the importance map of the occluder *IM* for $C_{desired}$, equation 3 yields to the amount of missing contrast required to maintain the initial conspicuity of important occluding elements. This can be seen in Figure 3. The saliency of the incoming video stream is interpreted as *IM* and used to compute the alpha map applied to generate the initial ghosted view $S(G_0)$ (shown in the lower left image). Afterwards, we compute the saliency map of the ghosted view $S(G_0)$ which we compare to the saliency map of of the incoming video (IM). The difference between both saliency maps is visualized in the red color channel in the image to the right in the lower row.

Note that C encodes a per-fragment value of missing contrast. But conspicuity depends on the values of surrounding fragments. Hence, while this approach is able to maintain the conspicuity it does not consider the contrast at the border between occluder and occludee, mainly because it has no information about object boundaries.Notice how in Figure 4 no lack of contrast is identified at the



(a) (b) (c)

(d) (e) (f)

Figure 3: Difference contrast map. The video feed (a) is used to detect important features (b) which are turned into alpha values (c) and blend with hidden virtual structure yielding to the ghosted view (d). To compute missing contrast, we compute a saliency map from the initial ghosted view (e) which we then compare to the saliency map of the video feed. The resulting differences are presented in the red channel in (f)

silhouette of the hidden red wheel, which defines the border to its occluding surroundings.

Therefore, the simple difference operation between importance values from pre- and post-blended imagery does not allow to control saliency values from the initial view, i.e. to increase contrast. The desired contrast depends on the scene (e. g., clutter) or on the user's perception, it is often not sufficient to maintain the amount of initial contrast only. Therefore, we provide an optional increase of the desired contrast map to support the perception of ghosted views also in scenarios with less salient features. To incorporate the silhouette of hidden structure we allow adding a user controlled value $tc_{silhouette}$ to $C_{desired}$ for silhouette pixels where blending occurs, i.e .where $0 < \alpha < 1$.

To further provide setting a minimal amount of desired contrast and scaling the desired contrast we allow setting a user controlled multiplier γ and a value $tc_{minimal}$ in $C_{desired}$ at all other pixels where blending occurs, i.e. where $(0 < \alpha < 1)$. Those fragments cover the critical area between occluder and occludee, thus it is essential that they be distinguishable in the final ghosted view. Using $tc_{minimal}$ in $C_{desired}$ allows to define a minimal amount of contrast all over the occludee, however, this approach ignores differences in contrast between important features above this threshold. Therefore, we further allow to incorporate the initial importance map in $C_{desired}$ by using $\max(tc_{minimal}, IM)$ at pixels where $0 < \alpha < 1$.

Notice again that the necessary amount of contrast to perceive all important information depends on a number of conditions, such as the amount of clutter, the AR display, the environment where the application is used (indoor or outdoor) or the perception of the user. Therefore, a number of methods may exist for setting reasonable values for $C_{desired}$.

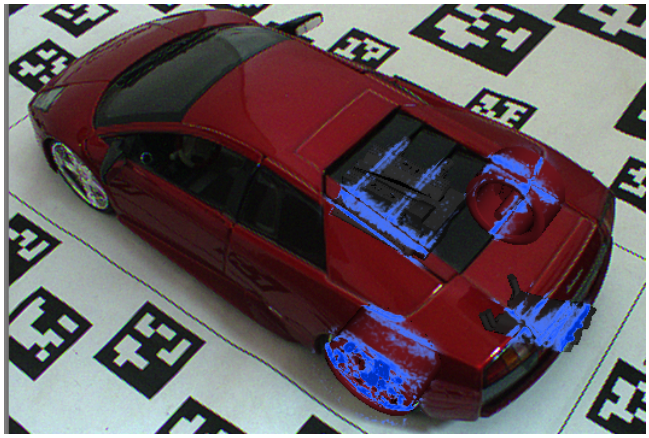


Figure 4: Color coded contrast map rendered on top of the initial ghosted view. By using a difference map only, we are not able to detect missing contrast at the silhouette of hidden structure.

3.3 Emphasis

If C is greater than zero, the color of occluder and occludee are too similar for a viewer to perceive the difference. As a consequence, we want to increase the conspicuity for these fragments (i. e., increase the contrast between occluder and occludee). To that end, we modulate the color of the occluder in *CIE* $L^*a^*b^*$ space. We chose *CIE* $L^*a^*b^*$ because it closely reflects visual perception. The goal of the modulation is to make important fragments in the adapted ghosted view sufficiently conspicuous.Therefore, based on the con-

trast map C, we determine which fragments should be modulated and how much. Because the conspicuity of a fragment depends on the contrast to its surroundings, we run a Gaussian kernel GK on C to distribute this information to neighboring fragments. Notice that this operation additionally removes any hard edges in C.

Similarly, we also have to consider the surrounding of occluded elements. Therefore, we also apply Gaussian smoothing to the occludee before we transform into *CIE* $L^*a^*b^*$ color space, resulting in D_{GK}^{LAB}.

$$D_{\mathrm{GK}}^{\mathrm{LAB}} = LAB(GK(D)), \tag{4}$$

where LAB defines the transformation to *CIE* $L^*a^*b^*$ color space.

To obtain sufficient modulation, we first determine which point O_{max}^{LAB} in the *CIE* $L^*a^*b^*$ space will achieve the highest contrast to D_{GK}^{LAB}:

$$O_{max}^{\mathrm{LAB}} = k_0^{\mathrm{LAB}} + sgn(k_0^{\mathrm{LAB}} - D_{\mathrm{GK}}^{\mathrm{LAB}}) \cdot k_1^{\mathrm{LAB}}/2, \tag{5}$$

where k_0^{LAB} is the center of the color space (usually $L^* = 50$, $a^* = 0$, $b^* = 0$), k_1^{LAB} is the extent of the color space and sgn is the sign operator.

Using the initial color of the occluder, we can determine the direction dir of the delta vector which will result in the highest contrast:

$$dir = norm\left(O_{\mathrm{max}}^{\mathrm{LAB}} - O^{\mathrm{LAB}}\right), \tag{6}$$

where $norm$ normalizes the vector.

Then, we move the initial color O^{LAB} of the occluder towards this maximum contrast value O_{max}^{LAB} by using the previously derived direction of the delta vector dir:

$$O_{\mathrm{A}}^{\mathrm{LAB}} = O^{\mathrm{LAB}} + GK(C) \cdot dir, \tag{7}$$

The adaptive ghosted view G_A is then calculated as:

$$G_A = LAB^{-1}(O_A^{\mathrm{LAB}})\alpha + D(1-\alpha). \tag{8}$$

3.4 Limitations

This method computes emphasis for a region, without knowledge of object coverage (i.e., what fragments belong to an object or another). It produces compelling results in scenes with rather uniform color distributions (Figure 1), but has certain caveats when applied in scenes with more arbitrary colors. The problem is that emphasis is still computed per fragment, fragments of a single object may move in different directions and/or by different amounts (see Figure 5(a)). The method needs to guarantee coherent changes accross the fragments of an object. The GK operations applied above ensure that changes are smooth. Still, after a pilot study we identified aditional constraints:

Direction of Delta Vector - Lightness Only Modulating only the L^* component delivers more perceptually pleasing results than changing the tone of the occluder.As an example, consider a light blue occluder, with a light green occludee underneath. The color furthest from light green is dark red. Choosing this extreme color to increase the contrast between occluder and occludee strongly alters the impression of the occluder. By changing only the L^* component, light green changes to a dark green, yielding an increase in contrast, and maintaining the tone of the occluder's color.

Direction of Delta Vector - Unified Sign Simultaneously increasing and decreasing lightness values for a single object alters its shading in different directions, leading to adverse changes in its perception. For example, Figure 5(a) shows an image where contrast has been increased between occluder and occludee, resulting in a misleading perception of shape of the head of the dog. To avoid this effect, we unify the direction over an area of lightness modifications. As can be seen in Figure 5(b), this better preserves the original shading of the head.

(a)

(b)

Figure 5: Consistent Lightness Modulation (a) Simultaneously increasing and decreasing pixel in front of the same object may influence its perception. Notice the artifacts in the shading of the head of the dog in the left image. (b) We avoid altering shading information by consistently modulating lightness values of the occluder.

4 Implementation

Performance has been an obstacle in the deployment of saliency oriented methods. The present approach is entirely implemented in OpenGL Shading Language (GLSL), it runs in full frame, in real-time. Our saliency analysis represents an accelerated variant of the technique described by Mendez et al. [22] and Veas et al. [28], which is itself is an adaption of the center-surround saliency analysis presented by Itti et al. [16]. In both approaches, a hierarchical multi-channel contrast measure is computed for every image pixel.

Veas et al. consider conspicuities in three dimensions: lightness (L), red-green (O_r), and blue-yellow color opponents (O_b). For every pixel in the input image, a conversion to *CIE* $L^*a^*b^*$ space yields these three values $k \in L, O_r, O_b$. Subsequently, an image pyramid with p levels is created on top of the full scale *CIE* $L^*a^*b^*$ image and the saliency for each pixel is determined as follows:

$$c_k(x,y) = \frac{\sum_{n=0}^{2}\sum_{m=n+2}^{n+4} k_n(x,y) - k_{n+m}(x,y)}{p}, \tag{9}$$

where $k_i(x,y)$ is the conspicuity $k \in L, O_r, O_b$ at pixel x, y and image

Figure 6: Experimental scenes from the first study. We compare alpha blending (left) and common ghosted views (middle) with emphasized ghosted views (right) in scenes with conflicting visual characteristics. Notice the emphasized structure covering the virtual objects marked with red arrows in the lower row and the emphasized structure within the enlarged inset in the upper row.

pyramid level i. Compared to the approach by Itti et al., Veas et al. dropped the absolute value operator in the inner sum to allow for their subsequent saliency modulation.

The filter pyramid computation of level k_n can be defined in a recursive manner as the down-sampled convolution of the next lower level with the pyramid filter f_p:

$$k_n(x,y) = \sum_{n_x,n_y=-\infty}^{\infty} k_{n-1}(n_x,n_y) \cdot f_p(2x-n_x, 2y-n_y). \qquad (10)$$

A fast saliency analysis is achieved through the built-in mip-mapping support of current GPUs. Hereby, f_p becomes a box filter, which may lead to block artifacts in the final saliency map. By using texture hardware, the saliency computation requires only a few very fast accesses to the mip-map textures, while the generation of the mip-maps takes comparably long. Note that if a high quality image pyramid is required, the box filter can be replaced by a Gaussian filter.

We derive a single filter for $c_k(x,y)$ by combining equation 9 with 10 and evaluating the sums. If f_p is a box filter, the convolution of multiple box filters again yields a box filter with wider support. The sum of the differently sized box filters yields a stair step function. If a Gaussian pyramid is used, we end up with a sum of convolutions of the Gaussian pyramid filter. While a convolution of two Gaussians yields a Gaussian, there is no simple function to describe the sum of different Gaussians. Nevertheless, we can approximate this filter by the sum of two Gaussians, where one represents all terms with positive sign in equation 9 and the other all terms with negative sign. The two Gaussian filters are separable and thus can be computed in two iterations of the image. If the filters are pruned to a size of approximately 40 samples, the saliency computation is even faster than the mip-mapping version.

5 Validation

Our intention is to overcome difficulties in depth perception arising from conflicting visual characteristics of the AR ghosted view, by preserving and even enhancing depth perception cues. We also strive to achieve at least the performance of a state-of-the-art approach in every other situation (i.e., when there are no visual conflicts). To validate our approach, we designed two experiments where we compared it to the mainstream (alpha blending) technique and the state-of-the-art image-based ghosted views in AR. Hence, each experiment counted three conditions: adaptive ghosting (G_A), *IM* blending (G_0), and alpha blending ($\alpha(x)$). Our intention was to compare the techniques at the perceptual level on the basis of the spatial understanding elicited by each of them. The experiments differed in the data and participants while the execution and procedure was the same.

5.1 Experiment 1: Enhancing Conflicting Regions

In this experiment we compared performance with G_A, G_0 and $\alpha(x)$ in scenes presenting conflicting visual characteristics. We defined regions of conflicting visual characteristics where colors of occluder and occludee clash and become indistinguishable from one another.

We hypothesized that participants could determine the spatial organization of the scene from importance guided visualizations (G_0 and G_A), with higher performance (less errors) when applying emphasis (G_A).

H1 - Adaptive ghosting leads to significantly less errors in spatial understanding for conflicting regions.

H2 - Adaptive ghosting improves the representation of important features.

5.1.1 Data

Our technique aims at preserving the occlusion cue, so we chose still images as a basis for comparison, as this prevents depth perception from motion cues and helps control factors resulting from unstable tracking during the experiment. Three thematic scenes were chosen for the experiment, observing a main physical object with virtual objects inside and a context made of physical objects

Figure 7: Experimental scene from the second study. We compare alpha blending (left) and common ghosted views (middle) with emphasized ghosted views (right) in a scene with little conflicting areas.

and ambient photography. Scene topics were selected to keep participants engaged in the task, while representing examples of real-world applications for ghosted views (Figure 6). We chose not to represent floor and back planes in hidden objects to restrict depth perception to occlusion and relative size cues (i.e., reducing the influence of perspective cues, vanishing points, etc.). We created the scenes by registering the (virtual) camera with our in-house natural-feature tracking library to a scene with physical objects. The scene and registration were recorded to generate the composition for each condition. The colors of occluded objects were chosen to introduce regions with background/foreground interference. The parameter x for the alpha blending $\alpha(x)$ condition was the average α produced by importance blending. To estimate it, an *IM* was computed for each scene yielding the average α as the sum of the importance for pixels in the region of the occludee.

5.1.2 Procedure

The study was divided in two stages: interpretation and comparison. In the interpretation stage, participants had to interpret a scene and respond to questions involving spatial reasoning: identify objects inside the occluder, spatial relations between occluded objects and occludee, and mapping information to the physical object. Mapping was done by pointing the location of hidden objects relative to a physical object that was only revealed at this point. Thereafter, they received a questionnaire for self assessment. Participants alternated the different scenes for each technique. The order of technique and scene were randomized.

In the comparison stage, participants were presented with the three alternatives for each scene and they had to compare their qualities in terms of perception and aesthetics, under the premise of representing x-ray visualization. The overall duration was approximately twenty-five minutes.

5.1.3 Participants

We recruited 17 people to take part in this experiment from the university population (16 male, 1 female). The study followed a repeated measures design, where each participant performed with all $\alpha_{(x)}$, G_0, G_A, each in a different scene. Each scene had six spatial reasoning targets to assess spatial understanding.

5.1.4 Results

We quantified performance per scene by averaging the number of targets found with the number of total targets. This resulted in a performance rate per scene. To analyze the effect of technique and scene we performed a repeated measures ANOVA. We found no effect of scene on performance. We found a significant effect of technique on performance rates ($F(1,11) = 4.239$, $p < 0.01$). LSD comparisons showed that G_A ($M = 0.7252$) lead to significantly higher performance than $\alpha_{(x)}$ ($M = 0.530$, $p < 0.02$, Bonferroni $\alpha = 0.017$) and also significantly better than G_0 ($M = 0.510$, $p < 0.01$, Bonferroni $\alpha = 0.017$), thus we retain H1. There was no significant difference between $\alpha_{(x)}$ and G_0.

Our subjective comparative test delivered categorical data, which was analyzed using a Friedman test and post-hoc Wilcoxon signed-rank tests. A Friedman test revealed a statistically significant difference with respect to clear representation of structurally important features on the occluder ($\chi^2 = 7.0$, $p < 0.03$) (Q: "The technique clearly represents important features (windows, frames, buttons, labels) in the occluder"). A post-hoc analysis with Bonferroni adjustment ($\alpha = 0.017$) showed no statistical differences between $\alpha_{(x)}$ and G_0 trials ($Z = -2.294$, $p = 0.022$) or between G_0 and G_A trials ($Z = -0.33$, $p = 0.795$), despite a perceived improvement in the G_0 vs. $\alpha_{(x)}$ trials. However, there was a statistically significant improvement in the G_A vs. $\alpha_{(x)}$ ($Z = -2.60$, $p < 0.009$). Hence, we partially retain H2.

5.2 Experiment 2: Emphasized Ghosting

Having achieved a significant improvement for conflicting regions, we evaluated our emphasis technique in scenes without color interference (Figure 7). Again, we compared performance with G_A, G_0 and $\alpha(x)$ based on perception of spatial configuration. We expected G_0 to lead to higher performance than $\alpha(x)$ in this case. Further, we hypothesized that our technique (G_A) would perform at least as well as G_0.

H3 - Adaptive ghosting does not introduce errors in spatial understanding for non-conflicting regions

The procedure for this experiment was exactly the same as for Experiment 1.

5.2.1 Data

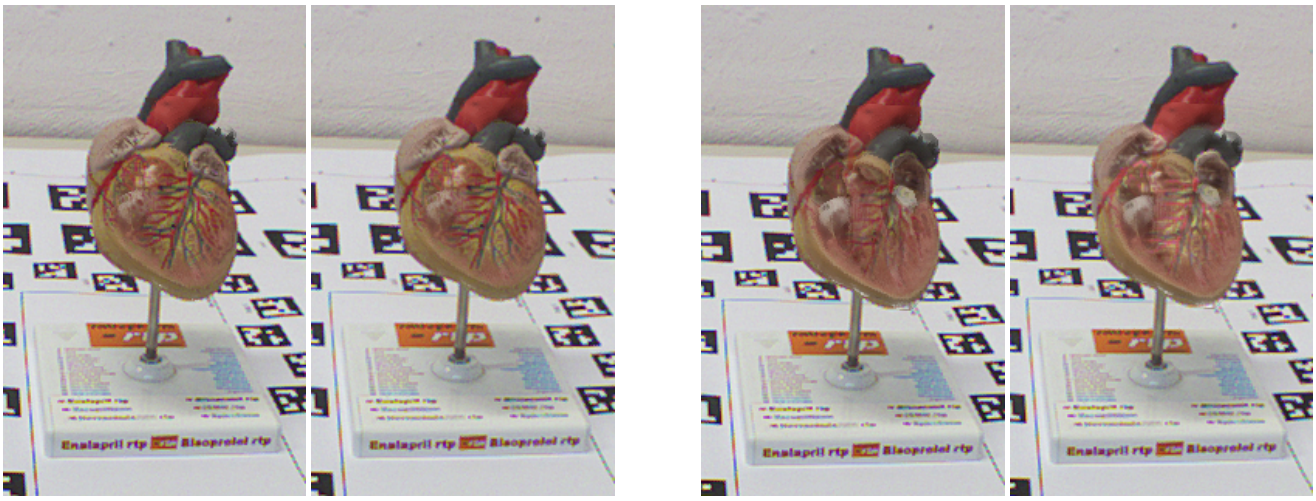
For this experiment we used the same scenes as in the previous experiment but modified the virtual objects to remove background/foreground interference. Thus the colors of virtual objects were chosen to be clearly distinguishable from those of the foreground object.

5.2.2 Participants

We recruited 16 participants to take part in this experiment (16 male, 1 female). As previously, the study followed a repeated measures design, randomizing scene and technique for each participant.

5.2.3 Results

We found a significant effect of technique on performance rates ($F(1,14) = 7.742$, $p < 0.005$). LSD comparisons showed that G_0 ($M = 0.829$) lead to significantly higher performance than $\alpha_{(x)}$ ($M = 0.482$, $p < 0.002$, Bonferroni $\alpha = 0.017$). G_A ($M = 0.722$) also lead to significantly higher performance than alpha ($M =$



(a) (b) (c) (d)

Figure 8: Comparison between state of the art Ghosted View and Adaptive Ghosted View in two different transparency settings. (a) Ghosted View generated from a low transparency value. (b) Adaptive Ghosted View using the same low transparency value as in (a). Since the contrast of important features occluding features is almost not affected after blending the ghosted occluder and the occludee almost no difference between state of the art and adaptive ghosted view is noticeable. (c) Ghosted View generated from a high transparency value. Notice how occluding arteries and veins become difficult to distinguish from the inner structure of the heart. (d) Adaptive ghosted view using the same high transparency value as in (c). Due to the emphasis the adaptive ghosted view better preserves the visibility of important occluding structure such as the arteries and veins.

0.482, $p < 0.01$, Bonferroni $\alpha = 0.017$). There was no significant difference between G_A and G_0. Hence, we retain H3.

Our subjective evaluations showed that participants were significantly more confident to "estimate the location of hidden objects" with both ghosting versions as compared to alpha blending ($M(G_A) = 2.14, M(G_0) = 2.21, M(\alpha_x) = 3.93, F = (1,14) = 6.597, p < 0.001$). This supports the results obtained for performance.

5.3 Discussion

The results of our experiments support the hypotheses that our technique does not introduce errors in spatial assessment for scenes with no color interference. More interestingly, they support the assumption that in cases where there is interference between foreground and background colors, the state-of-the-art technique (ghosting) degrades. In experiment one, People performed with ghosting as badly as with alpha blending. Conversely, our emphasis technique enhances features of the occluder in these cases, enabling participants to perform significantly better than in the other two conditions.

In the second experiment, subjective assessment indicated that participants feel significantly more confident to estimate spatial relationships and location of hidden objects with ghosting techniques. The first experiment indicated that the emphasis clearly retains important structure of the occluder and is thus preferred.

6 Examples

With increasing transparency, state of the art ghosted views often fail to provide enough contrast between occluding and occluded structures. This is demonstrated in Figure 8. It shows a ghosted views of the anatomy of the heart in two different transparency settings (its setup is illustrated in Figure 9). Figure 8(a) and Figure 8(b) allow to compare state of the art with adaptive ghosted view using a low transparency value. Since in this example important elements, such as the occluding arteries and veins have not been directly blend with inner structure the adaptive ghosted view does not alter the rendering. Therefore, both visualizations are very similar.

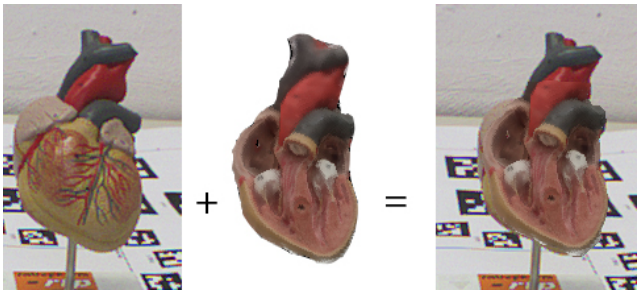


Figure 9: Scene setup of the example shown in Figure 8. We augment a physical model of the human heart (left) with a textured 3D CAD model of its inner anatomy (middle).

Figure 8(c) and Figure 8(d) demonstrate both techniques using a high transparency value. The value causes the ghosted view to blend hidden structure with highly important occluding elements only. In contrast, the adaptive ghosted view in Figure 8(d) is able to provide more contrast between very important occluding elements and the inner anatomy of the heart.

7 Conclusion and Future Work

Ghosted view visualizations are known to enhance the understanding of the relationships of elements for x-ray vision. This paper described a scheme that can be parameterized to obtain compelling ghosted views. It demonstrates the need to account for interference between occluder and occludee, and overcomes the problem by defining a minimum distance required between the two (t_c).

Our scheme further defines a contrast map representing regions that require enhancement. We also show how to apply the enhancement in a way such that visual coherence is maintained. Our initial study shows that our technique significantly improves perception

of important features of the occluder, while still showing enough of the occludee. We believe that these findings highlight the viability of the proposed solution and deem it a contribution worthy of any x-ray visualization toolkit. In the future we will extend our approach to multiple level of occluding elements.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers. This work was supported in part by the Austrian Science Fund (FWF) under contract P24021 and P23329, and the Austrian Research Promotion Agency (FFG) FIT-IT project Construct (830035).

REFERENCES

[1] R. Achanta, S. Hemami, E. Francisco, and S. Suessstrunk. Frequency-tuned Salient Region Detection. In *IEEE International Conference on Computer Vision and Pattern Recognition*, 2009.

[2] B. Avery, C. Sandor, and B. H. Thomas. Improving spatial perception for augmented reality x-ray vision. In *Proc. of the IEEE Conference on Virtual Reality*, pages 79–82, 2009.

[3] C. Bichlmeier, F. Wimmer, H. Sandro Michael, and N. Nassir. Contextual anatomic mimesis: Hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In *ISMAR '07: Proceedings of the International Symposium on Mixed and Augmented Reality*, pages 129–138, 2007.

[4] S. Bruckner and M. E. Gröller. Exploded Views for Volume Data. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1077–1084, 9 2006.

[5] M. Burns and A. Finkelstein. Adaptive cutaways for comprehensible rendering of polygonal scenes. In *Proceedings of ACM SIGGRAPH Asia*, pages 1–7, New York, NY, USA, 2008. ACM.

[6] F. C. Crow. Shaded computer graphics in the entertainment industry. *Computer*, 11(3):11–22, 1978.

[7] J. Diepstraten, D. Weiskopf, and T. Ertl. Transparency in interactive technical illustrations. *Computer Graphics Forum*, 21:317–325, 2002.

[8] S. K. Feiner and D. D. Seligmann. Cutaways and ghosting: satisfying visibility constraints in dynamic 3d illustrations. *The Visual Computer*, 8:292–302, 1992.

[9] J. L. Gabbard, V. Tech, J. E. S. Ii, V. Tech, S.-j. Kim, V. Tech, and V. Tech. Active Text Drawing Styles for Outdoor Augmented Reality : A User-Based Study and Design Implications. In *IEEE Virtual Reality Conference*, pages 35–42, 2007.

[10] B. Gooch and A. Gooch. *Non-Photorealistic Rendering*. AK Peters Ltd, 2001. ISBN: 1-56881-133-0.

[11] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and S. Dieter. Image-driven view management for augmented reality browsers. In *Proc. of the International Symposium on Mixed and Augmented Reality*, pages 3–12, 2012.

[12] E. R. S. Hodges, editor. *The Guild Handbook of Scientific Illustration*. John Wiley & Sons, Hoboken, NJ, 2nd edition, 2003.

[13] R. Ingram and S. Benford. Legibility enhancement for information visualisation. In *Proceedings of the 6th conference on Visualization '95*, VIS '95, pages 209–, Washington, DC, USA, 1995.

[14] V. Interrante, H. Fuchs, and S. Pizer. Enhancing transparent skin surfaces with ridge and valley lines. In *Proceedings of IEEE Visualization*, pages 52–59, 1995.

[15] V. Interrante, H. Fuchs, S. M. Pizer, and S. Member. Conveying the 3d shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, 3:98–117, 1997.

[16] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, Nov 1998.

[17] D. Kalkofen, E. Mendez, and D. Schmalstieg. Interactive focus and context visualization for augmented reality. In *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 191–200, Nov. 2007.

[18] D. Kalkofen, E. Mendez, and D. Schmalstieg. Comprehensible visualization for augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 15(2):193–204, 2009.

[19] S. Kastner, P. De Weerd, M. A. Pinsk, M. I. Elizondo, R. Desimone, and L. G. Ungerleider. Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology*, 86(3):1398–1411, 2001.

[20] J. Kruger, J. Schneider, and R. Westermann. Clearview: An interactive context preserving hotspot visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):941–948, 2006.

[21] M. A. Livingston, A. Dey, C. Sandor, and B. H. Thomas. Pursuit of x-ray vision for augmented reality. In *In Human Factors in Augmented Reality Environments*, pages 67–107. Springer, USA, 2013.

[22] E. Mendez, S. Feiner, and D. Schmalstieg. Focus and context in mixed reality by modulating first order salient features. In *Proc. of the 10th international conference on Smart graphics*, SG'10, pages 232–243, Berlin, Heidelberg, 2010. Springer-Verlag.

[23] S. D. Peterson. *Stereoscopic Label Placement: Reducing Distraction and Ambiguity in Visually Cluttered Displays*. Phd, Linköping University, Norrköping, Sweden, 2009.

[24] E. Rosten, G. Reitmayr, and T. Drummond. Real-time Video Annotations for Augmented Reality. In *International Symposium on Visual Computing*, 2005.

[25] C. Sandor, A. Cunningham, A. Dey, and V.-V. Mattila. An augmented reality x-ray system based on visual saliency. In *Proc. ISMAR*, pages 27–36, 2010.

[26] D. D. Seligmann and S. Feiner. Automated generation of intent-based 3d illustrations. In *Proceedings of ACM SIGGRAPH*, pages 123–132, New York, NY, USA, 1991. ACM Press.

[27] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.

[28] E. Veas, E. Mendez, S. Feiner, and D. Schmalstieg. Directing attention and influencing memory with visual saliency modulation. In *Proc. of the international conference on human factors in computing systems*, 2011.

[29] S. Zollmann, D. Kalkofen, E. Mendez, and G. Reitmayr. Image-based ghostings for single layer occlusions in augmented reality. In *Proc. of ISMAR*, pages 19–26, 2010.